

January 2016

A historical and sociolinguistic approach to language change in Mandarin Chinese: Corpus evidence for the development of YOU-MEI-YOU

Wenfeng Li
Purdue University

Follow this and additional works at: https://docs.lib.purdue.edu/open_access_dissertations

Recommended Citation

Li, Wenfeng, "A historical and sociolinguistic approach to language change in Mandarin Chinese: Corpus evidence for the development of YOU-MEI-YOU" (2016). *Open Access Dissertations*. 1255.
https://docs.lib.purdue.edu/open_access_dissertations/1255

This document has been made available through Purdue e-Pubs, a service of the Purdue University Libraries. Please contact epubs@purdue.edu for additional information.

**PURDUE UNIVERSITY
GRADUATE SCHOOL
Thesis/Dissertation Acceptance**

This is to certify that the thesis/dissertation prepared

By Wenfeng Li

Entitled

A HISTORICAL AND SOCIOLINGUISTIC APPROACH TO LANGUAGE CHANGE IN MANDARIN CHINESE: CORPUS EVIDENCE FOR THE DEVELOPMENT OF YOU-MEI-YOU

For the degree of Doctor of Philosophy

Is approved by the final examining committee:

Atsushi Fukada

Co-chair

John Sundquist

Co-chair

Elaine Francis

Felicia Roberts

To the best of my knowledge and as understood by the student in the Thesis/Dissertation Agreement, Publication Delay, and Certification Disclaimer (Graduate School Form 32), this thesis/dissertation adheres to the provisions of Purdue University's "Policy of Integrity in Research" and the use of copyright material.

Approved by Major Professor(s): Atsushi Fukada, John Sundquist

Approved by: Felicia Roberts

Head of the Departmental Graduate Program

7/26/2016

Date

A HISTORICAL AND SOCIOLINGUISTIC APPROACH TO LANGUAGE CHANGE
IN MANDARIN CHINESE: CORPUS EVIDENCE FOR THE DEVELOPMENT OF
YOU-MEI-YOU

A Dissertation

Submitted to the Faculty

of

Purdue University

by

Wenfeng Li

In Partial Fulfillment of the

Requirements of the Degree

of

Doctor of Philosophy

August 2016

Purdue University

West Lafayette, Indiana

For my family and friends

ACKNOWLEDGEMENTS

First, I would like to express my great appreciation to my two co-advisors, Professor Atsushi Fukada and Professor John Sundquist, who have provided thorough and constructive suggestions and comments throughout my dissertation research. I am also greatly indebted to Professor Felicia Roberts for her continuous guidance and support through my PhD years at Purdue and for reading and revising the very first draft of the first two chapters of my dissertation. Many thanks are also owed to Professor Elaine Francis, who has provided valuable feedback on the dissertation and on a related project I conducted for a class I took with her.

Secondly, I would like to offer my special thanks to April Grotberg, who has helped me in many different ways with my dissertation as an editor and friend. Without her help, I would not have been able to meet all the deadlines for my dissertation.

Finally, I feel especially indebted to my family, who have been supportive throughout my dissertation writing. I would also like to thank my friends and colleagues for their kindness, patience and encouragement over the years. The completion of my studies and this work would not have been possible without them.

TABLE OF CONTENTS

	Page
LIST OF TABLES	viii
LIST OF FIGURES	x
LIST OF ABBREVIATIONS.....	xi
ABSTRACT.....	xiii
1. INTRODUCTION	1
1.1 Background on the Construction.....	3
1.2 Theoretical Assumptions about the Origin and Nature of Preverbal YOU-MEI- YOU	6
1.2.1 YOU-MEI-YOU+VP Developed from YOU MEI YOU+NP	7
1.2.2 Alternative Analysis #1: YOU-MEI-YOU+VP Developed from YOU+VP ...	10
1.2.3 Alternative Analysis #2: YOU-MEI-YOU+VP Developed from MEIYOU+VP	13
1.2.4 YOU-MEI-YOU as a Single Unit	17
1.3 RQ1: What Is the Grammatical Status of Preverbal YOU-MEI-YOU?	19
1.3.1 Grammaticalization as Gradual Change	19
1.3.2 Auxiliaries and the Verb-to-TAM Grammaticalization Chain.....	21
1.4 RQ2a: What Stage of Development is Preverbal YOU-MEI-YOU Currently in? .	24

	Page
1.5 RQ2b: What Is the Distribution of Preverbal YOU-MEI-YOU over Time?	29
1.6 Chapter Summary and the Organization of the Dissertation.....	30
2. LITERATURE REVIEW	31
2.1 Previous Research on Preverbal YOU-MEI-YOU.....	31
2.1.1 Grammatical Status of Preverbal YOU-MEI-YOU in Earlier Research.....	36
2.1.2 Development of Preverbal YOU-MEI-YOU: Results and Findings from Previous Research.....	39
2.1.3 Linguistic Analysis of YOU-MEI-YOU+VP in Previous Studies	42
2.2 Grammaticalization: Definition and Measurement	48
2.2.1 What Defines a Particular Grammatical Category?	51
2.2.2 Gradient-based Approaches to Linguistic Categorization.....	54
2.2.3 Characterizing Stages of Grammaticalization: Features and Parameters	61
2.2.4 Frequency of Grammaticalized Forms	64
2.2.5 Grammaticalization of Periphrastic Perfect: Cross-linguistic Evidence	68
2.3 Aspect: A Semantic Account	70
2.3.1 Viewpoint Aspect and Situation Aspect: Expression, Classification and Relation.....	73
2.3.2 Perfective Viewpoints and Perfective Morphemes in Standard Mandarin: GUO, LE, and MEI(YOU).....	77
2.4 Word-class Categorization in Standard Mandarin	83
2.5 Chapter Summary.....	88
3. METHODOLOGY	91

	Page
3.1 Data Collection.....	91
3.1.1 The CCL Database.....	92
3.1.2 The CNC Database.....	94
3.2 Data Processing: Initial Filtering	96
3.3 Data Processing: Second Filtering	98
3.4 Data Analysis	105
3.4.1 Data Analysis to Determine the Formal Status of Preverbal YOU-MEI-YOU	105
3.4.2 Data Analysis to Identify the Developmental Stage of YOU-MEI-YOU+VP114	
3.4.3 Data Analysis to Present the Distribution of Preverbal YOU-MEI-YOU.....	123
3.5 Chapter Summary.....	127
4. RESULTS AND DISCUSSION.....	129
4.1 Status of Preverbal YOU-MEI-YOU in Standard Mandarin	129
4.1.1 The Function of YOU-MEI-YOU in the Marking of Sentence Type	132
4.1.2 The Role of YOU-MEI-YOU in the Expression of Viewpoint Aspect.....	136
4.2 Developmental Stage of Auxiliary YOU-MEI-YOU: Evidence from Diachronic Comparisons.....	143
4.2.1 Desemanticization: A Diachronic Comparison of Preverbal YOU-MEI-YOU	144
4.2.2 Decategorialization: A Diachronic Comparison of the Complements of YOU- MEI-YOU	147

	Page
4.2.3 Types of VP Complement of Auxiliary YOU-MEI-YOU in Different Time Periods	154
4.3 Temporal Distribution of Auxiliary YOU-MEI-YOU	157
4.3.1 Distribution of Auxiliary YOU-MEI-YOU in Mainland China Chinese	161
4.3.2 Distribution of Auxiliary YOU-MEI-YOU in the 2000s	164
4.4 Chapter Summary.....	166
5. CONCLUSION.....	168
5.1 Findings about the Use and Development of YOU-MEI-YOU+VP	168
5.2 Theoretical Implications.....	170
5.3 Limitations	171
5.4 Future Research.....	174
REFERENCES	177
VITA.....	188

LIST OF TABLES

Table	Page
Table 1 Relative chronology of YOU-MEI-YOU, MEI, MEIYOU and YOU.	17
Table 2 Summary of the CCL database.	93
Table 3 Summary of the CNC database.	95
Table 4 Data distribution before and after initial filtering	97
Table 5 Data distribution after second filtering	105
Table 6 Sub-categorization of VP complements	122
Table 7 Distribution of different uses of preverbal YOU-MEI-YOU	131
Table 8 Distribution of interrogative YOU-MEI-YOU	132
Table 9 Distribution of perfective YOU-MEI-YOU	137
Table 10 Distribution of different uses of preverbal YOU-MEI-YOU in different time periods	145
Table 11 Distribution of complements of preverbal YOU-MEI-YOU in different time periods	150

Table	Page
Table 12 Distribution of VP complements of YOU-MEI-YOU in different time periods	155
Table 13 Base size and number of clauses in different time periods.....	158

LIST OF FIGURES

Figure	Page
Figure 1: Relative frequency of use of auxiliary YOU-MEI-YOU over time.....	159
Figure 2: Frequency of use of auxiliary YOU-MEI-YOU in the CCL-CC corpus (all regions)	160
Figure 3: Frequency of use of auxiliary YOU-MEI-YOU in the CCL-CC corpus (Mainland China).....	163

LIST OF ABBREVIATIONS

1	first person
2	second person
3	third person
AC	Ancient Chinese
AP	adjective phrase
AspP	aspect phrase
Aux	auxiliary
CC	Contemporary Chinese
CCL	Center for Chinese Linguistics, Peking University
CLF	classifier
CNC	State Language Commission Chinese National Corpus
IPFV	imperfective
NEG	negative/negation
NP	noun phrase
O	object
OE	Old English
PL	plural
Q	question particle

RVC	resultative verb complement
S	subject
SG	singular
V	verb
VP	verb phrase

ABSTRACT

Li, Wenfeng. Ph.D. Purdue University, August 2016. A Historical and Sociolinguistic Approach to Language Change in Mandarin Chinese: Corpus Evidence for the Development of YOU-MEI-YOU. Major Professors: Atsushi Fukada and John Sundquist.

This dissertation introduces corpus-based analyses of a syntactic construction in Standard Mandarin, YOU-MEI-YOU (or ‘have-not-have’)+VP, which is used to form perfective questions. The purpose of the study is to (i) find evidence for the claim that preverbal YOU-MEI-YOU, i.e. YOU-MEI-YOU found in the new construction, is grammaticalizing into an auxiliary unit, and (ii) to investigate its historical development, including the stage of development that it has reached and its distribution over time. Using data from two databases, the present study first looks at the percentage of preverbal YOU-MEI-YOU conveying a certain grammatical meaning, i.e. sentence type and aspect. Next, the study compares the percentage of three linguistic features of this construction, namely, the grammatical meaning(s) conveyed by preverbal YOU-MEI-YOU, the general types of complement it takes, and the specific types of VP complement it takes, between different 20-year periods. The study also makes a comparison of the frequency of use of preverbal YOU-MEI-YOU between different 10-year periods. The results of the first type of analysis show that preverbal YOU-MEI-YOU helps to form constructions conveying either grammatical meaning in the majority of the clauses,

lending support to the claim that it is grammaticalizing into an auxiliary unit. The diachronic comparisons of the three features of the new construction indicate that preverbal YOU-MEI-YOU has reached Stage III as outlined in Heine (1993). The comparison of the frequency of use between different time periods shows no upward trend in the use of (auxiliary) preverbal YOU-MEI-YOU.

1. INTRODUCTION

This dissertation presents a historical and semantic account of Standard Mandarin YOU-MEI-YOU (or ‘have-not-have’)+VP, a syntactic construction which is used to form perfective questions (often about the completion of an action).^{1,2} One example of this construction is given below:

- (1) ta you-mei-you lai?
 3SG have-not-have come
 Did he come?

I will refer to YOU-MEI-YOU in (1) as preverbal YOU-MEI-YOU.³

A corpus-based study of this construction in literary work was conducted by Kuang (2000), who concluded that the spread of this form has been a recent phenomenon which “occurred in the past two decades” (p. 71), although Dong (2004) notes that it has been marginally attested since the 19th century. Apart from Ding (1961) and Ota (1958/1987), it had not received much scholarly attention until the past twenty years.⁴

¹ In the prose, Chinese pinyin will be represented in capital letters.

² Perfective questions can also be about the termination of an action.

³ If two or more Chinese characters/morphemes form one phrase, the morphemes will not be separated by spaces or hyphens unless they are prosodically separated in the sentence. As for the construction under investigation in this dissertation, i.e. preverbal YOU-MEI-YOU, the morphemes will be separated by hyphens.

⁴ Ota (1958/1987) was one of the earliest books on Chinese grammar that noted the existence of YOU-MEI-YOU+VP. He speculated that this construction had developed from YOU MEI YOU+NP but conducted no investigation into its development. Ding (1961) also made a brief mention of the appearance of YOU-MEI-YOU+VP in Standard Mandarin but offered no details about its use or development.

According to Cheung, Liu, & Shih (1994), this new way of forming perfective questions with preverbal YOU-MEI-YOU is but a dialectal or regional variant. They observed that “[i]n some of the southern dialects, this is indeed an acceptable pattern, and may be used as an alternative to ‘verb+LE+MEIYOU’ (verb + aspect marker LE + negation morpheme MEIYOU)” (p. 236). However, Shi & Li (2001), Xing (1990) and Wang, Wang, & Jiang (2006) document abundant examples of this new question type in the speech of Northern Mandarin speakers and in speech produced by people who supposedly speak Standard Mandarin, including news reporters in China’s Central Television and Northern and Beijing writers and linguists. Therefore, closer attention to the development and grammatical status of this construction is warranted since such research may inform us of the development of Standard Mandarin, a standardized language with its grammatical and usage norms being continuously regulated.

The purpose of the study is to investigate the development of preverbal YOU-MEI-YOU and to answer three central questions: What is the grammatical status of preverbal YOU-MEI-YOU in Standard Mandarin? What is the current development stage of this construction? What is the distribution of this construction over time? Answering these questions will help in assigning this relatively new construction to a proper category, give us a better understanding of its path of development, and provide insight into some relevant processes and factors underlying language change in Standard Mandarin.

1.1 Background on the Construction

The use of preverbal YOU-MEI-YOU, as illustrated in (1), was first discussed in the 1960s (Chao, 1968; Ding, 1961). It is believed to be historically related to questions about the possession or existence of something using the lexical verb YOU ('have', 'there is/exists') in the A-not-A construction (Dong, 2004; Kuang, 2000; Ota, 1958/1987; Shi & Li, 2001; Xue, 2010), with which it shares a very similar surface structure. The only structural difference between these is the complement type, with the VP complement type developing from the NP type through reanalysis and/or analogy. Compare the VP complement in (1), repeated here as (2a), with the use of YOU MEI YOU+NP complement in (2b):

- (2) a. ta you-mei-you lai?
 3SG have-not-have come
 Did he come?
- b. ta you mei you qian?
 3SG have NEG have money
 Does he have money?

While YOU MEI YOU+NP is a well-established, standard usage, traditional grammars of Standard Mandarin dismiss the YOU-MEI-YOU+VP construction as a borrowing from substandard variety of Mandarin (Chao, 1968; Cheung, Liu, & Shih, 1994; Lü, 1985);⁵ it is probably still regarded as characteristic of Mandarin as spoken in

⁵ However, one type of YOU-MEI-YOU+VP structure was listed as an acceptable usage in the *Syllabus for Putonghua Proficiency Test* (Liu, 1998). The sentence listed in Zheng (2001, p. 115; a textbook for Putonghua proficiency test takers modeled on the basis of the syllabus) was:

Southern China. Readers attempting to learn Standard Mandarin were warned not to use it by Cheung, Liu, & Shih (1994), who stated that “[t]here may be a variety of ways of asking the following question in Standard Mandarin, but this is definitely not one of them” (p. 236). For the purpose of comparison, examples (3a-c) illustrate the officially recognized/prescriptive ways of forming perfective questions alongside a semantically equivalent sentence employing YOU-MEI-YOU (3d) that has been rejected by Cheung, Liu, & Shih (1994). Also listed are two examples (3e, f) showing the conventional ways of forming perfective statements in Standard Mandarin.

(3) a. ta lai-le ma?

3SG come-LE⁶ Q

Did he come?

b. ta lai-le meiyou?

3SG come-LE NEG

Did he come?

c. ta lai mei lai?

3SG come NEG come

Did he come?

(a) ni you-mei-you chi-guo fan?

2SG have-not-have eat-RVC meal

Have you eaten?

This is given in contrast to the non-standard, unacceptable

(b) *ni you chi-guo fan meiyou?

2SG have eat-RVC meal NEG

⁶ The grammatical morphemes in Standard Mandarin (e.g. perfective LE, and perfective GUO) are written in capital letters in the glosses.

- d. ta you-mei-you lai?
 3SG have-not-have come
 Did he come?
- e. ta lai-le.
 3SG come-LE
 He came.
- f. ta mei(you) lai.
 3SG aux.NEG come
 He did not come.

As shown in example sentences (3a-d), the standard ways of forming perfective questions either employ a different morpheme, LE, which acts like a suffix following the main verb (as in 3a, 3b), or require a special construction of the main verb (or the first morpheme of a verb in most cases), namely the A-not-A construction (as in 3c).⁷ In contrast, in the construction under study, YOU-MEI-YOU precedes the main verb (as in 3d).

All the perfective questions listed above (3a-d) elicit the same response, which uses the same morphemes found in perfective statements (3e, f): an answer in the affirmative consists of the (first morpheme of the) main verb and the morpheme LE, and an answer in the negative is a simple MEI/MEIYOU.⁸ This is illustrated in example sentences (4a) and (4b), respectively.

⁷ For example, for verbs with two morphemes like JIEHUN ‘to marry’, the perfective question using the A-not-A construction would be: *tamen jie mei jiehun?*

⁸ It can also be the combination of a negative morpheme MEI and the first morpheme of the main verb or the whole main verb.

(4) a. lai-le.

come-LE

He came. / He has come.

b. meiyou.

aux.NEG

He did not/hasn't.

The morpheme LE in (3e, 4a) is a common aspect marker in Standard Mandarin and has a well-established status in both prescriptive and descriptive grammars. The status of MEI/MEIYOU as an aspect marker or an auxiliary, as used in sentences like (3f, 4b), has also long been established in Chinese (see Chao, 1968; Cheng, Huang, & Tang, 1997; Ernst, 1995; Huang, Li, & Li, 2009; Lü, 1985; Ma, 2010; Shi, 2002, among others). A detailed discussion of the aspectual meanings conveyed by these morphemes and the differences in their functions and meanings will be carried out in the next chapter.

1.2 Theoretical Assumptions about the Origin and Nature of Preverbal YOU-MEI-YOU

This section will lay out the two assumptions that underlie the analysis. The first is that YOU-MEI-YOU+VP grammaticalized from YOU MEI YOU+NP constructions. Section 1.2.1 lays out several positions that favor these endpoints, though the specific path from the one construction to the other varies by author. Section 1.2.2 and 1.2.3 address two possible alternatives to this position—that YOU-MEI-YOU+VP derives from an auxiliary YOU+VP and that it derives from MEIYOU+VP—and argues against the plausibility of each one. The second major assumption of this paper is that YOU-

MEI-YOU in its grammaticalized state (i.e., when used with a VP complement) behaves as a single unit. This will be discussed in Section 1.2.4.

1.2.1 YOU-MEI-YOU+VP Developed from YOU MEI YOU+NP

As mentioned earlier, previous studies on YOU-MEI-YOU+VP have argued that this construction developed from the YOU MEI YOU+NP structure and that preverbal YOU-MEI-YOU is a case of grammaticalization (Dong, 2004; Shi & Li, 2001). For instance, Shi & Li (2001) argued that the emergence of YOU-MEI-YOU+VP in Standard Mandarin has been made possible by the appearance of YOU MEI YOU+NP and the existence of MEIYOU ('lack', 'not exist'), which was first used as a lexical item and later used before VPs as a negator in Chinese in the 16th century. Drawing on Ota (1958/1987), they reasoned that YOU-MEI-YOU grammaticalized into a preverbal unit through the mechanism of analogy: since MEIYOU was found to precede NPs and VPs, and YOU MEI YOU, which was functionally similar to the former, was also found to precede NPs, it seemed natural for YOU MEI YOU to start appearing before VPs. Shi & Li (2001) also attempted to reconstruct the path of development of YOU-MEI-YOU+VP and speculated that the intermediate stage consisted of sentences in which the NP following YOU MEI YOU was modified by a VP (i.e. YOU MEI YOU+NP+VP). They hypothesized that the YOU-MEI-YOU+VP surface structure resulted from preposing the NP or eliding it from the YOU MEI YOU+NP+VP structure. The structure Shi & Li proposed as the intermediate stage between YOU MEI YOU+NP and YOU-MEI-

YOU+VP is similar to the *habban* ‘have’+NP+Past Participle VP structure in Old English, from which the auxiliary *have* in English has developed.⁹

Dong (2004) introduced a competing hypothesis about the grammaticalization path of YOU-MEI-YOU+VP. She proposed that the intermediate stage between YOU MEI YOU+NP and YOU-MEI-YOU+VP was clauses in which referential VPs precede YOU-MEI-YOU. According to Dong, referential VPs denote (a series of) events, and they appear before a separate clause formed by YOU-MEI-YOU (with no subject or object). Such clauses are questions about the occurrence of the event(s). In this context, it is difficult to tell whether the VPs are the preposed complement of YOU-MEI-YOU or whether the original clause contains an NP functioning as the complement of YOU MEI YOU that is omitted. Dong illustrated her hypothesis with examples taken from ancient novels, one of which is listed below:

- (5) *lianjinban de erxi, shuo ni hen teng ta, gei ta haoxie*
 Lianjinban DE¹⁰ Erxi, say 2SG very be.fond.of 3SG, give 3SG much
 dongxi, zai ni jia zhu-le yi ye, you mei you?
 thing, at 2SG home live-LE one night, have NEG have/have-not-have
 Erxi from Lianjinban said that you were very fond of him, gave him a lot of stuff,
 and he stayed in your house for one night, did this happen? /Did you do what he
 had said? /Did he really stay one night in your house?

One interpretation of the sentence above is that an NP complement of YOU MEI YOU, e.g. ZHEXIESHI (‘these thing’), is omitted. In another interpretation, the subject of

⁹ See Section 2.2.5 for a detailed description of the grammaticalization of verbs meaning possession in this structure.

¹⁰ The morpheme, DE (的) in this sentence is used to mark possession, and is similar to the English apostrophe-*s* (’s) in terms of function.

YOU-MEI-YOU is NI 2SG or TA 3SG and the complement of YOU-MEI-YOU is the corresponding VP, which is mentioned earlier in the discourse and omitted to avoid repetition in this particular clause.

Another possible path whereby YOU-MEI-YOU+VP developed from YOU MEI YOU+NP is through YOU-MEI-YOU followed by a phrase that is ambiguous in interpretation between an NP and a VP. Some examples of this nature were found in the data collected for this study, one of which is presented here for illustration:^{11,12}

(6) ta de zuoyong you-mei-you / you mei you gaibian ne?

3SG DE function have-not-have / have NEG have change NE

Did its function change? / Is there any change in its function?

In example (6), GAIBIAN can be interpreted as a VP meaning ‘to change’. It can also be interpreted as an NP: the quantifier YIXIE ‘some’ can be added to modify it without changing the core meaning of the sentence. The development of YOU-MEI-YOU+VP from YOU MEI YOU+NP could also have been made possible through such ambiguous sentences.

The relative chronology of YOU-MEI-YOU+VP and YOU MEI YOU+NP is also part of the basis on which previous studies have adopted the assumption that YOU-MEI-YOU+VP developed from YOU MEI YOU+NP (e.g. Dong, 2004; Ota, 1958/1987; Shi & Li, 2001). Shi & Li (2001) showed that YOU MEI YOU+NP occurred earlier than YOU-MEI-YOU+VP and was preceded by other more common ways of asking questions about the possession or existence using the lexical verb YOU ‘have’, ‘there is/exists’ (e.g.

¹¹ See Section 4.2.2 in Chapter 4 for a detailed discussion of such ambiguous cases.

¹² The morpheme, DE (的), in this sentence has the same function as it does in (5). The morpheme, NE (呢), is a sentence final particle, often used in an A-not-A questions.

YOU+NP+MEIYOU, YOU+NP+MEIYOU+NP). The use of YOU MEI YOU+NP was found to be frequent in the 18th century, which preceded the discovery of isolated examples of YOU-MEI-YOU+VP in the early 19th century.

1.2.2 Alternative Analysis #1: YOU-MEI-YOU+VP Developed from YOU+VP

One alternative hypothesis about the origin of YOU-MEI-YOU+VP is that it developed from YOU+VP. At first glance, this hypothesis seems plausible, since the application of the A-not-A rule in Chinese to an auxiliary YOU could yield the target construction. However, “Standard Mandarin does not have the YOU+VP structure” (Shi & Li, 2001, p. 268) and therefore Shi & Li as well as Dong (2004) reject this analysis.

To begin with, according to Shi & Li (2001), preverbal YOU-MEI-YOU bears surface resemblance to the A-not-A construction of a modal in Standard Mandarin—e.g. NENG ‘be able to’, HUI ‘can’—but is different from the latter in a fundamental way. While questions containing the A-not-A construction of a modal can be answered using the Modal+VP structure, given the lack of a YOU+VP structure, YOU-MEI-YOU+VP questions cannot be answered in a syntactically analogous way to the Modal+VP option. Along similar lines, Dong (2004) stated that “a positive counterpart, i.e. YOU, does not exist for the (auxiliary) preverbal YOU-MEI-YOU and that YOU could not be used as an auxiliary in Standard Mandarin” (p. 1). Instead, YOU-MEI-YOU+VP questions are answered with VP+LE (ex. 4a), with the perfective LE acting as the positive counterpart

of YOU-MEI-YOU in these questions.¹³ Hence, “preverbal YOU-MEI-YOU can only be included in the dictionary as an inseparable item” (Dong, 2004, p. 1).

Furthermore, if YOU-MEI-YOU+VP had been an interrogative form of YOU+VP, they should have developed in a parallel fashion. However, this is not the case. Findings from previous studies on the use of YOU+VP in certain dialects or varieties of Mandarin have shown that YOU-MEI-YOU+VP developed earlier than YOU+VP in Standard Mandarin spoken in Mainland China.

The existence of an auxiliary YOU or use of YOU+VP in Mainland Standard Mandarin was found to have only occurred in recent years, with only anecdotal or patchy evidence for this use. Wang (2012) reports that the use of YOU+VP by Standard Mandarin speakers from outside the Cantonese, Min or Hakka areas started in the 1980s.¹⁴ In his extensive survey of contemporary Standard Mandarin use, Lü (2000) pointed out that the use of YOU+VP specifically as an answer to YOU-MEI-YOU+VP questions was quite uncommon. Some other studies pointed out that YOU+VP appeared more frequently in Mainland Standard Mandarin beginning in the mid-1990s (Cai, 2009) or in the early 2000s (Lu, 2010), with more examples of YOU+VP being found in recent years (Wang, Wang, & Jiang, 2004). Cai (2009) conducted a survey on the attitude toward and use of YOU+VP among college students in Shanghai and found a negative correlation between age and the level of acceptance and use of YOU+VP (p. 85).

Using newspapers and magazines published in 2011, along with contemporary novels and other literary works (published in recent decades), Diao (2012) investigated

¹³ A more detailed discussion of the perfective morpheme, LE, can be found in Chapter 2.

¹⁴ According to Wang (2012), the use of YOU+VP was a characteristic of Standard Mandarin spoken by native speaker of Cantonese, Min or Hakka, and was introduced into Standard Mandarin by these speakers.

the use of YOU+VP in four different regions. He concluded that YOU+VP was quite rare in Mainland Standard Mandarin: it was attested in writing but just barely, especially compared with YOU-MEI-YOU+VP, which he found was less developed than MEIYOU+VP but more common than YOU+VP in his data. He also claimed that, compared with the use of YOU+VP in Hong Kong, Taiwan and Macau, the use and distribution of YOU+VP in Mainland Standard Mandarin were rather simple and seemed artificial to a certain extent: “it seemed that this structure was introduced into Mainland Standard Mandarin by choice or developed with an intention in order to complete a link” (p. 48), as illustrated below:

Interrogative		Statement
YOU-MEI-YOU+VP	Negative:	MEIYOU+VP
have-not-have+VP		aux.NEG +VP
	Affirmative:	YOU+VP
		aux+VP

In contrast to the late appearance of YOU+VP, the first documented use of the YOU-MEI-YOU+VP structure was much earlier (1808) and abundant examples of this structure have been documented in Mainland Standard Mandarin (Ding, 1961; Ota, 1958/1987; Shi & Li, 2001; Xing, 1990; Wang, Wang, & Jiang, 2006, to name just a few). YOU-MEI-YOU+VP became a common usage in the 1980s (Xing, 1990), and “has become more or less established in its development” (Dong, 2004, p. 3), a statement supported by the findings from studies using large corpora (Wang, Wang & Jiang, 2006; Shi & Li, 2001). Based on the findings above, I agree with previous scholars that

preverbal YOU-MEI-YOU cannot be analyzed as the A-not-A construction of an auxiliary YOU in Mainland Standard Mandarin.

1.2.3 Alternative Analysis #2: YOU-MEI-YOU+VP Developed from MEIYOU+VP

Another speculation about the origin of preverbal YOU-MEI-YOU is that the negative MEIYOU (which marks perfectivity) is the source, as it can be analyzed as MEI and YOU and can form YOU-MEI-YOU if one adds YOU before it. This possibility can, however, be ruled out for the following reasons: first, MEI and YOU fused into a single lexical item around the 12th century, meaning ‘lack’ or ‘not exist’. MEIYOU later developed into a negator, which was used before VPs. Preverbal MEIYOU is a functional head and cannot be separated. This can be shown by applying syntactic tests. In the following sentences, the same test is applied to a compound lexical verb YOUYONG ‘swim’ (7a-b) and to MEIYOU ‘lack’, ‘not exist’ (8a-b):

(7) a. ta zaoshang youyong.

3SG morning swim

He swims in the morning.

b. ta zaoshang you-guo yong (le).

3SG morning swim-GUO swim (LE)¹⁵

He already swam in the morning.

(8) a. ta meiyou zhe zhong jingli.

3SG not.have this kind experience

He does not this kind of experience.

¹⁵ The morpheme LE is used as a sentence-final particle, indicating current relevance of the state of affair. This is a different from the perfective morpheme LE.

b. ta meiyou-guo zhe zhong jingli.

3SG not.have-GUO this kind experience

He has/had never had this kind of experience.

Note that the aspect marker GUO is inserted between YOU and YONG, although the two morphemes form a single lexical verb meaning ‘swim’. In contrast, the aspect marker is attached to MEIYOU and cannot separate the (historically) two morphemes in MEIYOU. Similarly, when MEIYOU is used before a VP as a negator, as in (9a), the aspect marker is attached to the first morpheme of the main verb, as in (9b):

(9) a. ta zaoshang meiyou youyong.

3SG morning aux.NEG swim.

He did not swim in the morning.

b. ta zaoshang meiyou you-guo yong.

3SG morning aux.NEG swim-GUO swim

He never swims/swam in the morning.

To summarize, the inseparability of MEIYOU used as both a lexical and a function verb made it an unlikely source of preverbal YOU-MEI-YOU. If the A-not-A rule were applied to preverbal MEIYOU, one would not be able to derive the target construction (i.e. preverbal YOU-MEI-YOU).

Another point I want to make about MEIYOU is that there is a great interval of time between the use of the MEIYOU+VP structure and the first appearance of YOU-MEI-YOU+VP. In order to get a fuller picture of the development of MEIYOU, it is important to also look at the development of MEI, a word that is closely related to the emergence and development of MEIYOU.

In ancient Chinese, before the 8th century or earlier, MEI was used as a lexical verb meaning ‘sink’ or ‘bury’ and YOU was used as a verb meaning ‘have’ or ‘there is/exists’.¹⁶ According to Shi & Li (2001), in around the 8th century, MEI underwent semantic shift and developed the new meanings ‘lack’ or ‘not exist’. Examples of MEI ‘not exist’ taken from poems written in the 8th century were given in Huang (2015). Shi & Li also claimed that MEI and YOU fused into a lexical item meaning ‘lack’ or ‘not exist’ in the 14th -15th century, but Shen (2007) does list one example of MEIYOU+NP from a story written in the 12th-13th century. Huang (2015) also finds examples from this earlier time period, though the usage was quite rare. Both MEI and MEIYOU underwent subsequent grammatical change and came to be used as functional heads with VP complements, but their lexical meanings persist and they continue to be used with NP complements.

With respect to the time when MEI and MEIYOU started to be used before VPs, scholars also hold different opinions. According to Shi & Li (2001), MEI grammaticalized into a functional head in the 15th century and the 16th century saw more consistent appearance of MEI+VP. Nevertheless, other studies have found earlier instances of MEI+VP. For example, Wu (1995) and Yang (1999) argued that MEI+VP started in the second half of the Song Dynasty—i.e. 12th-13th century—and listed examples of MEI+VP from documents written in that period. Huang (2015), Liu (2010) and Xu (2003) noted that MEI+VP was found in texts written between the Tang Dynasty (7th-10th century) and the Song Dynasty (10th-13th century) but it was not widely used

¹⁶ YOU in ancient texts may have other meanings but for the purpose of the current discussion, only two meanings are listed.

until between the Yuan Dynasty (13th-14th century) and Ming Dynasty (14th-17th century) (Huang, 2015), or around the 15th century (Liu, 2010).

As for MEIYOU+VP, Shi & Li (2001) mentioned that this usage started after the 16th century. Huang (2015) and Xu (2003) claimed that MEIYOU+VP could be found in the Song Dynasty (10th-13th century). However, Huang (2015) gave no examples of such usage and Xu (2003) quoted one sentence from the 12th-13th century, but the meaning of MEIYOU in that particular sentence remains to be settled. Liu (2010) listed sentences from texts written in the 13th-14th century in which MEIYOU was used unambiguously as a negator of verbs (e.g. MEIYOU CHIFAN ‘not eat’; MEIYOU JIAN ‘not see’).

It should be borne in mind that the earliest instance of preverbal YOU-MEI-YOU was found in the early 19th century, after the appearance of YOU MEI YOU+NP. The interval in development between these YOU-MEI-YOU+VP and MEIYOU+VP is considerable. It cannot be used as evidence for arguing against the claim that preverbal YOU-MEI-YOU developed from preverbal MEIYOU, but it points to the unlikelihood of preverbal MEIYOU being the source of preverbal YOU-MEI-YOU. A timeline for the development of YOU-MEI-YOU, MEI, MEIYOU and YOU is laid out in Table 1:¹⁷

To sum up, based on the above findings, it is reasonable to claim that YOU-MEI-YOU+VP did not develop from MEIYOU+VP or YOU+VP. The existence of MEIYOU+VP might have served as the basis of analogous thinking that led to the development of YOU MEI YOU+NP to preverbal YOU-MEI-YOU, given the fact that both MEIYOU+NP and MEIYOU+VP were possible in Chinese. As for YOU+VP, various researchers speculated that it could eventually enter Standard Mandarin grammar,

¹⁷ The timeline for the usage of YOU is also listed here for comparison.

Table 1 Relative chronology of YOU-MEI-YOU, MEI, MEIYOU and YOU

	MEI		MEIYOU		YOU		YOU-MEI-YOU
???	verb: 'sink'		---		verb: 'have/there exists'	---	---
8 th c.	verb: 'lack', 'not exist'		---			---	---
12 th -13 th c.	'lack', 'not exist'	NEG	verb: 'lack', 'not exist'			---	
13 th -14 th c.			'lack', 'not exist'	NEG		---	
early 18 th c.			---			---	YMY + NP
early 19 th c.			---			---	YMY + VP
early 20 th c.			---		---		
1980s				aux + VP			

based on the development of YOU-MEI-YOU+VP (Liu, 2008; Lu, 2010; Shi & Li, 2001; Sun, 2003). Following Dong (2004), Shi & Li (2001), Liu (2008), and Xue (2010) and others, I will argue that YOU-MEI-YOU+VP has developed from YOU MEI YOU+NP and is best studied in the framework of grammaticalization (e.g. Heine, 1993; Hopper & Traugott, 2003; Lehmann, 2002 [1982]).

1.2.4 YOU-MEI-YOU as a Single Unit

The other assumption about preverbal YOU-MEI-YOU adopted in this dissertation is that it forms one unit. This assumption is based on the reasoning in previous studies about the functions and special status of preverbal YOU-MEI-YOU in the Chinese grammars (Chao, 1968; Dong, 2004; Li & Tang, 1991). This assumption also follows from the argument that preverbal YOU-MEI-YOU developed from YOU MEI YOU+NP. As Bybee (2003) states, during the process of grammaticalization, “a frequently used sequence of words or morphemes becomes automated as a single

processing unit” (p. 603).¹⁸ The description of grammaticizing constructions fits well with the development of preverbal YOU-MEI-YOU, which was derived from the sequence of YOU MEI YOU or the A-not-A construction of the lexical verb YOU ‘have’, ‘there is/exists’. The following sentence illustrate the inseparability of preverbal YOU-MEI-YOU and how it differs from the A-not-A construction of the lexical verb YOU:

(10) a. ta you mei you zhe zhong jingli?

3SG have NEG have this kind experience

Does he have this kind of experience?

b. ta you(-guo) mei you-guo zhe zhong jingli?

3SG have(-GUO) NEG have-GUO this kind experience

Did he ever have this kind of experience?

(11) a. ta you-mei-you jingli zhe zhong qingkuang?

3SG have-not-have experience this kind circumstance

Did he experience this kind of thing?

b. ta you-mei-you jingli-guo zhe zhong qingkuang?

3SG have-not-have experience-GUO this kind circumstance

Did he ever experience this kind of thing?

As shown in examples (10b), the aspect marker GUO is attached to the main verb YOU ‘have’, ‘there is/exists’ and the reduplicated main verb with the aspect marker can also appear in the A-not-A construction. In contrast, the aspect marker GUO cannot be attached to the preverbal YOU-MEI-YOU and has to be attached to the main verb JINGLI ‘to experience’. The fact that the aspect marker GUO has this behavior with

¹⁸ The term *grammaticization* is used as a synonym of *grammaticalization* in Bybee (2003).

respect to preverbal YOU-MEI-YOU in (11) suggests that preverbal YOU-MEI-YOU functions as one unit.

In the following sections, I will discuss my three research questions about preverbal YOU-MEI-YOU and discuss how they are motivated by findings and theories from grammaticalization studies.

1.3 RQ1: What Is the Grammatical Status of Preverbal YOU-MEI-YOU?

To motivate this question, I first define grammaticalization and the gradual nature of the process, overviewing several approaches to the notion of chaining. This lays the basis for my argument to use Heine’s (1993) approach to defining auxiliaries and ultimately motivates my claim that preverbal YOU-MEI-YOU is grammaticalizing into an auxiliary unit in Standard Mandarin.

1.3.1 Grammaticalization as Gradual Change

The term *grammaticalization* was coined by the French linguist Antoine Meillet to describe a process whereby new grammatical forms are created (1912, p. 131).¹⁹ The “newness” refers to both the new ways of expressing existing grammatical meanings or relations (e.g. new case endings derived from prepositions) and categories that are not expressed in a language prior to this process (e.g. a new tense category acquired by a language) (Kiparsky, 2012, p. 15). When free/autonomous words are grammaticalized and “take on the role of grammatical forms” (p. 133), they can be seen as grammatical

¹⁹ According to Meillet (1912), there are two processes whereby new grammatical forms are formed—analogy and what he called “grammaticalization” later in the same article: “Ce deux procédés, l’innovation analogique et l’attribution du caractère grammatical à un mot jadis autonome, sont les seuls par lesquels se consistent des formes grammaticales nouvelles” (p. 131).

elements in this use, even though they remain free words in other uses (e.g. the French verb *être* ‘to be’, in Meillet, 1912, p. 131). Put in another way, “when a content word assumes the grammatical characteristics of a function word”, it is said to be grammaticalized (Hopper & Traugott, 2003, p. 4).

It is reasonable to argue that when certain items undergo grammaticalization and start to assume grammatical functions, it becomes difficult to decide if they should still be seen as members of the original categories to they belonged prior to the grammaticalization. According to Lehmann (2002 [1982]), “in the course of such grammaticalization processes, there must be a point of shift” (p. 151), beyond which the grammaticalized items can be seen as members of a different syntactic category—or “syntactic reanalysis” in Lehmann’s words. However, such binary decisions are hard, and it has to be acknowledged that there are gradual transitions (p. 151), since “grammaticalization is a process of gradual change” and “its products may have different degrees of grammaticality” (p. 22).²⁰

Lehmann (2002 [1982]) used the term “grammaticalization scale” to describe the relationship between the “functionally similar signs types” with different degrees of grammaticality “as measured by certain parameters” (p. 33) (cf. the concept of a “cline” in Hopper & Traugott, 2003). It is clear from the examples discussed in Lehmann (2002 [1982]) that such categories are historically related or that a category found at one end of the scale is historically earlier or less grammaticalized than another category closer to the opposite end of the scale. Therefore, the difference between “two grammatical categories

²⁰ Following Lehmann (2002 [1982]), the word *grammaticality* is used here to refer to the degree of grammaticalization which an element has reached (p. 19). The same term has been commonly used to mean the degree to which something conforms to the rules of grammar.

connected on a grammaticalization scale” is gradual and “there is no clear-cut dividing line” between them (p. 40). This is a point Lehmann made about the distinction between main verbs and auxiliaries that have evolved from the former, which also applies to the distinction between other categories connected on a grammaticalization scale.

1.3.2 Auxiliaries and the Verb-to-TAM Grammaticalization Chain

Similarly, in his book on the genesis and development of auxiliaries, Heine (1993) pointed out the limitations of linguistic categorization based on necessary and sufficient conditions for membership. It was argued that such traditional taxonomies on the basis of discrete categories could not provide a satisfactory description of or explanation for such phenomena and problems as gradience, linguistic ambiguity and polysemy, etc. For instance, “the transitional nature of auxiliary-main verb relations” (p. 80) could not be adequately accounted for in the framework of classical generative grammar built on discrete categorization, as complained by Reis (1976, cited in Heine, 1993, p. 80).

Drawing on the findings in cognitive linguistics and grammaticalization, Heine (1993) put forward the Overlap Model, drawing on such notions as continuum (Garcia, 1967) and gradience (Bolinger, 1980) in his depiction of the relations between linguistic categories that are (historically) connected by the process of grammaticalization. In this model, the grammaticalization process of a linguistic entity was conceptualized as a chain, which is comprised of successive, intermediate links corresponding to “the different stages of the lexical-to-grammatical development” of the linguistic entity (Kuteva, 2001,

p. 10).²¹ Auxiliation, the development of auxiliaries out of lexical verbs as originally defined by Benveniste (1968, cited in Kuteva, 2001, p. 1), is seen as a grammaticalization chain in this work.²² Heine focused on what was referred to as the Verb-to-TAM auxiliation chain, which consists of “a verbal structure at one end and a grammatical marker of tense, aspect, modality, etc. at the other” (p. 53). The notion of grammaticalization chains, especially that of Verb-to-TAM (Tense, Aspect, Modality) chain, is of special relevance to the present study as it lays the foundation for addressing the first issue of interest, i.e. the categorial status of preverbal YOU-MEI-YOU.

Although it is difficult to pinpoint stages along such grammaticalization chains as discussed earlier, it can be stated safely that in the case of the Verb-to-TAM chain, the starting point is more or less easily identifiable, as “[t]he moment a verb is given an infinitive complement, that verb starts down the road of auxiliariness” (Bolinger, 1980, p. 297).²³ The verbhood of the construction under investigation, preverbal YOU-MEI-YOU, is difficult to test. However, similar to modal verbs in Standard Mandarin, it behaves differently from preverbal constituents (e.g. adverbs) and regular verbs.²⁴ It also exhibits other patterns of use (cf. 3d) that are congruent with the notion of Verb-to-TAM chain (Heine, 1993; Kuteva, 2001) and with the recent functional definition of auxiliaries proffered by Krug (2011), which built on and expanded Heine’s definition.²⁵ Therefore,

²¹ A more detailed description of the characteristics of a grammaticalization chain is offered in the next chapter.

²² See Heine (1993) and Kuteva (2001) for a refined/revised definition of auxiliation.

²³ Although, as Bolinger (1980) added, “[i]t may make no more than a start or travel all the way” (p. 297).

²⁴ Tests showing that preverbal YOU-MEI-YOU has auxiliary or verb-like characteristics are in Section 2.4 in Chapter 2.

²⁵ For Heine (1993), an auxiliary occupies “some range of uses along the Verb-to TAM [Tense, Aspect, Modality] chain” (p. 70). According to Krug (2011), an auxiliary developed from a verb and takes a verbal complement “which is less than fully finite” (p. 558).

the analysis of the preverbal YOU-MEI-YOU in terms of categorial status will be grounded on the definitions in Heine (1993) and Krug (2011).

Heine's concept of auxiliiation chains is of particular relevance to the analysis of the categorial membership of preverbal YOU-MEI-YOU. Also relevant to such an analysis is the idea that "the fate of a category in grammaticalization is largely predetermined once we know two things: 1) its meaning and 2) its syntactic function" (Lehmann, 2002 [1982], p. 33), or what is called a grammaticalization channel. The channels are illustrated with examples taken from various languages in Lehmann (2002 [1982]). The extent to which this notion is consistent with historically attested facts and new discoveries is beyond the scope of my investigation, but suffice it to say that Lehmann's discussion serves as a starting point for proposing a claim that will later be tested.

Since preverbal YOU-MEI-YOU is being used in perfective questions in which it takes VP complements, based on the previous observations and analysis about grammaticalization, auxiliiation and auxiliaries sketched above, the research presented here will show that preverbal YOU-MEI-YOU is grammaticalizing into an auxiliary unit in Standard Mandarin. The current study will follow Heine's (1993) approach to defining auxiliaries while also taking into account the way auxiliaries are characterized and/or defined in more recent work (e.g. Kuteva, 2001; Krug, 2011). In the analysis of data intended to produce evidence for this claim, the current study will also consider studies on aspect (e.g. Comrie, 1976; de Swart, 2012; Filip, 2012; Smith, 1997), a construct that bears on defining auxiliaries in the studies mentioned above.

In addition, the present study will consider the fact that Standard Mandarin is a language lacking morphological inflections, which means it is difficult to classify words in the first place. Therefore, it is important to look at and draw on methods and theories proposed for linguistic categorization of Standard Mandarin in prominent works (Guo, 2002; Li & Thompson, 1989; Yuan, 2010 and others). It should also be borne in mind that the ambiguity that is a natural consequence of grammaticalization may further complicate the analysis.

In summary, the current study will look at whether and to what extent preverbal YOU-MEI-YOU conveys a certain grammatical meaning in the data. If the construction is found to convey some grammatical meaning in the majority of the data, it will constitute evidence that substantiates the claim that the construction is grammaticalizing into an auxiliary unit. The definition and operationalization of grammatical meanings will be discussed in more detail in the next chapter.

1.4 RQ2a: What Stage of Development is Preverbal YOU-MEI-YOU Currently in?

Having looked at approaches to categorizing a linguistic unit undergoing grammaticalization, I now turn to an equally important issue, the question about its development stages and the pertinent parameters for positing and analyzing such stages.²⁶ I will be working with a dynamic perspective that takes into account both the synchronic and historical dimensions of grammaticalization (e.g. Bybee, Perkins, & Pagliuca, 1994; Heine, 1993; Hopper & Traugott, 2003; Kuteva, 2001; Lehmann, 2002 [1982]; Nichols &

²⁶ The attempt to define stages or focal points of development along a grammaticalization chain is merely meant for descriptive convenience, as are the notions used to characterize some salient stages, as explained by Heine (1993, pp. 54, 57).

Timberlake, 1991; Traugott & Heine, 1991); therefore, the historical trajectory of YOU-MEI-YOU is also an important facet of understanding its present behavior.

Tracing the change in various aspects of the linguistic form being grammaticalized, or the stages of development of the grammaticalized item, has long been of central concern for those working in grammaticalization. Efforts have been made to identify the linguistics areas affected by grammaticalization (e.g. morphology, phonology, syntax, semantics, etc.) and to investigate the shifts in these areas as well as how the shifts correlate with the degree/advancement of grammaticalization. While some prominent studies on grammaticalization focused on morphosyntax and morphophonology (e.g. Heine & Reh, 1984; Lehmann, 2002 [1982]), changes in meaning, i.e. in semantics and pragmatics, have gained increasing research attention since Givón's (1979) work on morphologization and syntacticization (synonymous with grammaticalization) in creoles proposed a grammaticalization path that started at the discourse level (e.g. Bybee, 1985; Bybee, Perkins, & Pagliuca, 1994; Heine, Claudi, & Hünemeyer, 1991; Heine, 1993).

With respect to the development of grammaticalization, some researchers have laid out or attempted to determine a set of relevant parameters to measure degrees of grammaticalization. For instance, Lehmann (2002 [1982]) proposed six parameters which were primarily concerned with the morphosyntax of the grammaticalized element, but the correctness of some has been questioned (see Hopper & Traugott, 2003) and it is difficult to operationalize some parameters, as Lehmann himself admitted (pp. 124, 144); I will discuss the problems with these parameters in the second chapter. It is important to note, nevertheless, that the challenges associated with the application of Lehmann's parameters

well exemplify the difficulty in deciding on cross-linguistically relevant, easy-to-operationalize parameters for the aforementioned purpose.

The parameters against which the data will be evaluated in this dissertation are desemanticization and decategorialization proposed in Heine (1993). They are considered two chains of grammaticalization and the stages along each of the chains are summarized below.

Three stages were proposed for the desemanticization chain, which are mainly differentiated by the types of concept (lexical versus grammatical) conveyed by the grammaticalized item, and the type of subject it has. During the initial stage, “the subject is typically human” (p. 54), and “the verb expresses a lexical concept” (p. 54). The second stage is characterized by a complement (of the grammaticalized verb) expressing a dynamic situation. During the third stage, “the subject is no longer associated with willful/human referents, and the verb acquires a grammatical function” (p. 54). The analysis in the present study will only focus on the type of meaning conveyed by YOU-MEI-YOU as Mandarin Chinese allows the omission of the subject in a sentence. If YOU-MEI-YOU in the new construction conveys a grammatical function in the majority of the data, it constitutes evidence that it has reached Stage III.

With respect to the decategorialization chain, five stages were listed in Heine (1993). The development of auxiliaries is mainly characterized by the loss of verbal properties of the grammaticalized item and the morphosyntax of the complement it takes. During the initial stage, “the verb exhibits a fully verbal morphosyntax” (p. 55), and an

adverbial or a noun phrase forms the nucleus of its complement.²⁷ As it proceeds to the second stage, “[i]nstead of a noun, the complement consists of a nominalized/finite verb” (p. 55). When it reaches Stage III, it “loses its verbal properties” (p. 55) and “it may no longer have a noun as its complement nucleus” (p. 55). The fourth stage is characterized by the further loss of verbal properties of the grammaticalized item, as well as the loss of “nominal (and adverbial) properties” (p. 55) of the complement. At Stage V, “[t]he verb loses virtually all remaining verbal properties, and the complement acquires the morphosyntax of a main verb” (p. 55).

It is clear that one of the distinguishing features of the five stages along the decategorialization chain—i.e. the complement the grammaticalized item takes, hinges on the notion of morphological inflections. This makes it difficult to apply this feature in the analysis of Mandarin Chinese data, as Mandarin Chinese is morphologically impoverished. The description about the complement in Heine (1993) may be useful in identifying the developmental stage that YOU-MEI-YOU has reached, if the focus is on what types of complement (noun phrase versus verb phrase) YOU-MEI-YOU takes, instead of on the morphology of the complement. Therefore, if YOU-MEI-YOU in this new construction is found to take VPs as complements in the majority of the data, it can be seen as evidence that it has reached at least Stage III.

The other distinguishing feature, the verbhood of the grammaticalized item, is also subsumed under the description/definition of auxiliaries in Heine (1993) and Krug (2011). However, it is also difficult to apply in the current analysis, because, as

²⁷ The definition of “nucleus” is associated with the morphology of the complement. As Heine (1993) explains, “in cases where the complement is an adverbial phrase rather than a noun phrase, some adverbial morphology is involved in addition” (p. 55).

mentioned earlier, the structure of YOU-MEI-YOU does not allow many conventional tests of verbhood. As it is claimed in this dissertation that preverbal YOU-MEI-YOU is grammaticalizing into an auxiliary unit, I will discuss how tests can be applied to justify this claim and to differentiate it from regular verbs and other preverbal constituents (e.g. adverbs) in the second chapter (Section 2.4).

Apart from research efforts to propose cross-linguistically relevant parameters and stages of grammaticalization, another line of studies investigated language-specific phenomena and put forward various criteria for defining and/or characterizing stages of development of linguistic elements. Specifically relevant to the current investigation is research on linguistic units for which a (near-)parallel can arguably be drawn with preverbal YOU-MEI-YOU: although no exact parallels exist in other languages for preverbal YOU-MEI-YOU as an interrogative marker, some semantic and functional equivalents can be found in other languages, which are modals or auxiliaries helping to form perfect constructions or to convey a perfective viewpoint developed from the morpheme meaning possession (e.g. *have* in English, *haben* in High German).²⁸ An examination of studies on the grammaticalization paths of these units may shed light on the development of this new construction. The linguistic criteria these studies have employed to define developmental stages will supplement the general parameters for measuring the degree of grammaticalization.

To sum up, this dissertation will examine the use of preverbal YOU-MEI-YOU in relation to the parameters proposed in Heine (1993) pertaining to the developmental stage

²⁸ According to Smith (1997), the ‘perfect’ is not to be confused with the ‘perfective’; “[t]he former refers to a construction with particular temporal and aspectual characteristics, [while] the latter refers to a closed aspectual viewpoint” (Smith, 1997, p. 106).

of auxiliaries and the linguistic criteria used in cross-linguistic studies on possessive perfects. By comparing the relevant characteristics of preverbal YOU-MEI-YOU with such parameters and linguistic criteria, the study will try to identify the stage of development the construction is currently in.

1.5 RQ2b: What Is the Distribution of Preverbal YOU-MEI-YOU over Time?

As the use of YOU-MEI-YOU before a verb phrase is a relatively new construction and co-exists with some functionally similar constructions that are well-established in Standard Mandarin, the question arises as to whether this construction is maintaining or even gaining momentum or is in fact gradually falling out of use. In grammaticalization research, the frequency of a form “has long been recognized informally a concomitant of grammaticalization” (Hopper & Traugott, 2003, pp. 126-127) and has recently gained more attention in empirical studies. In effect, “increased frequency of a construction over time is [assumed to be] *prima facie* evidence of grammaticalization” (Hopper & Traugott, 2003, p. 129) in diachronic studies of frequency, while synchronic studies often use the evidence from the overall distribution or frequency of a new function (of an existing form) to support their conclusion or argument about a change in progress.²⁹

In light of this, an analysis of the data collected for this dissertation will be carried out to capture the temporal distribution of preverbal YOU-MEI-YOU, with an aim to adding to the evidence of its grammaticalization. An increase in the frequency of use of

²⁹ Hopper & Traugott (2003) noted that the results from such synchronic studies cannot be taken as direct evidence of a change under way, and will have to be compared with studies showing a diachronic change.

preverbal YOU-MEI-YOU over time in the data will be taken as evidence that this construction is advancing in its grammaticalization.

1.6 Chapter Summary and the Organization of the Dissertation

In summary, the current study has been inspired by a dynamic approach to grammaticalization and attempts to examine the synchronic and diachronic dimensions of preverbal YOU-MEI-YOU. More specifically, the study will address two issues raised earlier, namely, the current grammatical status and the development of preverbal YOU-MEI-YOU by drawing on existing research on grammaticalization, auxiliatation, linguistic categorization and aspect, as well as studies on (near-)parallel constructions in other languages and on the historically related construction in Chinese. The goal is to better understand the synchronic uses as well as diachronic processes and uses with respect to the new construction. The analysis will be based on data from several written and spoken corpora.

The dissertation is organized as follows: The second chapter provides an overview of earlier research on the topics listed above, highlighting theoretical frameworks and concepts useful for the analysis of the target construction. The third chapter introduces the procedures and methods of data collection and analysis, and discusses the merits and challenges of the methods and concepts applied in data analysis. Chapter 4 presents the results and findings of the data analysis and will attempt to answer the research questions raised above. Chapter 5 summarizes major findings, discusses the limitations of the current study, and suggests directions for future research.

2. LITERATURE REVIEW

In this chapter, I will first review previous studies on preverbal YOU-MEI-YOU, focusing on how they have failed to address some important issues concerning this construction. I will then provide an overview of grammaticalization research, highlighting the theoretical and methodological concepts in the grammaticalization framework that inform the research question(s) and data analysis in the present study. Next, I will explain the semantic notion of aspect, which is closely related to the first research question, outline the major types and subtypes of aspect, and introduce a particular type of aspect morpheme in Standard Mandarin. Lastly, I will discuss some challenges inherent in the analysis of Chinese data, some methods for categorizing word classes in Standard Mandarin, and the problems with these methods.

2.1 Previous Research on Preverbal YOU-MEI-YOU

The systematic investigation of preverbal YOU-MEI-YOU has only recently started, and two lines of research have been conducted. In the first, focus is on the processes underlying the formation of this new construction and it is often tied into the discussion of language contact and grammaticalization (Xing, 1990; Shi & Li, 2001;

Zhou, 2012).³⁰ In the other type of research, focus is placed on the distribution and patterns of use of this construction. For instance, one kind of question asked by the second line of research concerns pragmatic/discourse functions and propositional attitudes expressed by the YOU-MEI-YOU+VP construction (e.g. emphasis, interrogation, rhetorical question); another question asks about the sentence(s)/clause(s) in which this construction can appear (e.g. main clause, subordinate clause) (Dong, 2004; Wang, Wang, & Jiang, 2006). These two lines of inquiry are complementary and are often pursued simultaneously by researchers investigating the new construction, such as Dong (2004), Shi & Li (2001), Wang, Wang, & Jiang (2006) and Xing (1990).

In much of the research conducted so far, YOU-MEI-YOU+VP has been analyzed in relation to and in comparison with other types of perfective questions, such as VP+LE+Q (see example 3a), VP+LE+MEIYOU (see example 3b), and A-not-A question (see example 3c).

Previous research on YOU-MEI-YOU+VP has provided snapshots of the linguistic and regional variation in the use of this construction. However, two important issues have not been taken up or adequately addressed in previous studies. The first issue that hasn't received due attention is the grammatical status of the construction, i.e. preverbal YOU-MEI-YOU: its status in the grammar remains unclear or unsettled in these studies. The ways this construction has been defined or treated with respect to its categorial membership and the problems with these approaches will be discussed in detail in Section 2.1.1. Suffice it to say that there has not been a systematic, theory-based

³⁰ In this study, the term *construction* is used in a general sense to refer to a grammatical unit or phrase.

analysis of its synchronic usage, which is essential for eventually assigning the new construction to a proper category.

Another issue that remains to be settled concerns the historical development of preverbal YOU-MEI-YOU. Studies conducted so far on this construction have not offered a diachronic account of its appearance, which is important for locating the construction along its path of grammaticalization, and identifying the processes and factors underlying language change. The lack of such diachronic analyses can be attributed to the types of data used and the types of research questions asked in previous studies. I will discuss this in more detail later in Section 2.1.2.

The reason for the insufficient attention given to the issue concerning the grammatical status of preverbal YOU-MEI-YOU is twofold: the focus was on the usage of the new construction. For instance, Dong (2004, p. 1) briefly defined YOU-MEI-YOU as an auxiliary and then downplayed the issue as secondary to describing the usage of the new construction.

Furthermore, determining the grammatical status of preverbal YOU-MEI-YOU is challenging, given the fact that the core of preverbal YOU-MEI-YOU is a lexical verb-YOU ‘have’, ‘there exists’ which is used as such in Standard Mandarin, taking only noun phrases as its complement.³¹ The fact that there does not exist an auxiliary YOU in Standard Chinese used in Mainland China makes it impossible to argue that preverbal YOU-MEI-YOU is the standard A-not-A construction (Dong, 2004; Shi & Li, 2001). Some researchers have postulated that an auxiliary YOU is likely or is starting to take shape through a process of back-formation, whereby the addressee chooses to use

³¹ Except, perhaps, in certain regions such as Taiwan and some Min dialect areas; as discussed in Zhou (2012).

YOU+VP instead of VP+LE in his/her answer to YOU-MEI-YOU+VP questions (e.g. Dong, 2004; Xue, 2010), or as a result of contact-induced analogy (e.g. Taiwanese merging with Standard Mandarin spoken in Taiwan, as proposed in Zhou, 2012).^{32,33} However, since only patchy evidence has been found for the use of YOU+VP in Standard Mandarin used in Mainland China, and given that YOU+VP is highly uncommon as an answer to YOU-MEI-YOU+VP questions in Mainland Standard Mandarin, it is hard to sustain the argument that YOU-MEI-YOU+VP is an implementation of the A-not-A construction of a lexical verb in Standard Mandarin.

Similar to the issue about the grammatical status of preverbal YOU-MEI-YOU, the question about its historical development has not been the focus of research in most of the previous studies. There are studies that have employed a corpus-based analysis but such studies generally set out to probe issues other than the historical aspect of the target construction. For instance, Wang, Wang, & Jiang (2006) looked at examples taken from a large corpus that covered various genres, but the research questions asked and the analysis conducted primarily concerned the structural properties and behavior of this construction, as well as the differences in function and distribution between the target construction and other constructions that have been in use for similar semantic functions. Some researchers have tried to describe the development of this new construction in a certain period, using a small collection of sentences, corpora that spanned a relatively short period of time (e.g. Diao, 2012) or data from selected groups of works or speakers (e.g. Kuang, 2000; Shi & Li, 2001). Nevertheless, the data currently available does not

³² Although the VP+LE is still the most likely answer and YOU+VP is quite rare as they observed.

³³ Standard Mandarin or Standard Chinese spoken in Taiwan is called Guóyǔ ‘national language’.

provide adequate coverage for the type of historical analysis needed to answer the question about its historical development proposed in this dissertation.

In light of how far the previous research on YOU-MEI-YOU+VP has gone in describing the semantic function of the construction, it seems reasonable that the next step would be to determine the grammatical status or category to which preverbal YOU-MEI-YOU belongs, and to identify how it came to acquire that status over time. The theoretical and practical benefits that can be attained from doing so will add to our understanding of the new construction and facilitate future analyses. First and foremost, it is essential that we recognize the different syntactic and semantic patterns of the preverbal YOU-MEI-YOU as manifestations of a shift in categorial membership. Acknowledging such a shift should naturally be followed by efforts to identify the direction towards which it is heading, i.e. the new grammatical category, and the extent to which it has advanced in this new development. Doing so will help keep the analysis of preverbal YOU-MEI-YOU in perspective since such an analysis will then be situated in preexisting discourse on this topic; to be more specific, it will be grounded on the theories and discussion of grammaticalization.

In sum, as part of the effort to better understand an ongoing historical process, that is, grammaticalization, which has been shown to be a primary force that shapes how language develops, the current study will conduct an analysis of preverbal YOU-MEI-YOU, specifically investigating the grammatical shift that underlies the development of this construction.³⁴ To this end, I will draw on some important works on grammaticalization, in particular theories and concepts proposed to portray and account

³⁴ See Meillet (1912, pp. 133), Lehmann (2002 [1982], p. 119).

for the nature of the grammaticalization process, the grammaticalization of auxiliaries, parameters that are germane to stages of grammaticalization of auxiliaries.

Next, I will give a critique of previous studies that discussed the grammatical status of preverbal YOU-MEI-YOU in Section 2.1.1. Section 2.1.2 will cover how the development of this construction was studied in earlier studies on this topic and Section 2.1.3 will discuss the linguistic analyses on preverbal YOU-MEI-YOU in previous studies.

2.1.1 Grammatical Status of Preverbal YOU-MEI-YOU in Earlier Research

Assuming that preverbal YOU-MEI-YOU is not the A-not-A structure, the question that naturally follows is: What is it? A careful look at previous work on preverbal YOU-MEI-YOU reveals that it has been labeled differently but no conclusive evidence has been put forth to help decide its grammatical status. Further, no justifications have been offered for the labels used to describe preverbal YOU-MEI-YOU in individual studies. Some of the prominent studies acknowledge the auxiliary status of preverbal YOU-MEI-YOU in some way (Dong, 2004; Shi & Li, 2001). For instance, Dong (2004) refers to it as an auxiliary, but makes no explanation on why this label had been chosen, apart from the brief comment that “the construction was similar to an auxiliary in English in function and thus could be seen as belonging to the category of auxiliaries, a functional category” (p. 1).

In some other studies, the construction in question is regarded as a combination of the lexical verb YOU and its negative counterpart, MEIYOU ‘not have’, ‘there is not/does not exist’ (Xing, 1990; Wang, Wang & Jiang, 2006) or a combination of a

positive adverb YOU and a negative adverb MEIYOU (Kuang, 2000; Xing, 1990). Many other studies offer no definition or analysis of preverbal YOU-MEI-YOU per se or choose not to take a stance on this issue, and instead examine the use of YOU-MEI-YOU+VP at the sentence or clause level in various contexts (e.g. Shen, 2006; Sun, 2003; Xue, 2010).

In light of the many competing hypotheses that have been presented, it is crucial to understand the criterion for putting preverbal YOU-MEI-YOU in a specific category (auxiliary or other). Whereas Dong (2004) did not describe her criteria, it is clear that Xing (1990) based his judgment mainly on semantic grounds. In his analysis, YOU-MEI-YOU+VP could be divided into two types: YOU-MEI-YOU+VP used in describing a static situation and YOU-MEI-YOU+VP used in depicting a dynamic situation. The YOU-MEI-YOU used in describing a static situation was considered a combination of a lexical verb YOU 'have', 'there is/exists' and its negative form MEIYOU 'not have', 'there is not/does not exist'. The YOU-MEI-YOU used in describing a dynamic situation was seen as a combination of an adverb YOU and its negative form MEIYOU.

For each of the two constructions, Xing's (1990) judgment was made on the basis of different criteria. For the first type of YOU-MEI-YOU+VP, a syntactic test was applied: a degree adverb was added to the verb YOU in the positive answer proposed by the author to the original question. This showed that YOU was a verb and the original question contained this verbal YOU and its negative form. For the second type of YOU-MEI-YOU+VP, Xing claimed that YOU and MEIYOU were purely adverbial and similar to 'ever/once' and 'never' in meaning and hence should be treated as adverbs. However, the VP nature of what follows YOU-MEI-YOU in the example sentences provided in

Xing (1990) is disputable. In fact, some of these phrases are ambiguous and can be classified as nouns in the original contexts (i.e. following YOU-MEI-YOU). For example, one of the verb phrases is ZENGJIA ‘increase’, and a quantifier like YIXIE ‘some’, ‘a little’ can be added between YOU-MEI-YOU and this VP in the example sentence in which it appears in Xing (1990):

- (12) renshu you mei you yixie zengjia?
 number.of.people have NEG have some increase
 Is there some increase in the number of people?

The fact that such phrases can take a quantifier without making the sentence(s) ungrammatical or changing the core meaning of the sentence(s) suggests that they are actually nominal. Thus, for the second type of YOU-MEI-YOU+VP in Xing (1990), the quasi-semantic criterion adopted was in itself problematic.

Xing’s (1990) argument of YOU and MEIYOU being purely adverbial and hence non-verbal is also debatable since the notion of an adverb YOU and its negative counterpart MEIYOU is highly speculative, given that YOU in isolation is only used as a verb in Mainland Standard Mandarin.³⁵ In addition, YOU-MEI-YOU was treated as a monolithic marker in children’s acquisition of YOU-MEI-YOU+VP (Li & Tang, 1991). I will argue that YOU-MEI-YOU, when it appears in unambiguous YOU-MEI-YOU+VP questions, should be understood as expressing the aspect of the sentence, which would be reflected in the response in some way (e.g. VP followed by an aspect marker, LE, see

³⁵ The fact that YOU (‘have, ‘there is/exists’) can be suffixed with aspect markers shows that it is used as the main verb in the sentence(s), as in the following example:

(c) women you-guo meihao de huiyi.
 1PL have-GUO nice DE memory
 We had good memories together.

example sentence 3a). Under this interpretation, it is reasonable to argue that YOU-MEI-YOU should be treated as one unit.

The question remains as to how to determine the category YOU-MEI-YOU falls into when it precedes a VP. To answer this question, I will look at how YOU-MEI-YOU has taken on a new pattern (of distribution) and new (grammatical) functions, and how the new grammatical status of the construction could be defined by examining the new features it exhibits as well as the processes in which such characteristics have developed and been developing.

2.1.2 Development of Preverbal YOU-MEI-YOU: Results and Findings from

Previous Research

Another area where previous research on YOU-MEI-YOU+VP has been slow to get off the ground concerns the diachronic development of this construction. That is not to say this issue has escaped researchers' attention or has been neglected to a significant extent compared to other research questions. In earlier research, the issue appeared in different forms and various analyses were conducted comparing YOU MEI YOU+NP (the old use) and YOU-MEI-YOU+VP (the new use) and/or tracking the distribution of preverbal YOU-MEI-YOU. Some studies also turned to historical texts and speculated about the genesis of preverbal YOU-MEI-YOU.

One example of this kind of research was Kuang's (2000), which looked at the use of YOU-MEI-YOU in the classic works of eight well-known Chinese writers. These writers were considered to be roughly representative of two geographical/linguistic regions (three from Southern China and five from Northern China) born at various times

throughout the 20th century. The five writers from Northern China were each born in different periods; their work represents diachronic slices of written vernacular Chinese or Baihuawen.³⁶ The other three authors, born approximately 18-20 years apart, were speakers of various dialects of Southern China and their work roughly represents synchronic slices of the Baihuawen, according to Kuang.³⁷ She examined these literary works for the use of YOU-MEI-YOU with several different complement clauses, including verb phrase (VP), noun phrase (NP), noun phrase+verbal predicate (NP+VP), and adjective phrase (AP). The study found that, among these constructions, YOU MEI YOU+NP was the most common across time and region, while YOU-MEI-YOU+AP was the least common. Preverbal YOU-MEI-YOU appeared later than all other uses; it was also found that the use YOU-MEI-YOU+VP was more frequent in the works by writers from Southern China and increased over time in both regions (p. 71).

Based on her findings, Kuang (2000) made some hypotheses about the processes and mechanisms underlying the formation and spread of YOU-MEI-YOU+VP. She hypothesized that this new sentence structure had entered Standard Mandarin due to the influence of Southern dialects (e.g. Cantonese, Min dialects) in which parallel constructions exist.³⁸ Kuang's work has produced interesting findings about the

³⁶ Here, Baihuawen is used interchangeably with Baihua to refer to the vernacular literary language used in contemporary China. The emergence and development of Baihua or Baihuawen are closely related to the New Culture Movement and the May Fourth Movement which started at the beginning of the 20th century. It has replaced Wenyan, or the classical literary language, as the standard written Chinese in contemporary China. Baihua is based on Northern Mandarin and has been undergoing changes since the 1920s, incorporating elements from other Mandarin varieties, old Chinese and foreign languages (Chen, 1999).

³⁷ One of them was born in 1920, the second born in 1938 and the third born in 1958. They were all famous contemporary Chinese writers.

³⁸ One example of such parallel constructions is YOUMOU+VP in Cantonese (YOU 'have', 'there is/exists'), MOU, meaning the opposite of YOU). One such example in Min dialects is YOUWU+VP or YOU+VP+WU+VP (YOU 'have', 'there is/exists', and WU meaning the opposite of YOU)

sociolinguistics of this phenomenon and pointed to new venues for investigation into this construction. Since the historical development of YOU-MEI-YOU+VP is only part of what Kuang set out to investigate, the data she collected—although highly valuable—does not paint a full enough picture of the origins of YOU-MEI-YOU+VP to answer the questions posed here.

The same is true of Shi & Li (2001), who surveyed about one hundred influential contemporary Chinese writers in their inquiry into the historical and regional development of YOU-MEI-YOU+VP. Compared with Kuang (2000), who used works by a selected and small group of speakers to represent contemporary Chinese writings, Shi & Li looked at a much wider range of writers, thus providing a more representative sample of this particular population; nonetheless, in both studies, the corpora used were restricted to one genre, i.e. literary works. Therefore, it is questionable whether and to what extent the findings in these studies can be generalized to a larger population, or different genres or forms of speech.

Some studies examined data taken from literary works and other sources (e.g. newspapers and magazines), but such data generally spanned a relatively short period of time. For instance, in Xing (1990), examples of the use of YOU-MEI-YOU+VP were taken from literary works and field notes of what he observed or heard in daily conversations or TV programs. The time-span of his personal observations was not specified but the written examples in the study were from 1984 to 1989; so one could reasonably guess that the observations were made during the same 6-year period in which the data were collected.

Another example of this kind of study is Wang, Wang, & Jiang (2006). Based on data taken from TV programs as well as from some books and newspapers, they offered a detailed description of the distribution, structure, discourse functions and propositional attitudes conveyed by the use of YOU-MEI-YOU+VP and of YOU+VP, as well as of the new usage of an established construction, MEIYOU+VP.³⁹ Most of the data used in this study were from China Central Television, and only a small portion were from local televisions, books and newspapers. Like Xing (1990), the authors did not specify the time when the TV programs, books and newspapers, were first aired or released, but judging from the examples used for illustration in the study, they were taken from a roughly 4-year time span (2001-2004).

Studies such as those as mentioned above cannot possibly provide an adequate account of the historical development of YOU-MEI-YOU+VP, regardless of the size of the dataset, due to the limited time span studied.

2.1.3 Linguistic Analysis of YOU-MEI-YOU+VP in Previous Studies

It should be pointed out that the use of the construction under investigation has been analyzed from a linguistic perspective in previous studies, but such analyses are purely descriptive in nature. For example, Shi & Li (2001) briefly discussed the features of the verbal elements that followed YOU-MEI-YOU, and mentioned that all verbal complements of this kind in the examples they had collected for the study were what they referred to as complex verbal phrases, which include but are not limited to: serial verb constructions, verbs taking a clausal complement, idioms/set phrases. Along similar lines,

³⁹ Wang, Wang, & Jiang (2006) found that MEIYOU+VP was used to express volition, in a way that is similar to the use of BU+VP, in a small number of sentences.

Wang, Wang, & Jiang (2006) examined the verbal complements YOU-MEI-YOU took, and claimed that such verbal complements were often special, or complex, usually consisting of multiple parts. They also found a couple of sentences in which YOU-MEI-YOU preceded what they considered verb phrases expressing some kind of mental action (e.g. *know, worry, fear, believe*) or volition or possibility (e.g. *want, is possible*).

The kind of linguistic analysis as presented in these studies was by and large concerned with how YOU-MEI-YOU+VP differed from the more established structures with parallel functions in terms of structure and pragmatic functions. For instance, in both Shi & Li (2001) and Wang, Wang, & Jiang (2006), the focus was on (i) whether YOU-MEI-YOU+VP could be reworded as VP+(LE)+MEIYOU, (ii) the types of VPs that could be used in the former which were not possible in the latter, and (iii) the differences in speakers' attitudes expressed by these two structures. Xing (1990) also evaluated the possibility of rewording YOU-MEI-YOU+VP as two other types of structures, one of which was VP+MEIYOU. He stated that a lengthy VP in VP+(LE)+MEIYOU questions would cause listener effort and that YOU-MEI-YOU could supplement the former in this sense: YOU-MEI-YOU+(lengthy)VP better expresses the intended meaning by presenting the focus of the question (i.e. the verbal phrases) immediately after the question marker (i.e. YOU-MEI-YOU).

Analyses such as those sketched above produced interesting findings about the linguistic behavior of preverbal YOU-MEI-YOU in comparison with other conventional structures, but were more descriptive than analytical in the sense that they defined the complexity of the verbal complements primarily based on their length. Nevertheless, one study looked beyond the lengthy VPs following YOU-MEI-YOU. In Wang, Wang, &

Jiang (2006), the examples of complex VPs listed included verb phrases taking a clausal complement, verb phrases taking a lengthy adjunct (adverbial phrase), serial verb constructions, and reduplicated verb phrases. The fourth kind of complex VP conveyed, as they argued, an imperfective meaning and was not compatible with perfective morphemes GUO and LE, making it an impossible candidate in the VP+MEIYOU structure that usually has either GUO or LE between the VP and MEIYOU.⁴⁰

Wang, Wang, & Jiang (2006) also looked at the semantics of other types of verb phrases and presented examples in which YOU-MEI-YOU was found to precede (i) what they called mental action verbs (examples listed earlier), (ii) verbs expressing volition or possibility (e.g. XIANG ‘want’, HUI ‘is possible’), (iii) verb phrases that have a negative morpheme or are negative in meaning. In Wang, Wang, & Jiang (2006), the co-occurrence of YOU-MEI-YOU with such a variety of verb phrases is taken as evidence that YOU-MEI-YOU+VP is not simply an alternative to VP+(LE)+MEIYOU but has developed beyond that and has taken on new functions and new types of complements that are not possible in the latter. Putting aside the question about how representative their examples are, it is worth examining the validity of their analysis against relevant linguistic concepts. One problem with the analysis in Wang, Wang, & Jiang (2006) is the lack of internal consistency in the semantic criteria adopted for grouping and labeling verb phrases that follow YOU-MEI-YOU, which resulted in unconvincing reasoning and arguments about the linguistic features associated with the co-occurrence of certain verb phrases and YOU-MEI-YOU. The second problem with the semantic analysis is that

⁴⁰ The examples of reduplicated verb phrases given in Wang, Wang, & Jiang (2006) are XIANG-YI-XIANG, and WEN-WEN. According to Smith (1997), the reduplication of the main verb (without the YI ‘one’ in between) presents a closed situation and conveys a perfective viewpoint.

some words were put into a group on the basis of their semantic similarity, and their structural differences were not taken into account or discussed in the analysis. I will briefly discuss each of these in more detail in the following paragraphs.

One semantic test used by Wang, Wang, & Jiang (2006) for differentiating between verb phrases following YOU-MEI-YOU is the compatibility of the verb phrase with the negation morphemes BU (expressing volition or some stative situation) and MEIYOU (often describing some dynamic situation) in constructions that do not contain YOU-MEI-YOU. This is also used to determine the interpretations of the sentences in which YOU-MEI-YOU takes various mental action verbs: if a mental action VP is compatible with both negation morphemes when used in isolation, the VP or the sentence is said to have two readings: an imperfective reading (expressing stative or ongoing situation, as associated with BU) and an inchoative reading (referring to the beginning of such as state as denoted by the VP, as associated with MEIYOU); in contrast, if a mental action VP in isolation is not compatible with MEIYOU, the combination of YOU-MEI-YOU and this VP is said to have only the imperfective reading.

The problem with this test is that, the verbs that were found in their example sentences were mostly bisyllabic words, with no other words following them in the sentence(s), and therefore could also be interpreted as NPs (as diagnosed by their ability to take a quantifier in the original contexts, e.g. DANXIN, ‘concern, fear’, HAIPA, ‘fear’). For instance, one of the examples given is

(13) ni you mei you danxin?

2SG have NEG have concern

Are you afraid?

A quantifier YIDIAN ‘a bit of’ can be inserted between YOU-MEI-YOU and DANXIN, without changing its meaning.

(14) ni you mei you yidian danxin?

2SG have NEG have a.bit.of concern

Are(n’t) you a little concerned?

It is difficult to determine if the speaker was using YOU-MEI-YOU+VP or simply the A-not-A construction of the lexical verb YOU in the questions.

The other problem with the semantic analysis in Wang, Wang, & Jiang (2006) concerns the group of verb phrases. For example, one type of verb phrase proposed in their study was termed “VP of negation.” Included as examples of this type were verb phrases that contained a negative morpheme (BU) and a verb phrase that was negative only in meaning but not in form—GUOSHI ‘outdate’. Such a grouping is problematic for two reasons. First, it has been established that there is a distinction between “words that are explicitly negative, or syntactic negation, and words that are implicitly negative, or morphological or inherent negation” (Hidalgo-Downing, 2000, pp. 43-44), in terms of their syntactic behavior or co-occurrence with other forms.⁴¹ Second, Jespersen (1917) points out that the practice of assigning a negative value to a word which is not syntactically or morphologically marked for negation) and subsequently labeling it as an inherent negative is, from a theoretical standpoint, to some extent, arbitrary: we can do the same to its antonym and call it an inherent negative instead since “nothing hinders us from logically inverting the order” (p. 43). For example, the verb “succeed” can be

⁴¹ See Hidalgo-Downing (2000) Section 2.3.1 for a review of literature on approaches to identifying and classifying negative words.

thought of as meaning “not fail” just as the verb “fail” can be thought of as “not succeed” (Hidalgo-Downing, 2000, p. 43).

In summary, a systematic linguistic analysis of the use of YOU-MEI-YOU+VP is lacking hitherto. It is important to conduct such analyses in order to gain a better understanding of the structural properties and semantic variation of this construction, which is in turn crucial to identifying its developmental stage(s). Also lacking in previous studies on this construction is a broadly applicable framework for characterizing its developmental stages. In order to understand the structural properties of YOU-MEI-YOU as it is being used in Mainland Standard Mandarin, this study will draw on grammaticalization research that has proposed a cross-linguistic methodological framework for determining the stage(s) of grammaticalization of linguistic items.

A critical examination of studies on grammaticalization is necessitated by the issues of interest in the present study: the shift in categorial membership YOU-MEI-YOU is undergoing and its historical development. As discussed in the previous chapter, YOU-MEI-YOU in YOU MEI YOU+NP is the A-not-A construction of the lexical verb YOU ‘have’, ‘there is/exists’. In YOU-MEI-YOU+VP, YOU-MEI-YOU takes verb phrases as its complement and is used to form perfective questions. In other words, it has a purely grammatical function. This kind of process is generally studied in the research framework of grammaticalization.⁴² In fact, theories from research in this area are part of what drives the first research question of the present study-and will provide methodological underpinning for the data analysis necessary for answering the research questions about the formal status and historical development of YOU-MEI-YOU+VP.

⁴² Here, the term *grammaticalization* refers to a research framework, and elsewhere it is used to refer to the language phenomena studied in this framework.

In this section, I provided a critical examination of previous studies on preverbal YOU-MEI-YOU, with respect to the claims concerning the grammatical status and historical development of this construction. I also pointed out the need for a systematic analysis of the synchronic and historical manifestations of this construction based on theories and methodologies from a broadly applicable research framework (i.e. grammaticalization). In the following section, I will give a broad overview of some of the major theories in grammaticalization research associated with the two issues mentioned above and discuss how they inform the research questions and the methodology. This will be followed by a summary of studies on grammaticalization of possessive perfect/perfective markers (perfect or perfective morphemes that have developed from lexical verbs meaning ‘to possess’), constructions that can be considered near-parallel to the target construction. The goal is to identify linguistic tests that the current study can draw on.

2.2 Grammaticalization: Definition and Measurement

Before proceeding to a detailed discussion about grammaticalization, it is necessary to point out that, in this study, the term grammaticalization is used to refer to the “process whereby a lexical item changes into a grammatical one and a less grammatical item becomes more grammatical (Lehmann, 2002 [1982]).⁴³

As mentioned in passing in the previous chapter, this term is attributed to Meillet (1912), but the notion of linguistic entities undergoing changes to become (more)

⁴³ Here, the term *grammatical* is used to “signify that which belongs to, is part of, the grammar, as opposed to, e.g. what belongs to the lexicon, to stylistics or discourse” (Lehmann, 2002 [1982], p. 8). It is not to be confused with the other sense of the word, which refers to grammatical correctness or conformity to the rule of grammar.

grammatical forms and the ideas behind it can be traced back to a much earlier time. As early as in the 18th century, scholars working on evolutive/evolutionary typology claimed that certain grammatical forms (e.g. personal endings of verbs, prepositions) had evolved from lexical/free forms (see Lehmann, 2002 [1982] for a detailed review on these works).

Two important scholars whose works have had a lasting influence on the study of need to be mentioned here: Wilhelm von Humboldt, and Georg von der Gabelentz.

Humboldt, a German philosopher, gave a lecture in 1822 (cited in Lehmann, 2002 [1982]), in which he put forward a sophisticated account of the evolution of grammatical forms, outlining the stages of evolution of grammatical categories. Humboldt's work was an important contribution to the development of agglutination theory, a theory pursued and/or promoted by some Neogrammarians. Georg von der Gabelentz, a German Neogrammarian, expanded this theory by offering an explanation for grammaticalization and positing the non-linear nature of the evolution of grammatical forms:

grammaticalization was seen as the result of two competing forces, “the tendency towards ease of articulation, and the tendency towards distinctness” (Gabelentz, 1891, p. 251, quoted in Lehmann, 2002 [1982], p. 3); this process was conceived of as cyclical, or spiral, rather than linear.

In essence, the work of Humboldt and Gabelentz sketched out (the basic principles of) what would become the theory of grammaticalization, according to Lehmann (2002 [1982]). Their idea that words with material or concrete meanings would begin to take on formal (structural or grammatical) meanings, and undergo semantic and other changes (e.g. phonological and morphological) to eventually become (purely) grammatical markers in a language is still basic to grammaticalization studies. It is clear

that Meillet's (1912) use of the term *grammaticalization* has developed from the insights of these predecessors, but from Meillet's perspective, the focus was on the transformations of these forms underwent rather than their origins. Meillet also asserted that grammaticalization is one of the two processes whereby new grammatical forms come into being (the other being analogy), and that for either process, "the details may be complicated in each particular case, but the principles are always the same" (p. 131).⁴⁴

In later works on grammaticalization, the conceptualization of this process remained largely unchanged. For example, Lehmann (2002 [1982]), grammaticalization is seen as a process "which may not only change a lexical into a grammatical item, but may also shift an item from a less grammatical to a more grammatical status" (p. 10). In another important grammaticalization study, Bybee, Perkins, & Pagliuca (1994) included in her discussion of *grammaticization* the changes whereby some lexical morphemes or combinations of lexical morphemes with lexical or grammatical morphemes shift to grammatical status, as well as the further development of such morphemes or combinations "after grammatical status has been attained" (p. 5).⁴⁵ In more recent works, grammaticalization has been defined along the same lines. For example, in Hopper & Traugott (2003), grammaticalization "as a term referring to actual phenomena of language, refers most especially to the steps whereby particular items become more grammatical through time" (p. 2). Heine & Kuteva's (2002) definition is another case in point: they defined grammaticalization as "the development from lexical to grammatical forms and from grammatical to even more grammatical forms" (p. 2).

⁴⁴ "[L]es faits de détail peuvent être compliqué dans chaque cas particulier ; mais les principes sont toujours les même" (Meillet, 1912, p. 131).

⁴⁵ The term *grammaticization* was usually used interchangeably with *grammaticalization*.

To sum up, the core concept behind grammaticalization is that a lexical unit takes on a grammatical meaning. The question is: what is a grammatical meaning? Or, what does it mean to be grammatical, as opposed to being lexical? Next, I will introduce how being grammatical is defined, and how linguistic categories are delimited in traditional approaches and in the framework of grammaticalization.

2.2.1 What Defines a Particular Grammatical Category?

In work on linguistic typology (e.g. Humboldt, 1822; Gabelentz, 1891), no definitions of the term *grammatical role/function* have been offered, though abundant examples have been given of a certain class of grammatical words that had evolved from lexical words. For instance, in early speculations about the origins of grammar, a range of grammatical formatives had been claimed to have evolved from lexemes or free forms, such as personal endings of the verb (de Condillac, 1746), articles, auxiliaries and indefinite pronouns (von Schlegel, 1818), etc.

In Meillet's example of the French verb *être*, grammaticalization or the transformation of a lexical word/unit into a grammatical element/role was associated with both its linguistic classes and meaning: The verb *être* 'to be' remains an autonomous word with an existential meaning in some sentences (15), but is a grammatical element in sentences like (16).

(15) je suis celui qui suis

1SG be.1SG DEM REL be.1SG

I am the one who is.

(16) je suis parti

1SG PRF.1SG leave

I left.

Meillet stated that the verb in the latter use was what people improperly called an auxiliary. However, Meillet hinted at what a grammatical element/role referred to, by saying that this use of *être* (what people improperly called an auxiliary) was itself simply part of a complex grammatical form expressing a tense.

That a grammatical role or meaning is what is peculiar to a grammatical class, or is what defines a grammatical class/category seems self-explanatory, but it turns out to be a more complex issue, which is tied up with the problem of linguistic categorization. This question was taken up by researchers working on grammar or grammaticalization, who proposed various criteria for defining grammatical meanings or grammatical categories as opposed to lexical meanings. For instance, quoting Boas, Jakobson made the assertion that “the obligatoriness of grammatical categories” was “the specific feature which distinguishes them from lexical meanings” (1959, p. 489). Lehmann (2002 [1982]) commented that the obligatoriness criterion was a useful but not absolute one, since “[s]omething is obligatory relative to the context” (p. 10), and that the same argument could be made with any other criterion one might propose to differentiate grammatical categories from lexical ones, because the property of being grammatical, or being products of grammaticalization which was a process of gradual change, means having different degrees of grammaticality (p. 11).

Instead of proposing criteria for differentiating between what is grammatical and what is not, Bybee, Perkins, & Pagliuca (1994) enumerated functional and/or

distributional features to define grammatical elements, describing grammatical elements as “closed-class elements whose class membership is determined by some unique grammatical behavior, such as position of occurrence, co-occurrence restrictions, or other distinctive interactions with other linguistic elements” (p. 2). The difference in meaning between grammatical and lexical morphemes was also discussed in Bybee, Perkins, & Pagliuca (1994), which described the former as being associated with more general, abstract or relational meanings and the latter as bearing rich and specific meanings. Their distinction between lexical and grammatical morphemes may be descriptively accurate, but was nonetheless too general and abstract to be applicable in a case-by-case analysis of a morpheme. In addition, as Lehmann (2002 [1982]) stated, the distinction between open sets of lexical items and closed sets of grammatical items is gradual (p. 119).

In fact, as Givón (1979) pointed out, it was never easy to set clear-cut boundaries between major word classes, just as it was difficult to separate grammatical from lexical categories. Givón observed that, even for lexical categories like verbs, nouns, and adjectives, there exists a universal phenomenon whereby such categories “occupy different areas of a continuum” (p. 14), whose scalar property was dependent on time-stability which was a matter of degree. Therefore, even within the same language, the lexical class of adjectives could exhibit “some ‘more noun like’ properties and some ‘more verb like’ ones” (p. 14). This is due to the nature of grammaticalization, which was essentially a gradual change of state (Lehmann, 2002 [1982], p. 11). Hence the inadequacy of grammatical models that depended on binary distinctions between categories.

2.2.2 Gradient-based Approaches to Linguistic Categorization

In recent grammaticalization studies, a static view to language with heavy reliance on discrete categorization was abandoned in favor of a more dynamic approach so as to address the issue of linguistic categorization, or more specifically the question about what makes a particular grammatical category. Heine (1993) pointed to the developments in research in discourse pragmatics, cognitive linguistics and grammaticalization, which lent support to the argument that language should be seen as a dynamic entity and linguistic behavior as a process rather than a state or a product (p. 3).⁴⁶ Heine also reviewed some models of categorization, including the classical model (with definitional criteria involving necessary and sufficient conditions), the prototype model, family-resemblance model. He argued that discrete categorization and prototypicality cannot adequately describe categories or members of a certain category undergoing grammaticalization, since the intermediate stages of their grammaticalization “combine the attributes of both earlier and later stages” (p. 115).

Assuming that grammaticalization is unidirectional process and conceiving of grammaticalization as a process with different stages of development that roughly correspond to successive intermediate links in a chain, Heine (1993) proposed his own approach to categorization, which was built on two parameters: its relative degree of grammaticalization, and its family resemblance. Put simply, the category membership of a linguistic entity is defined in terms of how close it is to the endpoint of a chain of grammaticalization, and how many attributes it has in common with other members of

⁴⁶ It is worth noting that such a dynamic view about language structure has been echoed and further developed in the complexity theory framework (e.g. Larsen-Freeman & Cameron, 2008).

the chain.⁴⁷ Such an approach aligns with his Overlap Model, which characterizes as a continuous and unidirectional process the shift (e.g. conceptual, morphosyntactic and phonological) a given linguistic entity undergoes in its transition from a lexical to a grammatical concept.

The notion of linear family resemblance, which defined Heine's (1993) approach to linguistic categorization, was said to be applicable to the analysis of grammaticalization chains, and particularly relevant to the analysis of auxiliaries, a category that was particularly difficult to describe using traditional definitional properties or criteria. There is as much debate about what types of criteria (e.g. syntactic, morphological, semantic) should be used to define auxiliaries as on whether auxiliaries constitute a linguistically valid category, which can be traced back to the theoretical backgrounds or models adopted by scholars as well as the wide range of functional and formal variation auxiliaries exhibit within and across languages, according to Heine.⁴⁸

Heine's approach, like other approaches that draw on notions like continuum, gradience (e.g. Bolinger, 1980; Garcia, 1967), assumes no hard-and-fast distinction between auxiliaries and main verbs and determines how good or prototypical a certain auxiliary is based on a set of relevant features. More importantly, Heine argued that, when defining auxiliaries, one should adopt a pan-chronic perspective and take into consideration both their synchronic and diachronic manifestations, and consequently, auxiliaries should be defined with reference to the process whereby they came about, a

⁴⁷ According to Heine (1993), non-peripheral members share more attributes with others than peripheral members, but two endpoints typically do not share any attributes (p. 116). See Heine (1993) for the complete list of properties of family resemblance categories used to describe grammaticalization.

⁴⁸ See Heine (1993) for a detailed review.

process Benveniste (1968) called *auxiliation*.⁴⁹ In accordance with the Overlap Model, auxiliaries that have historically developed from main verbs are seen as forming chains, with links corresponding to the different stages of the verbs-to-auxiliaries development. Such chains are described in terms of the functional developments in the verb-to-auxiliary transition, namely, their transformation into markers of tense, aspect, and modality—hence the name Verb-to-TAM chains.

In effect, the Verb-to-TAM chain discussed in Heine (1993) is a typical example of “grammatical channels”,⁵⁰ or “frequently recurring routes signs in different parts of grammar take in grammaticalization” (Lehmann, 2002 [1982], p. 22). Quantitative studies and large-scale survey of grammar and grammaticalization in world languages (e.g. Bybee Perkins, & Pagliuca, 1994; Givón, 1979; Kuteva, 1995; Lehmann, 2002 [1982]; Meillet, 1912, to name a few) have shown that the direction or route a linguistic item or category takes in grammaticalization can be predicted with a certain degree of confidence, once two conditions are known: its meaning and its syntactic functions. Such channels of grammaticalization that linguistic items or categories typically go through as observed in different languages are informative and can be used as the basis for making hypotheses with respect to a particular linguistic item or category undergoing grammaticalization.

To reiterate, Heine (1993) emphasizes that a certain linguistic item or category should be defined as inseparable from the process of grammaticalization comprised of different stages of development towards a certain direction, thus connecting its

⁴⁹ See Kuteva (2001, pp. 1-2) for a definition of the term *auxiliation*.

⁵⁰ According to Lehmann (2002 [1982]), the term *grammatical channels* was used in Givón (1979) and Heine & Reh (1984), but there is no mention of when and where it was first used.

earlier/historical and later uses and allowing for intermediate stages or ambiguity. Accordingly, auxiliaries are defined as “linguistic item[s] covering some range of uses along the Verb-to-TAM chain” (Heine, 1993, p. 70). Kuteva (2001) adopted and extended Heine’s approach to auxiliiation and auxiliaries in his investigation of cognitive forces and discourse-pragmatic factors at work in auxiliiation.

Drawing insights from previous empirical and theoretical work, Krug (2011) put forward a functional definition of auxiliaries. That an auxiliary takes a nonfinite complement, as assumed in Heine’s and others’ definitions, is included as a property of an auxiliary.⁵¹ In addition, in line with Heine’s (1993) idea that “an auxiliary is no longer a full verb but not yet a grammaticalized inflection either” (p. 86), Krug maintains that an auxiliary “needs to have a synchronic allomorph that is either a free form or a clitic” (Krug, 2011, p. 555), in order for the definition to be descriptively inclusive without being all-encompassing: to be specific, morphemes that are more grammaticalized (e.g. clitics) are seen as qualifying as auxiliaries but that clear cases of inflection will not be included in the discussion of auxiliarihood.

Instead of simply adopting the perspective on auxiliaries in Heine (1993) and Kuteva (2001), Krug (2011) aims to broaden it and includes in his definition of auxiliaries other properties in order that it “can account for more notions than TAM” (p. 557). According to Krug, “[a]n auxiliary helps to form grammatical constructions that convey cross-linguistically recurrent meanings beyond person, number and case marking”

⁵¹ Property (i) listed in Krug (2011) states that “an auxiliary is a (de)verbal entity with scope over a (de)verbal complement which is less than fully finite” (p. 558).

(Property iii, p. 558), which include but are not limited to tense, aspect and modality.⁵² In addition, Krug defines an auxiliary as “a verbal or de-verbal entity”, to “include etymological opaque items with verb-like behavior” (p. 558), which further distinguishes Krug’s definition from previous ones.

The present study will follow the linear family resemblance approach to linguistic categories in Heine (1993). This means two things: (i) relevant properties will be identified on the basis of analysis of prototypical members of a category and will be used, *inter alia*, to judge the categorial membership of a linguistic item; (ii) the direction and the endpoint(s) of a grammaticalization chain will be considered and a linguistic item will be measured against such endpoint(s) in terms of categorial membership. In the current study, the analysis will be focused on the semantic and syntactic features since morphological factors are largely irrelevant in the case of Standard Mandarin, a morphologically impoverished language, and it is difficult to determine the degree to which phonology figures in the grammaticalization of pre-verbal YOU-MEI-YOU given that the data used for analysis will be from written corpora. As for the direction of grammaticalization, this study will examine the meaning and syntactic function(s) of the construction under investigation, drawing on findings from cross-linguistic empirical and theoretical research on various channels of grammaticalization.

The construction of interest herein, preverbal YOU-MEI-YOU, when looked at against and compared with the “grams” and “gram-types” (Bybee, Perkins, & Pagliuca, 1994) in different languages, in terms of semantics and syntax, is found to bear a close

⁵² Some examples of non-TAM grammatical meanings include passive, negation, emphasis, sentence types (e.g. declarative vs. interrogative), irrealis, though some can be subsumed under modality, according to Krug (2011).

resemblance to what develops into markers of tense, aspect and modality (called “auxiliaries” or “auxiliary” verbs), as found in cross-linguistic studies (e.g. Bybee, Perkins, & Pagliuca, 1994; Givón, 1979; Lehmann, 2002 [1982]).⁵³ Upon a closer look, the semantic and syntactic behavior of this construction fits in with the conceptualization of auxiliaries in Heine (1993) and Krug (2011) in particular:⁵⁴ it takes a verbal phrase as its complement and serves to convey grammatical meanings, e.g. aspect and sentence type.⁵⁵

Against the backdrop of the findings stated above, the first research question, i.e. what is the grammatical status of preverbal YOU-MEI-YOU, will be phrased as a claim: preverbal YOU-MEI-YOU is turning into an auxiliary unit through the process of grammaticalization. The hypothesis will be tested by analyzing the linguistic features and behavior the construction exhibits in real language use. Underlying this hypothesis is the assumption that YOU-MEI-YOU is one unit, an argument made by Dong (2004).

It should also be noted that only YOU-MEI-YOU taking a VP complement is considered an auxiliary, since elsewhere YOU-MEI-YOU is still found to take nominal complements and convey only lexical but no grammatical meanings. In order to make sure to include only instances of preverbal YOU-MEI-YOU in the analysis, methods of word class categorization adapted from Guo (2002) and Yuan (2010) will be applied to eliminate sentences in which YOU-MEI-YOU is followed by an NP or other

⁵³ Bybee, Perkins and Pagliuca (1994) claim that grammatical morphemes, or “grams”, can be studied as “gram-types” that “can be viewed as analogous to the phonetic description of the type ‘voiceless bilabial stop’” (p. 149), which are substantive universals that develop in similar ways across languages. Examples of gram-types include future, past, perfective, and imperfective. However, that does not entail that “a language NEEDS a particular gram-type” (p. 298, [original emphasis]) since “no gram-types are universal” (p. 298).

⁵⁴ Following Heine (1993) and Krug (2011), this study adopts the position that auxiliaries constitute a universal functional category (cf. Kaisse, 1981; Reuland, 1983).

⁵⁵ It is used to express the viewpoint aspect of the sentence. I will elaborate on the topic in Section 2.3.

complements that are not of verbal nature.⁵⁶ These methods will be discussed in detail later this chapter and the benefits and problems with these methods will be illustrated with examples in the methodology section of the dissertation. The notion of aspect, an important grammatical meaning or function, which has a direct bearing on the categorization of preverbal YOU-MEI-YOU, will be elaborated in the last part of this chapter.

After preverbal YOU-MEI-YOU is categorized, which entails determining the direction towards which the grammaticalization of preverbal YOU-MEI-YOU proceeds, a natural next step is to investigate its development stages and trajectory. The major distinction between the first research question, i.e. the question about the grammatical status of preverbal YOU-MEI-YOU, and the second and third research questions, i.e. questions about its development stage and trajectory is that, the former is essentially concerned with the overall characteristic of the construction, while the latter deals with the minutiae or fluctuations that constitute and reflect the process of its development. That means, in order to gain insights into this process, we need to first look at such observable data or behavior as the construction exhibits across region, time and contexts to see if there are patterns of development.

⁵⁶ If a sentence/clause contains YOU-MEI-YOU preceding an adjective and preposition, it is considered an example of YOU-MEI-YOU+VP construction, as adjectives and prepositions can function as the predicate of a sentence and behave like verbs in Chinese; in fact, adjectives are considered as a type of intransitive verb by some Chinese linguists (see Li & Thompson, 1989). In a more fine-grained analysis of the complements of YOU-MEI-YOU necessary for answering the second research question, adjective phrases are put into a separate group.

2.2.3 Characterizing Stages of Grammaticalization: Features and Parameters

A number of studies on grammaticalization have provided language-specific or cross-linguistic description of the changes affecting various linguistic areas (e.g. morphology, phonology, semantics, syntax) in the process of grammaticalization (e.g. Bybee, 1985; Bybee, Perkins, & Pagliuca, 1994; Givón, 1979; Heine, Claudi, & Hünnemeyer, 1991; Heine & Reh, 1984; Lehmann, 2002 [1982]) and some have attempted to outline the linguistic features or changes characterizing different stages of development and parameters that can be used to delineate such stages at a cross-linguistic level.

Lehmann (2002 [1982]) was the first to have laid out a systematic set of parameters for measuring degree of grammaticality, but the validity and applicability of these parameters remain questionable. As Lehmann himself acknowledged, these parameters were abstract and difficult to quantify (e.g. the parameter of paradigmatic variability).⁵⁷ Even though he proposed aspects for each parameter that could be measured or quantified, he made the caveat that they would not be able to “stand up to the requirements of linguistic theory and methodology” (p. 144).

On top of the issue of operationalization, there are also questions about the validity or correctness of some of these parameters. Among the six parameters Lehmann proposed, the first one, the paradigmatic weight of a sign, is related to semantic bleaching, a notion that has been found problematic: grammaticalization leads to loss of lexical or content meaning, but with it also comes gain of grammatical meaning (Sweetser, 1988).⁵⁸

⁵⁷ Defined as “the freedom with which the language user chooses a sign” (Lehmann, 2002 [1982], p. 123).

⁵⁸ The ‘paradigmatic weight of a sign’ is defined as “its possession of a certain substance which allows it to maintain its identity, its distinctness from other signs” (Lehmann, 2002 [1982], p. 113).

In fact, “the process of demotion of some lexical meaning and promotion of others is characteristic of semantic change in general” (Hopper & Traugott, 2003, p. 98). Besides, the initial stages of grammaticalization involve “a redistribution or shift, not a loss, of meaning” (Hopper & Traugott, 2003, p. 94). As Hopper & Traugott contend, semantic bleaching “must therefore be taken to be a very relative notion and one that pertains almost exclusively to late stages of grammaticalization” (p. 98). Hopper & Traugott also questioned the third parameter, one that concerns the “obligatoriness” of using a sign for expressing a certain (grammatical) meaning, since historical processes like grammaticalization “are always ongoing, and furthermore are not deterministic”, meaning that it is impossible to predict if one sign would win out and be selected instead of other signs with equivalent functions in all contexts.

In spite of the above-mentioned problems with the parameters proposed in Lehmann (2002 [1982]), some of these notions have proved useful and cross-linguistically relevant. For example, although the parameter of paradigmatic weight was contentious in its original definition, the idea that a grammaticalized entity would undergo “gradual loss of phonological substance” (p. 113) and desemanticization, or a “decrease in semanticity” (p. 113), however “semanticity” was defined, was acknowledged and included in the general discussion of grammaticalization. In addition, Bybee, Perkins, & Pagliuca (1994) found that in the process of grammaticalization, the syntactic positioning of a gram became increasingly rigidified and grammatical classes tended to reduce in size, lending support to what Lehmann stated about the parameter of

syntagmatic variability and the parameter of paradigmaticity, which were found to be partly negatively correlated to the stages of grammaticalization.⁵⁹

Other researchers approached the question about degree of grammaticality or grammaticalization stages by looking at the changes within different linguistic areas. The project conducted by Bybee, Perkins, & Pagliuca (1994), for example, focused on the co-evolution of meaning and form (confined to phonetics in their study) in the grammaticalization of various grams or gram-types on particular grammaticalization paths and proposed parameters for measuring the degree of phonological and semantic development (e.g. the shortness parameter and the fusion parameter for measuring the phonetic reduction, and specificity vs. generality, concrete vs. relational meaning for measuring semantic reduction).

With respect to auxiliaries, Heine (1993) provided a systematic account of the development of auxiliaries and discussed four parameters in this verb-to-auxiliary development, involving the series of shifts that occur in four linguistic areas, namely, semantics, morphosyntax, morphophonology, and phonetics. The shift in each of the four areas was conceived of as a chain—hence the desemanticization chain, the decategorialization chain, the cliticization chain, and the erosion chain. Heine attempted to characterize the salient stages along each of the chains and examine the relevance and/or exemplification of the features characteristic of different stages along each chain in the overall Verb-to-TAM chain consisting of distinct stages. The notions of four sub-chains involved in the Verb-to-TAM auxiliatation, along with the general stages of such

⁵⁹ Syntagmatic variability is defined as “the possibility of shifting it around in its construction” (Lehmann, 2002 [1982], p. 110). Paradigmaticity is defined as “the degree to which it enters a paradigm, is integrated and dependent on it” (Lehmann, 2002 [1982], p. 110).

auxiliation proposed in Heine, will be looked at and utilized wherever applicable to guide the identification of the developmental stages of the construction under discussion in this study.

2.2.4 Frequency of Grammaticalized Forms

Another feature or factor that has recently attained more importance in grammaticalization research is the textual frequency of grammaticalized forms. As Hopper and Traugott (2003) observed, the findings from empirical studies on the frequency of forms not only provide evidence for unidirectionality in grammaticalization, but they are also helpful in understanding “how lexical forms move into grammatical roles” (p. 127). A distinction is made between two types of frequency: type frequency and token frequency. The former refers to “the number of items that are available to a particular class of forms” (e.g. different affixes for marking plurality in English; Hopper & Traugott, 2003, p. 127) and the latter concerns “the number of times a particular form...occurs in texts” (Hopper & Traugott, 2003, p. 127).

Frequency of forms that are suspected of or found to be “moving toward some kind of grammatical status” (Hopper & Traugott, 2003, p. 129) is treated differently in synchronic studies and diachronic studies, though they both draw on this kind of information, in one way or another. In synchronic empirical studies, the findings would be used to support their arguments about a change in progress or compared with “a well-described type of diachronic change” (Hopper & Traugott, 2003, p. 129). For instance, Barth-Weingarten & Couper-Kuhlen (2002) looked at occurrences of the English word *though* functioning as a discourse marker and the instances in which the word had a

different or “fixed” function and listed the difference in frequency (in percentage) of these uses. However, as Hopper & Traugott (2003) noted, such data need to be compared with other studies that show such a change is possible before they can be used as evidence for a change in progress.

In diachronic studies, frequency is taken as direct evidence for grammaticalization of forms. One example of this kind of diachronic study is Laury (1997), who examined the frequency of lexical nouns accompanied by a demonstrative in Finnish in texts from three historical periods. Laury showed that “the use of the demonstrative *se* (and its case forms) is becoming increasingly obligatory” (Hopper & Traugott, 2003, p. 130) and argued that it was evidence for the emergence of a definite article in Finnish, even though “Finnish was often characterized...as a language that lacks a definite article” (Hopper & Traugott, 2003, p. 130).

The assumption or belief that an increase in frequency is indicative of grammaticalization, however, is challenged in recent grammaticalization studies. For instance, Bybee (2003) argues that “grammaticalization is not only a result of grammaticalization, it is also a primary contributor to the process” (p. 602). She examined the grammaticalization of the English modal *can*, and discussed how the frequent use of this verb in general, its frequent use in various contexts (i.e. with verbs of different classes) and in some particular combinations lead to semantic bleaching at different stages of development. Similarly, Diessel (2007) claims that the loss of phonetic substance and/or pragmatic and semantic force of linguistic expressions undergoing grammaticalization are “due to frequent language use” (p. 117). Nevertheless, it is agreed that high frequency is not the only factor that drives grammaticalization. Bybee (2003)

acknowledges that while “[r]epetition is universal to the grammaticalization process” (p. 622), this factor alone “cannot account for the universals of grammaticalization” (p. 622).

Studies that look at frequency identified two major types of seemingly contradictory effects of frequency (e.g. Bybee, 2003; Diessel, 2007). On the one hand, high token frequency “leads to phonetic reduction and the development of new linguistic forms” (Diessel, 2007, p. 119) but on the other hand, “token frequency can be a conservative force protecting high-frequency structures from analogical leveling” (p. 119).

Heine & Kuteva (2007) counters the argument in Bybee (2003) and Diessel (2007) that frequency of use is “the main trigger of grammaticalization” (p. 38). Drawing on Heine, Claudi, & Hünemeyer (1991) and a comparison of two German verbs, they conclude that “frequency of use appears to be an epi-phenomenal product of extension rather than a trigger of it” (p. 39). Their conclusion is based on two findings about the instances of grammaticalization they looked at:

[O]verall, non-grammaticalized items that serve as the source of grammaticalization do not necessarily belong to the most frequently used words of a language, nor are grammaticalized items necessarily used more frequently than their non-grammaticalized counterparts. (p. 38)

Peng’s (2011) investigation of various strings of Chinese morphemes produced similar results. Nonetheless, Peng (2011) maintains a slightly different view from Heine & Kuteva (2007). He proposes a distinction between two kinds of token frequency, namely “critical frequency” and “non-critical frequency”. He defines “critical frequency”

in relation to the concept of “critical contexts” (Diewald, 2006), in which pragmatic inference is possible and new grammatical meanings can be derived. Comparing the grammaticalization of various structures in Chinese, Peng shows that it is not general token frequency, but “critical frequency” that plays a role in grammaticalization.

Since frequency could be a relevant factor in the process of grammaticalization, it is worthwhile to investigate the distribution of preverbal YOU-MEI-YOU in different historical periods. Attempts will be made to identify patterns of distribution across time and region and caution will be exercised when interpreting the results, given the debate on the role of frequency in the literature.

Next, I will review some studies on perfect or perfective markers that have grammaticalized from lexical verbs meaning possession, focusing on the linguistic tests or criteria used to characterize stages of grammaticalization. Since these linguistic elements in other languages are semantically or functionally equivalent to preverbal YOU-MEI-YOU (or to the verb YOU to be exact) and have undergone similar conceptual changes (from lexical units to auxiliary units), findings regarding their developmental paths may help us better understand the grammaticalization of preverbal YOU-MEI-YOU. To be more specific, the findings about the prominent features that characterize different stages of development of such linguistic items or units in these studies may also inform the data analysis in the current study for answering the research questions(s) about the historical development of preverbal YOU-MEI-YOU.

2.2.5 Grammaticalization of Periphrastic Perfect: Cross-linguistic Evidence

The grammaticalization paths of possessive perfects (from verbs of possession to aspect or tense markers) in European languages were given dedicated attention in Heine & Kuteva (2006), whose cross-linguistic study established the correlations between the structural changes and the stages of grammaticalization of verbs meaning possession (e.g. English *have*). Some important observations can be made from the figure in Heine & Kuteva (2006, p. 151) outlining the development stages of possessive perfects, summarized as follows:

- (i) The grammaticalization of a verb of possession begins; there is a noun phrase referring to the object possessed by the subject in the sentence; there is another verb in the form of past participle in the sentence, modifying the noun phrase.
- (ii) The grammaticalized morpheme (i.e. the verb of possession) goes beyond the first stage (or Stage 0 as Heine & Kuteva call it) and reaches the second stage (or Stage 1) at which it attains auxiliarihood; the past participle form of the other verb in the sentence acts as a main verb and the noun phrase becomes its direct object.⁶⁰
- (iii) The grammaticalized morpheme becomes an auxiliary and the sentence no longer requires the presence of a noun phrase functioning as the Patient (i.e. the under-goer of the action denoted by the predicate or verb phrase) (p. 151).

⁶⁰ According to Heine & Kuteva (2006), during Stage 0, the past participle, being the modifier of the noun phrase taking the thematic role of a Patient, shows inflectional agreement with the latter. At Stage 1, it acts as a main verb but the agreement may still be there and will not be lost until Stage 2 (p. 151).

Drawing on Heine & Kuteva (2006), Łęcki (2010) conducted an extensive and thorough investigation into *habban* ‘have’ + past participle in Old English (OE), Łęcki made a strong case for his assertion that this construction in the OE period was a full-edged perfect. He argued that the OE *habban*+past participle construction was placed in a relatively advanced stage of development, when evaluated with the parameters formulated by Heine & Kuteva (2006).

The linguistic criteria introduced and used in the aforementioned studies for determining the auxiliary status of the grammaticalized morpheme include the semantic role/agentivity of the subject, the inflectional agreement between the past participle-verb and the Patient NP, and the type of VP complement the grammaticalized morpheme takes, the (propositional) meaning of the grammaticalized construction, etc. Similar diagnostic tools were used in Coussé (2013), who also included the concreteness of the direct object, and telicity of the past-participle-verb in the tests applied to determine the degree to which the grammaticalized construction has extended contextually.

Along similar lines, Rittenhouse (2014) focused on the inflectional patterns and semantics of the past participle in three types of periphrastic perfect and passive constructions, namely *become/be/have*+past participle, in her corpus-based study on the development of periphrastic perfect and passive in Old High German and Old Saxon. The frequency of overt nominal inflection or lack thereof, as well as the patterns of aspectual properties exhibited by the past participles in these constructions in two historical texts was examined and used to argue for the stage of grammaticalization of these periphrastic verbal constructions.

The cross-linguistic research on the grammaticalization of periphrastic verbal constructions conveying aspectual or voice meaning has shown that the linguistic features or tests sketched above could help determine the developmental stages of a grammaticalized construction.

In this section, I discussed some important topics in grammaticalization research, including how the distinction between grammatical and lexical meanings/categories was determined, the different approaches to linguistic categorization, criteria/parameters for measuring degree of grammaticality, and briefly introduced studies on the grammaticalization of possessive perfects across languages, the methods used in which may be applied in the analysis of preverbal YOU-MEI-YOU.

The following section will cover some important studies (e.g. Comrie, 1976; de Swart, 2012; Filip, 2012; Smith, 1997) on the theoretical concept of aspect—an important concept in semantics which plays a key role in defining preverbal YOU-MEI-YOU. It will also briefly introduce the perfective aspect in Standard Mandarin, as well as various perfective morphemes. These will be investigated in chapters 3 and 4 as a means of better understanding the grammatical meaning and behavior of preverbal YOU-MEI-YOU.

2.3 Aspect: A Semantic Account

Comrie's *Aspect* (1976), one of the most influential works on aspect and related concepts, defined aspect as “ways of viewing the internal temporal constituency of a

situation” (p. 3). The definition was illustrated in the following sentence from French along with its English parallel (Comrie, 1976, p. 3):⁶¹

- (17) Jean lisait quand j=entrai.
 John read.IPFV.3SG when I SG=enter.1SG
 John was reading when I came in.

According to Comrie, the second verb presents the situation referred to in the second clause, i.e. my entry, as a whole, “with beginning, middle and end rolled into one” (p. 3); in contrast, the first verb phrase makes “reference to an internal portion of John’s reading” (p. 4), the situation referred to in the first clause, while “there is no explicit reference to the beginning or to the end of his reading” (p. 4). This leads to the interpretation that my entry happened at the same time as only one portion of John’s reading, and that his reading “both preceded and followed my entry” (p. 3). The first is said to have imperfective meaning, and the second perfective meaning.⁶²

In Comrie (1976), the noun *aspect* or the plural *aspects* is restricted to “referring to particular grammatical categories in individual languages that correspond in content to the semantic aspectual distinctions drawn” (p. 7) as illustrated by the examples above. He also discussed the inherent aspectual properties (e.g. durativity, telicity, stativity⁶³) of

⁶¹ Translation equivalents taken from five languages are listed in Comrie (1976), and only two are listed here for illustration purpose.

⁶² To be exact, the French verbs “lisait” and “entrai” exemplify the distinction between imperfectivity and perfectivity, but the difference between “was reading” and “entered” in the English example is that of Progressive and Non-Progressive. Comrie (1976) noted that, “provided we restrict ourselves to nonstative verbs and exclude habitual meaning, then the difference between the two forms is one of the imperfectivity versus perfectivity” (footnote in Comrie, 1976, p. 4)

⁶³ Durativity means the quality of “a situation that lasts for a certain period of time (or at least, is conceived of as lasting for a certain period of time)” (Comrie, 1976, p. 41) and is used in opposition to punctuality. A telic situation has a natural or built-in endpoint while an atelic one does not (Comrie, 1976, p. 44). A state differs from a dynamic situation in the sense that a state will continue “unless something happens to

various classes of lexical items but made a distinction between such semantic aspectual meaning and the aspectual categories in a language, thus excluding the semantic properties conveyed by particular lexical items from his definition of *aspect*.

Developing on Comrie (1976), Smith (1997) proposed a two-component theory of aspect and made a broad distinction between two types of aspect: viewpoint aspect and situation aspect. Viewpoint aspect is what was defined as aspect in Comrie, and is also called grammatical aspect in de Swart (2012): it presents “situations with a particular perspective or focus” (Smith, 1997, p. 2). Situation aspect, by contrast, refers to “classes of events and states” and bears on the inherent meaning of verbs, as illustrated in the following sentences (exx 4a-b from de Swart, 2012, p. 753):

(18) Bill was in love with Susan.

(19) Sarah wrote a dissertation.

The first sentence describes a state that holds during a certain period of time, while the second sentence describes a completed event. What differentiates the two sentences is that writing a dissertation, unlike being in love, is a process with an inherent endpoint.

Situation aspect is also called aspectual class or *Aktionsart* (in its new extended sense).⁶⁴

change the state” whereas “[w]ith a dynamic situation, the situation will only continue if it is continuously subject to a new input of energy” (Comrie, 1976, p. 49).

⁶⁴ The lack of generally accepted terminology in the discussion of aspect was noted in early studies like Comrie (1976), who mentioned the two kinds of distinction between aspect (in his definition) and *Aktionsart* (p. 7). See Filip (2012, p. 725) for an introduction of the change in the meaning of the term *Aktionsart(en)* and its use in linguistic works.

2.3.1 Viewpoint Aspect and Situation Aspect: Expression, Classification and Relation

Viewpoint aspect, or grammatical aspect, is often expressed by a grammatical morpheme, usually verbal (Smith, 1997, p. 2). According to Filip (2012), it can be expressed by “a grammatical marker on a verb in a given sentence” (p. 724), or syntactic constructions usually comprised of an auxiliary and a non-finite verb (e.g. *be+V-ing* in English). The expression of situation aspect is more complex: it was reasoned that not only the meaning of verb phrases, but also the meanings of the arguments (e.g. subjects, objects), and the thematic relation between these two are responsible for the aspectual character of the sentence (Verkuyl, 1972; Comrie, 1976, p. 45). Verkuyl argued that “[situation] aspect needs to be defined at the level of the predicate-argument structure (VP and S)” (cited in Filip, 2012, p. 754). Similarly, Smith (1997) states that “[s]ituation type is conveyed by the verb constellation”, which he defined as “a main verb and its arguments, including subject” (p. 2). In addition, he noted that “[t]he relation between verb constellation and situation types is not one-to-one” (p. 18), and suggested that one distinguish between basic-level and derived categorization of sentences in terms of situation types.⁶⁵

There is much variation in the classification of aspect, viewpoint and situation alike, in the literature on aspect. In the case of viewpoint aspect, it is relatively straightforward. Comrie (1976) was primarily concerned with the perfective versus imperfective opposition, though he also outlined subdivisions of imperfectivity and mentioned the existence of some language-particular categories, which “often combine

⁶⁵ According to Smith (1997), the basic-level categorization is always for a verb constellation, but “the derived level requires adverbial or other information from context” (p. 18).

aspect and some other category, most usually tense” (p. 9).⁶⁶ In Smith (1997), apart from the perfective and the imperfective, there is also a third type: neutral viewpoints.⁶⁷

According to Smith (1997), it is a default for sentences with no explicit aspectual morpheme, and it includes “the initial point and at least one stage of a situation” (p. 62). Smith also discussed perfective viewpoints in certain languages and two common types of imperfectivity, the general imperfective and the progressive. The tripartite classification of viewpoint aspect proposed in Smith (1997) will be followed in the data analysis in the current study.

The picture is much more complicated for situation aspect, or aspectual class. In event semantics, it is commonly classified into process, events and states, according to Filip (2012). Dowty (1979) also proposed three types of aspectual classes, namely state, indefinite change and definite change while Vendler (1957) listed four: activity, accomplishment, achievement and state.⁶⁸ Smith (1997) added semelfactives to Vendler’s list as a fifth situation type.⁶⁹ More recently, studies based on degree-based theories and mereological approaches to aspect discussed in Filip (2012) suggested two new types of verbs—incremental verbs and scalar verbs, which “are not aligned with [the] traditional lexical aspectual classes” (p. 744) introduced above. Again, the classification of situation aspect in Smith (1997) will be adopted in the data analysis, since it not only includes the four categories proposed or suggested in other studies that are commonly accepted but is

⁶⁶ Imperfective is divided into habitual and continuous, the latter of which is further divided into progressive and non-progressive (see the classification of aspectual oppositions in Table 1, p. 25).

⁶⁷ Perfective viewpoint, as defined by Smith, “focus a situation in its entirety, including both initial and final endpoints” (p. 3), while imperfective viewpoints “focus an interval, including neither initial nor final endpoints” (p. 3).

⁶⁸ Judging from the example given, indefinite change in Dowty’s classification roughly corresponds to activity in Vendler’s terms, and the definite change to accomplishment and achievement.

⁶⁹ This category was introduced in Comrie (1976, ch. 2) in his discussion of punctuality and durativity.

broader in range. The implications of the findings from the latest approaches to lexical aspect mentioned in Filip (2012) will be borne in mind but the two new categories will not be used in the present study for the sake of simplicity and economy.

The relation between viewpoint aspect and situation is a complex one. Smith (1997) argued that these components are seen as independent from each other, though they also interact with each other. It is much more difficult to tease these two apart when languages like Standard Mandarin and Russian are considered which do not have strict boundaries between them, according to de Swart (2012). It is generally agreed that the former “determines the aspectual nature of the sentence as a whole and may overrule certain semantic features of its internal aspectual make-up” (de Swart, 2012, p. 766).

Viewpoint aspect and situation aspect also interact in complex ways. Viewpoints have certain selection requirements and tend to occur with some types of situations but not others, but when there is an unusual combination of these two, there may be aspectual coercion.⁷⁰ When this happens, it will trigger a shift in situation type, resulting in a different reading of the sentence. For example, when a stative predicate like BING ‘sick’, which is normally incompatible with the aspectual morpheme LE in Standard Mandarin, is followed by LE in a sentence, it gives rise to an inchoative reading with the effect that the subject got sick. The findings of the relation and interaction between viewpoint aspect and situation aspect are informative and will be kept in mind when the data are analyzed with respect to aspectual meaning.

In addition to laying out definitions and classification of aspect, and how the interactions between different kinds of aspect affect the aspectual interpretation of a

⁷⁰ “Aspectual coercion...requires the eventuality description to shift its meanings to satisfy the aspectual selection requirements” (de Swart, 2012, p. 769) of the grammatical aspect.

sentence, some previous research on aspect also provided theoretical tools for determining the aspectual meaning of a given sentence. One typical example is Smith (1997), who used the technique of indirect proof by means of semantic tests. For instance, she conjoined a sentence with a clause asserting the non-completion of the situation in question to find out if the sentence presented a complete event. If the conjunction was reasonable, then he could adduce evidence that the original sentence had an open interpretation or the viewpoint was not perfective. In a similar light, questions about the continuation of a situation could be used to determine if the sentence presented an open situation. The use of indirect proof or semantic tests in the investigation of aspectual meaning conveyed in a sentence has important implications for the analysis of YOU-MEI-YOU, as its status as an auxiliary, or an aspect marker to be exact, has yet to be established. The current analysis will follow Smith's (1997) method of analysis and design suitable semantic tests to investigate and determine the aspectual meaning conveyed in sentences/clauses containing the YOU-MEI-YOU+VP construction.

The following is a summary of research on some of the perfective morphemes and/or constructions in Standard Mandarin. This will serve two purposes: first, a knowledge of such morphemes/constructions is necessary for understanding the above-mentioned semantic tests designed to determine the aspectual meaning of a YOU-MEI-YOU+VP clause, as these morphemes are indispensable to such tests. Second, a couple of such perfective morphemes and/or constructions are semantically and historically related to the preverbal YOU-MEI-YOU and an examination of these items might give us insights into the development of the target construction.

2.3.2 Perfective Viewpoints and Perfective Morphemes in Standard Mandarin:

GUO, LE, and MEI(YOU)

Smith (1997) mentioned three perfectives in Standard Mandarin: perfective LE⁷¹ and perfective GUO, both of which are verbal suffixes, as well as a lexical perfective called Tentative “formed with the reduplication of the main verb” (p.271). Apart from the difference in their syntactic properties, these perfective constructions also convey different aspectual meanings. Perfective GUO “presents a prior closed situation, and conveys that its final state no longer obtains” (p. 266), as illustrated in the following example:

- (20) ta shang ge yue qu-guo beijing.⁷²
 3SG last CLF month go-GUO Beijing

Last month he went to Beijing (he is no longer there).

In addition, this morpheme also carries experiential sense “presenting a given situation as a member of a set” (Smith, 1997, p.268), which requires a repeatable situation (see example 21 below):

- (21) Question: ni chi-guo pisa meiyou?⁷³
 2SG eat-GUO pizza NEG

Have you ever eaten pizza?

⁷¹ The perfective morpheme LE should not be confused with the sentence-final particle LE, which has developed from the verbal suffix LE (Wang, 1958, p. 445). The former conveys aspectual meaning and the latter signals the current relevance of some state of affairs to some particular situation (Li & Thompson, 1989, p. 240). It is hard to distinguish between these two forms in a sentence that is structurally ambiguous (e.g. Subject+VP+LE). A post-verb perfective LE and a sentence-final particle LE can co-exist in a sentence.

⁷² Example adapted from Smith (1997, Chapter 11, ex. 8b).

⁷³ Example adapted from Smith (1997, p. 268, example 12).

Answer: chi-guo

eat-GUO

Yes, I have. (presumably on multiple occasions)

In summary, the perfective viewpoint conveyed by GUO is functionally equivalent to the English Perfect, “presenting a situation prior to Reference Time and ascribing to an experiencer the property of having participated in the situation” (Smith, 1997, p. 269). It is compatible with verb constellations of any situation type.

By comparison, the LE perfective in Standard Mandarin, also conveying perfective viewpoint, presents “closed events of a single point or an arbitrary endpoint” (Smith, 1997, p. 266). Only events which can be bounded, i.e. which have initial and final endpoints, are compatible with this LE perfective. By contrast, if LE is applied to a sentence with a stative verb constellation presenting an undifferentiated period in which the situation is true and implying no initial or final endpoints (e.g. something equivalent to the English ‘Mary likes him’), it will cause the situation type to shift from stative to inchoative, showing that the perfective LE “does not appear with statives” (Smith, 1997, p.265). Another property that differentiates the perfective LE from perfective GUO is a discontinuity with the present or other Reference Time which is only required of the latter (i.e. GUO) but not of the former (i.e. LE), as illustrated in the following pair of sentences:

(22)/(20) ta shang ge yue qu-guo beijing

3SG last CLF month go-GUO Beijing

Last month he went to Beijing (he is no longer there).

(23) ta shang ge yue qu-le beijing.⁷⁴

3SG last CLF month go-LE Beijing

Last month he went to Beijing (he may be still there).

It is important to note that the perfective viewpoint conveyed by LE “semantically conveys termination, not completion” (Smith, 1997, p.264), although conventionally it suggests a completive interpretation. The same point was made in Li & Thompson (1989), who highlighted the disconnect between the perfective LE and the past time, and between LE and completion in various contexts.

For Li & Thompson (1989), the expression of a perfective viewpoint through the use of this verbal suffix means that the event is being viewed in its entirety, which is possible when the event is “bounded temporally, spatially, or conceptually” (p.185). According to Li & Thompson (1989), an event can be bounded in the following ways: “by being a quantified event, by being a specific or definite event, by being inherently bounded because of the meaning of the main verb and by being the first event of a sequence” (p. 185). The present study will utilize these ideas in the examination of sentences containing the preverbal YOU-MEI-YOU and determine the aspectual readings of the target sentences through semantic tests incorporating these perfective morphemes.

In addition to the two perfective viewpoint morphemes found in the literature, there is another morpheme in Standard Mandarin that has been treated as an aspectual marker: the negative morpheme MEI or MEIYOU which can be used interchangeably in Standard Mandarin.⁷⁵ In the present study, they will be treated as the variants of the same

⁷⁴ Example adapted from Smith (1997), Chapter 11, (8a).

⁷⁵ These two negation morphemes have undergone different processes of grammaticalization historically and have come into existence in different periods of time (see Shi & Li, 2001, Section 14.5 for more

negation morpheme and juxtaposed in the discussion (but see Wang, 1965 for a different view on MEI vs. MEIYOU).⁷⁶

Both MEI and MEIYOU can be used as the negative form of the lexical verb YOU, to negate the possession or existence of the NP complement in a canonical declarative sentence, as in the example sentence (24a); they can also take a VP complement, as in (24b):

- (24) a. wo mei(you) zhe ben shu.
 1SG not.have this CLF book
 I don't have this book.
- b. wo mei(you) mai zhe ben shu.
 1SG aux.NEG buy this CLF book
 I did not buy this book.

MEI/MEIYOU is often analyzed in comparison with BU, another negation morpheme in Standard Mandarin. Much controversy remains over the semantic functions and/or restrictions of MEI/MEIYOU on its verbal complements. According to Ernst (1995), MEI/MEIYOU and BU differ in terms of their aspectual selectional requirements: while BU aspectually selects a stative situation as its complement, the preverbal MEI or

details). As a negation morpheme, however, MEI can be used interchangeably with MEIYOU in almost any context, except that in an A-not-A question, the “not” element can only be MEI. At the end of the VP-NEG question, the element is usually MEIYOU. At the end of VP-NEG questions, the NEG element is usually MEIYOU, but if a sentence final particle is present, MEI can also be used the negation morpheme (Lü, 1985). This might be explained by factors other than the semantics of these two morphemes: phonological constraints might be one of such factors, as the “not” element can be either MEI or BU, both being monosyllabic words. An answer in the negative to the A-not-A question can be either MEI or MEIYOU. The use of MEI before a sentence-final particle in the VP-NEG questions might also be attributed to such phonological constraints. The use of MEI before a sentence final particle as the negation morpheme in VP-NEG questions might also be attributed to such phonological constraints. However, evidence from research is needed to provide evidence for my speculation.

⁷⁶ Lü (1985) considered MEIYOU as a variant of MEI.

MEIYOU, which is usually used as the positive counterpart of the perfective aspectual marker LE (see Shi & Li, 2001; Wang, 1965), selects an event as its complement.

There are other competing theories about the semantics of these negation morphemes: for instance, Xiang (2014) posits that the specifier of the syntactic projection AspP introduces an event variable closure, which is either existential or generic. BU is licensed by the semantic operator *Gen*⁷⁷ in sentences with a generic, habitual or universal reading, and is positioned below the AspP, whereas MEI is positioned above the AspP and takes scope above the existential or generic closure.⁷⁸ This implies that MEI/MEIYOU goes beyond the function of marking perfectivity.

Aside from the debate on the differences between the two negation morphemes (MEI/MEIYOU vs. BU), the existing literature agrees on the function of the preverbal MEI/MEIYOU in the marking of aspect (which might vary across contexts), and on the close association between MEI/MEIYOU and YOU-MEI-YOU in terms of functions, the former being the negative answer to a question containing the latter. Hence, it is reasonable to postulate that the grammaticalization processes of MEI/MEIYOU can inform us of the grammaticalization of the preverbal YOU-MEI-YOU.

Shi & Li (2001) tracked the grammaticalization of MEI and MEIYOU that spanned a long period of time (about 1,900 years for MEI and 700 years for MEIYOU), and put forward a hypothesis about the syntactic environments that have triggered their grammaticalization of MEI and MEIYOU: in their theory, the modern construction

⁷⁷ *Gen* is a quasi-universal quantifier.

⁷⁸ In my personal communication with the author, she confirmed that she followed Wang (1965) and Ernst (1995) and treated MEI as MEI(YOU) and that this analysis also applies to MEIYOU.

Subject+MEI+Verb+Object has developed from a syntactic environment structured like this: Subject+[MEI+NP] + [Verb+Object] (p. 250).⁷⁹

However, Shi & Li (2001) did not elaborate on the syntactic and semantic properties of the environments in which MEI/MEIYOU have been grammaticalized into negation morphemes; nor did they provide much detail about the stages of development of these morphemes, especially the later stages at which these two have taken on the function of negating the predicate of the sentence, except for the discussion on the change in their co-occurrence and/or compatibility with the aspectual marker LE based on a classic piece of literary work in Standard Mandarin.

The same problem was found in research on the positive form of the aspectual marker MEIYOU, namely YOU ‘have’, ‘there is/exists’.⁸⁰ Evidence, though scarce, that a preverbal YOU, has begun to emerge, especially in Standard Mandarin spoken by certain regional speakers, was found by some researchers (Diao, 2012; Zhou, 2012), but none has provided a detailed account of the development of this morpheme. This is what motivated me to look at grammaticalization research, in particular studies on constructions equivalent to a preverbal YOU, in other languages (e.g. *have* in English), as discussed in Section 2.2.3. It is hoped that knowledge gleaned from an examination of the general parameters proposed for measuring the degree of grammaticalization and the trajectories of developments of such foreign counterparts will facilitate the identification

⁷⁹ In the S+[MEI+NP]+V+O structure, MEI is part of the serial verb construction and takes an NP as its object. Shi & Li (2001) did not discuss what happened to the NP when this structure developed into the S+MEI+V+O structure.

⁸⁰ YOU (‘have’, ‘there is/exists’) is commonly used as a lexical verb, but when it occurs before a verb phrase could be considered as the positive counterpart for MEIYOU and preverbal YOU-MEI-YOU.

of the developmental stages and features of the preverbal YOU-MEI-YOU in Standard Mandarin.

In this section, I presented works by some important scholars on aspect, a key notion in the categorial assignment or definition of preverbal YOU-MEI-YOU, introduced the definitions and classification systems of aspect, the relation and interaction between different types of aspect, and mention one kind of aspect—perfective aspect in Standard Mandarin.

In the last section, I will sketch some typological features of Standard Mandarin, and summarize the methods in some important works on words and word class categorization in Standard Mandarin and discuss the ways the methods proposed in these studies are useful as well as problems with or challenges of using such methods in the current data analysis.

2.4 Word-class Categorization in Standard Mandarin

Chinese is morphologically highly impoverished. In addition, Lin (2012) listed some other syntactic properties of Standard Mandarin, including the existence of bare nominal predicates, lack of expletive subjects, lack of finite vs. nonfinite distinction, and lack of case-motivated movement.^{81,82} These typological characteristics of Standard Mandarin have made it challenging for Chinese linguists to classify words in terms of their parts of speech. The difficulty has been heightened by other factors such as the flexible word order and the pro-drop nature of the language that allows the omission of

⁸¹ In Chinese, nouns and adjectives can be put in the predicate position of a sentence, without a copula.

⁸² Since the finite vs. non-finite distinction is not relevant in the Chinese context, the part of Krug (2011)'s definition of auxiliaries concerning the (non) finite status of the (de)verbal complement will not be considered when data are analyzed to test the hypothesis in this study.

subject or object pronouns which can be pragmatically inferred. For decades, Chinese linguists have been grappling with this problem and concepts from mathematics, logic and linguistics such as distributional weight, set, prototypical theory, and degree of membership have been applied. Most of the existing words can be successfully categorized in the theoretical frameworks proposed in the influential works on this topic (Guo, 2002; Yuan, 2010); however, controversy still remains on the criteria and methods of classification (for a review, see Guo, 2002; Yuan, 2010).

One of the most important works on word class classification in Standard Chinese is Yuan (2010), whose method was based on mathematical concepts, logic and linguistic notions. He believed a system built on these concepts would be most reliable and practical for the classification of word classes of Standard Mandarin that were in nature fuzzy clusterings with no clear boundaries in terms of denotation and extension (Yuan, 2010, p. 118).

This method is similar to Heine's (1993) approach to linguistic categorization, which is based on family resemblance categories. With Yuan's (2010) method, each word can be put into the category/categories it fits best depending on how similar its behavior is to that of other members in the group(s). To evaluate the behavior of a word in relation to other words, Yuan chooses a set of distributional features/frames based on the grammatical behavior of the typical members of a category and uses this set of frames to determine the membership of a particular word. Each distributional feature is given certain weight, and depending on whether a word has the feature or not, the weight can be positive, neutral/zero, or negative. The choice of distributional features for each word-class and the weight given to each feature are built on empirical data: each set of

distributional features is chosen to maximally set apart one particular word-class from another and the weight can be different for each distributional feature in the same set that is shared to a varying degree by members in the same category (word-class).

This method of word-class classification is effective in alleviating the long-standing problem in Chinese, a language that cannot rely on most of the traditional morphosyntactic tests (e.g. inflections) for describing and/or determining the syntactic categories of its lexicon.

However, this method has a major flaw: To determine whether and to what extent a word belongs to a certain word-class, one needs to take it out of its original context, put it into a set of frames, and to test how well it fits in the distributional frames selected for that word-class/category. This means, if a word belongs to more than one class or part of speech, it is not possible to determine its categorial membership in a particular context, as the evaluation occurs outside of the original context. This problem is exemplified by words having dual categorial membership like KENENG ‘is possible’, ‘possibility’ in Standard Mandarin. I will elaborate on this in the chapter on methodology.

In light of this, the ideas proposed in Guo (2002) will be used to complement Yuan’s (2010) method. In Guo (2002), the issue of a word having an ambiguous membership in a certain position was addressed in the section on “nominal verbs and adjectives” or verbs and adjectives found in the object position or after a modifier (pp. 165-172). Guo maintained that there were two possibilities for this kind of word: (i) the word was nominalized at the syntactic level (in a particular context), but remained a verb or an adjective at the lexical level; and (ii) the word, which could be used as a verb/adjective, exhibited nominal features in the object position and should be treated as

having a dual membership (Verb/Adjective & Noun) at the lexical level. Guo's conclusion was based on some syntactic tests (e.g. adding the intensifier HEN 'very') directly applied in the context where the word was found whose syntactic category was difficult to determine.

Guo's (2002) approach of evaluating the syntactic status of a word in its local context offers an alternative way of examining the contextual behavior of a word, which is not possible within Yuan's (2010) analytical framework, and will be used to analyze data that the latter could not handle. It will be applied to determine the syntactic category of words with dual/multiple memberships.

Nonetheless, the approach used in Guo (2002), like Yuan's (2010) method, and other existing methods, focused on words in static, isolated contexts and have not adequately addressed those that appear in new contexts in which the manipulations/tests proposed in their frameworks cannot be applied for other (i.e. semantic or pragmatic) reasons. The problems are most apparent in the following scenarios: (i) a word belonging to more than a class or part of speech is found in a context in which the tests for both classes seem to be applicable, and (ii) a word belonging to an established class is undergoing a syntactic or semantic change and is taking on a new identity in certain new contexts while still keeping its old identity in other contexts.

When the tests proposed by these authors are applied to YOU-MEI-YOU, they are inadequate for determining its structural position within the question. Most conventional tests for verbhood also fail with this construction. For instance, the A-not-A rule cannot be applied to it, it cannot form imperatives, and cannot be negated. In light of this, syntactic tests from other studies (Li & Thompson, 1989; McCawley, 1992) are

employed to show that preverbal YOU-MEI-YOU behaves in a distinct way from regular VPs and non-verbal phrases that are also found to appear before VPs—e.g. adverbial phrases.

One test Li & Thompson (1989) proposed for differentiating auxiliary verbs and regular verbs is the VP-ellipsis test. According to Li & Thompson, “[a]n auxiliary verb must co-occur with a verb (or an ‘understood’ verb)” (p. 173). That means that when the V’ complement of an auxiliary is understood from context, the V’ complement can be omitted. McCawley (1992) compared the auxiliary verb NENG ‘can’ with the subclass of verbs that “take a surface complement V’ such as *dǎsuàn* ‘intend’” (p. 212), and concluded that “the sort of context that licenses the omission of the complement of *néng* does not license the omission of the complement of *dǎsuàn*” (p. 213). Since adverbial phrases in Chinese also cannot appear without the main verb or VPs they are attached to, this VP-ellipsis test is adopted here to justify the claim that preverbal YOU-MEI-YOU has auxiliary characteristics that makes it distinct from regular verbs and from preverbal adverbs.

The following sentences show the three-way contrast between preverbal YOU-MEI-YOU, the verb DASUAN ‘intend’, and the adverb FEICHANG ‘very’:

- (25) zhangsan mai-le na ben shu. lisi you-mei-you ne?
 Zhangsan buy-LE that CLF textbook. Lisi have-not-have NE
 Zhangsan bought that book. What about Lisi/did Lisi buy it, too?
- (26) zhangsan dasuan mai na ben shu. *lisi ye dasuan.
 Zhangsan intend buy that CLF book. Lisi also intend
 Zhangsan intends to buy that book. Lisi also intends to do so.

(27) zhangsan feichang xihuan qiaokeli. *lisi ye feichang.

Zhangsan very like chocolate. Lisi also very

Zhangsan likes chocolate very much. Lisi also likes it very much.

This test shows that preverbal YOU-MEI-YOU has auxiliary characteristics and can be differentiated from regular verbs like DASUAN. However, it is difficult to find tests showing that it has verb-like properties, for the aforementioned reasons. It cannot take aspect markers, like modal verbs (or what Li & Thompson (1989) call auxiliaries), although McCawley (1992) argues that the possibility of adding aspect markers “can be used only as a positive test for verbhood” (p. 213). This means that, “if an aspect marker can be added to something, it is a verb, but nothing can be concluded about an item’s part of speech from the fact that they can’t be added to it” (p. 214). The ways these classification methods are employed in this study will be discussed in more detail and the problems or challenges will be elaborated in the next chapter.

2.5 Chapter Summary

In this chapter, I pointed out that previous studies on preverbal YOU-MEI-YOU did not give enough attention to the status of this construction in the Chinese grammar, and also failed to give an adequate account of the historical development of this construction. I discussed the major problems with the methodologies and/or theoretical arguments made in these studies concerning the two issues of interest in the current study—i.e. the grammatical status and the historical development of preverbal YOU-MEI-YOU, respectively.

After presenting the reasons for conducting the current project, I gave an overview of grammaticalization research, focusing on how approaches to linguistic categorization, and Heine's (1993) definitions of auxiliaries motivate my first research question, and on the different parameters and features associated with developmental stages of grammaticalized forms that may be relevant to the analysis of the target construction. I also briefly mentioned the studies on grammaticalization of a certain type of auxiliary—i.e. possessive perfects or perfective, which is considered functionally similar to preverbal YOU-MEI-YOU. I discussed how the methodologies used in the cross-linguistic studies on possessive perfects or perfectives may be applied in the current study.

Lastly, I provided an overview of studies on aspect, a notion essential to defining the grammatical status of preverbal YOU-MEI-YOU, ways of determining aspectual meaning, and one type of aspect morpheme in Standard Mandarin. I concluded the chapter with a short discussion of typological features of Standard Mandarin that may pose a challenge to the categorization of words and word classes, the methods proposed for word class categorization in Standard Mandarin as well as problems with these methods.

In the following chapter, I will introduce the sources from which I collected my data and the procedures for sorting the data prior to data analysis involving the methods for categorizing words proposed in Yuan (2010) and Guo (2002). I will discuss the challenges with and the reasons for this kind of sorting, and how the data will be analyzed to answer the three research questions in this study. The analysis will draw on Heine's (1993) and Krug's (2011) definitions of auxiliaries and begin with an

examination of the grammatical meanings, in particular the aspectual meaning, conveyed by YOU-MEI-YOU to answer the first research question. I will conduct an analysis of the semantic and syntactic features of preverbal YOU-MEI-YOU and of the types of VP complements it takes in different historical periods. The analysis will be based on the parameters and linguistic tests put forward in Heine (1993) and cross-linguistic studies on the development of possessive perfects, to answer the second research question. The frequency of use of this construction will be approached by plotting its temporal and regional distributions, in order to answer the third research question.

3. METHODOLOGY

In this chapter, I will introduce the sources from which the data have been collected and the ways the data will be classified and analyzed to answer the research questions of the current study. First, I will give an introduction of the corpora used in the study, in particular the types of data and information available in these corpora. In the second and third sections of the chapter, I will outline how the data will be filtered, including the principles and methods used in the filtering. In the last section, I will discuss the particular ways the data will be analyzed to answer each of the three research questions.

3.1 Data Collection

Previous studies on grammaticalization have shown that it is most beneficial to examine data from corpora that allow researchers to trace the origins of instances of a grammaticalized form and the development of this form in different contexts. In this light, data were collected from two major Chinese databases that are publicly available.⁸³

⁸³ In one of the databases, only about 21 million words of data are made publicly available (for search) for one of its two corpora.

The two databases from which data were taken are the Center for Chinese Linguistics Peking University database (henceforth referred to as CCL, Center for Chinese Linguistics PKU, 2009), and State Language Commission Chinese National Corpus (henceforth referred to as CNC, State Language Commission, 1993-2012). These two databases provide corpora from both written and spoken text, with the latter taking up only a small portion in both cases. Both databases provide at least partial information about the sources of the entries in the search results (e.g. author, year of publication) when a search is conducted. Such information has made it possible to estimate the production/publication date for each entry. The entries for which the production/publication date could not be obtained were not included in the analysis.

I will introduce the corpora in each of these databases separately in the following sections. In particular, I will outline the time span of the data, the amount of data that is publicly available, and the types of information available in the corpora of the two databases.

3.1.1 The CCL Database

The CCL database spans about 3,000 years (11th c. BCE to 21st c. BCE), beginning with one of the earliest dynasties recorded in Chinese history and ending in 2013. It is divided into two sections: Ancient Chinese (hereafter referred to as AC) corpus and Contemporary/Modern Chinese (hereafter referred to as CC) corpus. The CCL-AC corpus covers the period from the Zhou Dynasty to the Minguo or the Republic of China period (1912-1949); the CCL-CC corpus spans the post-Minguo or New

China period (from 1949 to the early 21st century).⁸⁴ The CCL-AC corpus contains 201,668,719 characters, all from written text; the CCL-CC corpus contains 581,794,456 characters, about 97.66% of which are taken from written text.⁸⁵ Table 2 summarizes the basic information of the CCL database.

Table 2 Summary of the CCL database.

Database	Corpus	# of Characters
CCL	AC	201,668,719
CCL	CC	581,794,456

The search results in the CCL corpora are aligned with a conventional token-level alignment method. This means that, when a search is performed, sometimes only the immediate clause containing the keyword(s) (even if part of a longer sentence) will be displayed. In addition, the sentences—often the paragraphs, preceding and following the clause containing the keyword(s) can be retrieved and read by clicking on the “co-text” icon to the right side of each search result, although they will not be displayed as part of the search results. Nevertheless, the CCL database does not provide the whole source text (e.g. an article, a book). The CCL database allows users to download the search results as text documents, which include partial data about source information for each search result (e.g. name of the article/book, name of the author).

⁸⁴ In fact, writings from the early and later Minguo Period were also found in the CCL-CC corpus. Therefore, the CCL-CC corpus spans at least most of the Minguo period and the New China period.

⁸⁵ According to Wikipedia (“Character (computing)”, 2016), in computer and machine-based telecommunication, a character refers to a unit of information; it can be a printed or written letter, number, symbol, or a word in Chinese.

Data were collected from the CCL database by searching the two corpora with the key word YOU-MEI-YOU (corresponding to the Chinese characters 有没有 without a space in between). The search was performed in September 2013 and two text documents were obtained, one from the CCL-AC corpus and the other from the CCL-CC corpus. A total of 150 results were located in the CCL-AC corpus and 6,479 in the CCL-CC corpus.

3.1.2 The CNC Database

Similar to the CCL database, the CNC database is also divided into an Ancient Chinese corpus and Contemporary Chinese corpus. According to the introductory text of the database, the CNC-AC corpus, with about 100 million words, spans the period from the Zhou Dynasty (11th to 3rd c. BCE) to approximately the Qing Dynasty (1616-1912); the CNC-CC corpus covers the period from 1919 to 2002.⁸⁶ It contains about 100 million characters, about a fifth of which are available for search (approx. 19.5 million).^{87,88} As stated in the introductory text, the majority of the corpora in CNC are taken from written texts and the small portion of spoken corpus in this database contains coherent and clear

⁸⁶ New data are being added to the database every year after 2002, but the bulk of the data in the CNC-CC corpus are taken from period of 1919-2002. In addition, the data that are searchable by the outside users in the CNC-CC corpus span the period from 1919 to 1998, and data from certain years were not available for search, as I found out through my personal communication with the researcher who led the team in compiling this corpus.

⁸⁷ The introduction to the two corpora on the official website uses different wording to describe the size of each one: the CNC-AC corpus is said to have over 100 million “words” while the CNC-CC corpus is said to have about 19 million “characters”.

⁸⁸ However, this was slightly different from the information I obtained through my personal communication with the researcher who lead the team in compiling this corpus: the total number of characters in the texts used in this corpus, as listed by the researcher, was about 21.5 million.

speech that can be easily transcribed (e.g. public speeches, scripts). The basic information for the publicly searchable portion of the CNC database used is summarized in Table 3.⁸⁹

Table 3 Summary of the CNC database.

Database	Corpus	# of Characters
CNC	AC	about 100 million
CNC	CC	19,455,328

The CNC database is slightly different from the CCL database in terms of how search results are displayed: when a search is conducted in the CNC corpora, the whole sentence in which the keyword(s) is/are located, short or long, will be displayed with the keyword(s). However, the CNC database does not provide the clauses preceding and following the sentence that contains the keyword(s) or the whole source text (e.g. an article, a book). Like the CCL database, the CNC database also allows users to download the search results as text documents, although the text document obtained from the CNC-CC corpus only has the sentences containing the keyword(s) and the one from the CNC-AC corpus contains the name of the book/article in addition to such sentences. The information about the source texts from which these sentences are taken (e.g. name of the article/book, name of the author) can only be viewed online on both of the CNC corpora.

Data were collected from the CNC database in September 2013, by searching the two corpora with the same keyword, i.e. YOU-MEI-YOU (有没有). In the CNC database, the search produced 12 results in the AC corpus, and 388 in the CC corpus.⁹⁰

⁸⁹ For the sake of convenience, “character” instead of “word” is used in the table and the readers should keep in mind that the CNC-AC corpus has over 100 million words rather than characters. In addition, the number of characters available for search in the CNC-CC corpus as stated in the website is listed here, although the corpus actually has over 100 million words, as stated in the official introductory text.

3.2 Data Processing: Initial Filtering

The entries obtained from the two databases were filtered to locate clauses eligible for the purpose of data analysis in this study. Once the raw data were collected, the search results were filtered using the procedures and principles explained below and in the following section.

Before being analyzed, the data from the corpora were first filtered according to the following criteria: (i) the entry should be interpretable and is a faithful reproduction or transcription of the original text as identified in the source information; (ii) the keyword in each entry forms one unit (with YOU-MEI-YOU appearing in a sequence, followed by an NP, VP, AP or any other type of constituent); and (iii) the entry is produced in Mandarin Chinese by native speakers/users of Chinese. The first principle required that caution be exercised and suspicious entries that seemed to contain some kind of error (e.g. typo, mis-transcription) be compared with the source text, if the latter was available. The second principle was set up to eliminate entries containing fragments or parts of different constituents whose accidental adjacency produced what looked like the target keyword (i.e. YOU-MEI-YOU used as one unit). The main purpose of the third principle was to get rid of nonnative speech and writings, including speech produced in certain dialects (e.g. Cantonese) that has been transcribed in Mandarin Chinese.

Apart from these major principles, some other procedures were applied in the filtering of the data in order to eliminate unnecessary ambiguity and confusion. For instance, if an entry was found in the AC corpus and was labeled as part of an Ancient

⁹⁰ The initial search in the CNC-AC corpus produced a list of 3,317 results, but upon a closer look, only 12 of them actually contained the three words, i.e. YOU, MEI, YOU, in a sequence.

Chinese text, but was written in a distinct Modern Chinese style, it would be discarded, as it was most likely to have been taken from modern writings interpreting and/or commenting on the original Ancient Chinese text. Another filtering procedure involved the identification of the same entries appearing more than once within a corpus and/or across corpora. Such an entry, if found, would be used and counted in analysis only once.

These filtering procedures produced 124 (out of a total of 150) entries in the CCL-AC corpus that contained a YOU-MEI-YOU and 6,218 (out of 6479) entries containing a YOU-MEI-YOU in the CCL-CC corpus. Only three (out of 12) entries in the CNC-AC corpus, and 379 (out of 388) in the CNC-CC corpus remained after the filtering.⁹¹ The results of the filtering are listed in Table 4.

These entries then underwent another type of filtering, in order to be usable for the kinds of data analysis needed to answer the research questions of the current study. The next subsection introduces how the data were further processed, including the factors considered and the methods used.

Table 4 Data distribution before and after initial filtering

Database	Corpus	# of Entries Found in Search	# of Entries after Initial Filtering
CCL	AC	150	125
CCL	CC	6479	6221
CNC	AC	12	3
CNC	CC	388	379

⁹¹ 9 out of the 12 entries/sentences in the CNC-AC corpus are also found in the CCL-AC corpus and are counted only once, i.e. included in the 124 entries (after initial filtering) for the CCL-AC corpus. 1 entry found in the CNC-CC corpus is also found the CCL-CC corpus and is counted only once, i.e. included in the 6222 entries of the CCL-CC corpus.

3.3 Data Processing: Second Filtering

Since the research questions of this study concern the status and development of preverbal YOU-MEI-YOU, a prerequisite to data analysis is to identify and exclude entries in which YOU-MEI-YOU is followed by a non-verbal complement.⁹² Therefore, during this round of filtering, the methods for categorizing words and word classes in Mandarin Chinese introduced in the previous chapter were employed to eliminate instances that do not contain the target construction.

As discussed in the previous chapter, Yuan's (2010) method was effective in assigning words into categories to which they bear the most resemblance in most cases, but it was not very helpful in dealing with words with dual categorial membership or those whose categorial membership remains to be settled—i.e. those undergoing a shift in status. One such difficult case concerned the word/phrase, KENENG ('possible'). This word belongs to more than one class: when appearing after a verb, a quantifier or a modifier, it is a noun meaning 'possibility'; opinions diverge, however, as to the classification of this word when it appears between the subject and the main verb. In fact, this word was categorized as an adverb in some studies (e.g. Guo, 2002; Li & Thompson, 1989) and as a modal (verb) in others (Huang, Li, & Li, 2009; Lin & Tang, 1995; Yuan, 2010; Zhu, 1982). Analyzed in Yuan's (2010) framework, KENENG per se has a dual membership: Noun and Auxiliary.

When KENENG was used in a context involving some ambiguity, Yuan's (2010) method became irrelevant. For instance, this method was of little help in determining the

⁹² In this study, non-verbal complements refer to phrases other than verb phrases, adjective phrases and prepositional phrases. See footnote 56 for more details.

categorial membership of KENENG ('is possible', 'possibility') found before a verb and after MEI(YOU) 'not have', which could be viewed either as the negative counterpart of the main verb YOU 'have' or an aspect marker (see Examples 24a-b for the two functions of MEI/MEIYOU). The following sentences illustrates the scenario:

- (28) a. hounian genben mei(you) keneng wancheng zhibiao.⁹³
 next.year at.all aux.NEG possible complete quota
 (We) could not possibly meet the target in two years.
- b. hounian genben mei(you) keneng wancheng zhibiao.
 next.year at.all not.have possibility complete quota
 There is no possibility that the target can be met in two years.

The same problem arose with KENENG found before a YOU-MEI-YOU, which could be seen as the A-not-A construction of the main verb YOU or as an auxiliary, as hypothesized in this study. This can be illustrated with the following sentences:

- (29) a. shouji diu le, you-mei-you keneng zhao huilai?⁹⁴
 cellphone lose LE⁹⁵ have-not-have possible find return
 (My) cellphone was lost. Is it possible to get it back?
- b. shouji diu le, you mei you keneng zhao huilai?
 cellphone lose LE have NEG have possibility find return
 (My) cellphone was lost. Is there any possibility that I can get it back?

⁹³ This example was taken from the Internet.

⁹⁴ This example was taken from the Internet.

⁹⁵ In this sentence, LE can be a perfective marker or a sentence-final particle. Therefore, there is no hyphen connecting the verb DIU 'lose' and LE.

A quantifier YIDIAN ‘a little’ can be inserted between YOU-MEI-YOU and KENENG in (29b) without making the sentence ungrammatical or changing its core meaning.

Thus, the categorization based on Yuan’s (2010) method found itself in a dilemma: the syntactic category of YOU-MEI-YOU needed to be settled so as to determine that of the complement which was, in itself, information necessary for determining the syntactic category of YOU-MEI-YOU. The diagnostic tools proposed in Yuan (2010) and by some other scholars (e.g. Li & Thompson, 1989) could not be applied to directly test the status of YOU-MEI-YOU in this type of sentence for syntactic and semantic reasons. For instance, one common test for an auxiliary in Standard Mandarin is to see if the word in question can occur as the A element in an A-not-A question (Li & Thompson, 1989, p. 172). It is obvious that this test would not work in the case of YOU-MEI-YOU, which has developed from the A-not-A question. Another common test of verbhood is the negation test. However, when preverbal YOU-MEI-YOU is used with a Negative Polarity Item (e.g. RENHE ‘any’), the resulting construction is grammatical, as in the following example:

- (30) ni you-mei-you ting-dao renhe shengyin?
 2SG have-not-have hear-RVC any sound
 Did you hear any sound?

Therefore, something in (30) must be licensing the NPI and the only lexical item that could do this is preverbal YOU-MEI-YOU. Since preverbal YOU-MEI-YOU has a NEG feature, the negation test cannot be applied.

Presumably, indirect evidence could be obtained from tests applied to other related form(s), i.e. YOU and MEI/MEIYOU; however, one should bear in mind that the

positive form (YOU), still in the incipient stage of development towards auxiliarness best fits into the category of verbs at the current stage while the negative form MEIYOU can be categorized as either a lexical verb or an auxiliary in Yuan's (2010) framework.

In order to solve this problem, I turned to the method discussed in Guo (2002), and applied some syntactic tests in the original contexts in which KENENG and other words with dual categorial membership were found. This approach proved to be helpful in some cases: for instance, with Guo's approach, it was decided that KENENG found between YOU-MEI-YOU and a VP was best categorized as a noun in that particular context.⁹⁶

It should be pointed out that even with Guo's method, there still remained some uncertainty. The complement of YOU-MEI-YOU in some clauses could be analyzed as either a NP or a VP, even when assessed in the original contexts. One such example was taken from the CNC-CC corpus:

- (31) a. fushipin gongying qingkuang you-mei-you gaishan?
 nonstaple.food supply situation have-not-have improve
 Has the supply of non-staple foods improved?
- b. fushipin gongying qingkuang you mei you gaishan?
 nonstaple.food supply situation have NEG have improvement
 Is there any improvement in the supply of non-staple foods?

As shown in the example sentence (31), the sentence could be analyzed in two ways, due to the indeterminacy in the syntactic category of the word GAISHAN

⁹⁶ The test applied involved the use of a quantifier, DIAN/YIDIAN ('a little') or YIXIE ('some') before KENENG in the sentence. The resulting sentence was still grammatical and the meaning was not changed to a great extent.

(‘improve’/‘improvement’), which could be used either as a noun or an intransitive verb in other contexts. Guo’s (2002) approach was not effective in this case, either: on the one hand, the word seemed to fit in the noun category, as a quantifier (e.g. DIAN/YIDIAN ‘a little’, YIXIE ‘some’) could be inserted before it in this sentence. On the other hand, it was also possible to insert in front of GAISHAN an adverb typically used to modify a verb (e.g. XUNSU ‘quickly’ or ‘soon’), without making the sentence ungrammatical. I will discuss such ambiguous cases and how they inform the analysis and the research questions in the next chapter. Suffice it to say here that such instances should be distinguished from sentences that contained the target construction, namely YOU-MEI-YOU+VP and put into a separate group.

After applying the aforementioned analytical procedures, the remaining entries were categorized into three types: YOU MEI YOU+NP, YOU-MEI-YOU+VP (including Verb Phrase, Adjective Phrase, and Preposition Phrase) and YOU-MEI-YOU+Ambiguous. Among these, the first type of entry was discarded and the last two types were used in the data analysis. These three types of clauses are illustrated by the following examples:

(32) you mei- you duixiang?

have NEG have boyfriend/girlfriend

(Do you) have a boyfriend/girlfriend?

(YOU MEI YOU+NP)

(33) zuotian ni zai hai shang you-mei-you jian dao liang ge kaocha

yesterday 2SG at sea above have-not-have see-RVC two CLF expedition

duiyuan?

team.member

Did you see two members of the expedition team on the sea yesterday?

(YOU-MEI-YOU+VP)

(34) a. ta de zuoyong you-mei-you gaibian ne?

3SG DE⁹⁷ function have-not-have change NE

Did its function change?

b. ta de zuoyong you mei you gaibian ne?

3SG DE function have NEG have change NE

Is there any change in its function?

(YOU-MEI-YOU+Ambiguous)

After the entries have been categorized according to the type of complement YOU-MEI-YOU takes, the data from the YOU-MEI-YOU+VP and YOU-MEI-YOU+Ambiguous categories underwent another filtering procedure in which sentences from translated works were singled out. These sentence were also excluded from further analysis due to potential translation effects, since the possibility could not be ruled out that such sentences might have been produced with the motivation of mimicking speakers of the source languages that are not Chinese or that their production could have been influenced by the source language(s) and thus do not truly reflect the grammar of Standard Mandarin.

⁹⁷ The morpheme, DE (的), in this sentence is used to mark possession, and is similar to the apostrophe s ('s) in English in terms of function.

In addition, entries for which the year of publication/composition could not be determined were also discarded, since this piece of information was vital when the data was entered for further analysis.

A total of 991 entries remained after the second round of filtering. These entries, including the ones with YOU-MEI-YOU+Ambiguous were recombined into one big group as presented in Table 5 and were further analyzed in terms of the grammatical meaning conveyed.⁹⁸ The decision to include the ambiguous cases in this step was based on the consideration that a verbal interpretation of the complement of YOU-MEI-YOU was possible in such entries, meaning that they could be used as evidence for the development of YOU-MEI-YOU towards auxiliarihood.

It was observed that in some cases, one search result contained more than one instance of YOU-MEI-YOU+VP or YOU-MEI-YOU+Ambiguous. In order to have a comprehensive understanding of the use of this construction, all the clauses that were found to contain a preverbal YOU-MEI-YOU would be used in the data analysis. A total of 1,023 such clauses were found in the two databases. Table 5 presents the number of clauses in each corpus that were used for data analysis:

The next subsections explain how the data were further classified and/or analyzed in order to answer the three research questions. I will first introduce the tests for the aspect-related analysis of the data, which is crucial to addressing the first research question (and the other two research questions). I will also discuss how the complements

⁹⁸ A finer-grained analysis and discussion of the YOU-MEI-YOU+Ambiguous entries, as opposed to the ones with YOU-MEI-YOU+VP, will be conducted in the next chapter.

Table 5 Data distribution after second filtering

Database	Corpus	#of Clauses after Second Filtering
CCL	AC	1
	CC	979
CNC	AC	0
	CC	43

of YOU-MEI-YOU found in the 1,023 clauses were classified and analyzed (a step necessary for answering the second research question). Lastly, I will describe how the distribution of data was presented to answer the third research question.

3.4 Data Analysis

In this section, I will discuss how data were analyzed to answer each of the three research questions and introduce the theoretical and methodological tools that were used in the different types of analysis.

3.4.1 Data Analysis to Determine the Formal Status of Preverbal YOU-MEI-YOU

In order to answer the first research question, I inspected the semantic and syntactic functions of preverbal YOU-MEI-YOU in the data and tried to determine if it served to convey any of the grammatical meanings typically associated with the use of auxiliaries as outlined in Heine (1993) and Krug (2011). To this end, the data that had undergone the two prior rounds of filtering were first coded in terms of two grammatical

meanings expressed by the grammatical construction preverbal YOU-MEI-YOU helped to form: sentence types and aspectual meaning(s).

Since it has been observed that preverbal YOU-MEI-YOU appears in perfective questions, the clauses were coded as having sentence types of Interrogative or Non-Interrogative, and as having viewpoint aspect of Perfective or Non-Perfective.⁹⁹ The percentage of clauses in which preverbal YOU-MEI-YOU conveyed either type of grammatical meaning was used to support the claim about the auxiliary status of preverbal YOU-MEI-YOU. The coding for sentence type is illustrated by examples (35-36): (35) is a clause coded as Interrogative while (36) is coded as Non-Interrogative. The coding for aspect is illustrated by examples (37-38): (37) is coded as Perfective and (38) Non-Perfective.

(35) na nimen de gongzuo you-mei-you shou yingxiang?
 then 2PL DE¹⁰⁰ work have-not-have encounter influence
 Was your work affected?

(36) ta meitian dou hui shike zhuyi na xie wan you-mei-you
 3SG every.day all will often take.note that CLF bowl have-not-have
 xi-hao
 wash-RVC

Every day, he would pay close attention to whether the bowls were washed.

⁹⁹ Some studies make “a distinction between *interrogative* as a syntactic term and *question* as a semantic or pragmatic term” (Ceong & Saxon, 2013, p. 2). In this study, the coding was based on both syntactic and semantic grounds. Direct questions were coded as Interrogative and non-direct questions (e.g. relative clauses that contain preverbal YOU-MEI-YOU) were coded as Other. A detailed discussion is offered in this section.

¹⁰⁰ The morpheme DE is used to mark possession in this sentence.

(37) na ni dangshi you-mei-you xiang-dao shi feidian?

then 2SG at.that.time have-not-have think-RVC be SARS

Did it occur to you at that time that it was SARS?

(38) ni you-mei-you juede wo he yiban ren bu yiyang?

2SG have-not-have think 1SG and common person NEG same

Do you think I am different from ordinary people?

The decision to mark clauses containing preverbal YOU-MEI-YOU as Interrogative or otherwise was based on the differences between direct questions (appearing in the form of independent clauses) and indirect questions or embedded questions (being part of another clause) in Standard Mandarin (Lü, 1982; Shao, 1996; Tang, 1981, 1984) and in Korean (Ceong & Saxon, 2013). Given that preverbal YOU-MEI-YOU was found in both main/independent clauses and embedded clauses and in different discourse contexts, it is necessary to distinguish instances in which this construction is used to help form a direct question and those in which it is found in an indirect question.¹⁰¹ This is important because in direct questions, preverbal YOU-MEI-YOU, being a V-not-V form, marks Yes-No interrogatives in Standard Mandarin (Shao, 1996; Zhu, 1982), whereas in indirect questions, the clause that immediately contains it may be turned into a complement clause, depending on the main verb of the matrix clause, and the presence or absence of preverbal YOU-MEI-YOU will not change this fact.

¹⁰¹ Following Ceong & Saxon (2013), direct quotations parentheticals are not included under the term “embedded clauses” in this study. Punctuations (e.g. direct quotation marks, colon), the pronouns (referring to the addressee) and other cues in the original context(s) are used to tell such instances from embedded clauses defined in a narrow sense. The difficulty in identifying the boundary between them should be kept in mind.

Drawing on Shao (1996) and Tang (1981), a clause that immediately contains a preverbal YOU-MEI-YOU would be coded as Interrogative if (i) it was an independent clause, or (ii) appeared after what was called a “direct question verb” in Tang (1981) or a “question verb” in Lü (1982), Shao (1996)—for example, *WO WEN NI* (roughly equivalent to ‘I have a question for you’), *(WO) QING NI GAOSU WO* (‘please tell me’); (iii) it appeared after verbs meaning guessing or speculating with the second person pronoun as its subject (e.g. *NISHUO* ‘in your opinion’, *NICAI, NI RENWEI*, ‘(you) guess’); or (iv) it appeared after expressions like *BUXIAODE, BUZHI(DAO)* (roughly translated as ‘not sure’). The next chapter will discuss the sentence-type marking function of preverbal YOU-MEI-YOU in different contexts, drawing on findings and theories on clause types and questions (e.g. Ceong & Saxon, 2013; Garzonio, 2004; Shao, 1996; Tang, 1981, 1984) and on speech acts (Morgan, 1977).

Apart from marking sentence types, the expression (or lack thereof) of aspectual meanings by preverbal YOU-MEI-YOU in the clauses was also used to categorize the data. This type of categorization involved several steps. As an initial step, clauses were coded on the basis of the aspect morphemes in Standard Mandarin (e.g. GUO, LE) found in them, if any. That means if a clause contained a perfective GUO, it would be coded as Perfective and one that contained a progressive ZAI would be coded as Other, and so forth. During the second step, all the clauses, including the ones containing an aspect morpheme and the ones that did not, were examined. During this process, evidence was gathered via the technique of indirect proof (Smith, 1997, p. 63) to make further judgment.

This indirect proof technique involves the use of semantic tests that were similar to those used in Smith (1997). Such tests were designed to determine the aspectual viewpoint conveyed in a sentence/clause and different tests were applied to analyze the same sentence so that the categorization would be backed up by sufficient evidence.

One such test looked at positive answers to the question(s) containing YOU-MEI-YOU in natural speech, if answerable, and/or the actual answer in the original text, if available. For example, if a YOU-MEI-YOU+VP question contained both YOU-MEI-YOU and the morpheme GUO, and the positive response to it was composed of the predicate/verb and GUO attached to it, it would confirm that the sentence is in fact Perfective. In the following example, both the question and the response contain a perfective marker GUO and in the response, negation has scope over the VP:¹⁰²

(39) (Reporter) ni zuijin yi liang ge yue you-mei-you chu-guo
 2SG recent one two CLF month have-not-have exit-GUO
 yuanmen ne?

distant.door NE

Did you ever go to any faraway places in the recent one/two months?

(Luo) yi liang ge yue nei shi meiyou dao waimian zou-guo de.
 one two CLF month inside SHI NEG go outside walk-GUO DE
 (I) really did not travel outside in the past one/two months.

¹⁰² In this example, MEIYOU is positioned after SHI (是), which is used together with the sentence-final DE (的) to mark emphasis.

However, if GUO was found in the original question, it would raise the question about the specific role of YOU-MEI-YOU played in this kind of clause, or whether it functioned as the aspect viewpoint marker in a clause that contained a Perfective GUO or LE for that matter.¹⁰³ Therefore, clauses containing both preverbal YOU-MEI-YOU and the perfective morpheme GUO were coded as GUO Perfective and those containing preverbal YOU-MEI-YOU and the perfective morpheme LE were coded as LE Perfective. I will elaborate on this in Section 4.1.2.

The test that looked at the possible or actual answer(s) in the affirmative to questions containing preverbal YOU-MEI-YOU worked the same way for clauses without (other) aspectual morphemes like GUO or LE. For instance, if the positive answer involved the use of the perfective morpheme LE attached to the predicate/main verb of the original question that did not contain LE, the clause would be coded as Perfective. The following example illustrates this:

(40) Question: nali you-mei-you zai shu ya?
 there have-not-have plant tree YA¹⁰⁴

Were any trees planted over there?

Answer: xin zai-le yixie sangshu
 new plant-LE some mulberry.tree

Some mulberry trees were planted (there) recently.

¹⁰³ Clauses in which YOU-MEI-YOU did not seem to convey viewpoint aspect, and those in which it did not seem to be the (sole) marker of viewpoint aspect were excluded from this particular step in order to focus on what constituted cut-and-dried evidence for the development of an aspectual YOU-MEI-YOU.

¹⁰⁴ The morpheme YA (呀) is an allomorph of the sentence-final particle A (啊), which is used at the end of an interrogative or exclamatory clause.

Clauses like (40) that did not contain the perfective morpheme GUO or LE were coded as YOU-MEI-YOU Perfective, in order to differentiate them from the ones that contained the perfective GUO or LE.

In another test, the question was transformed syntactically in a way that best preserved the original meaning. This was done by asking the question in a different way. Recall that there are various ways to ask the same question in Standard Mandarin that do not have significant differences in meaning (see examples 3a-d, repeated as 41a-d below):

(41) a. ta lai-le ma?

3SG come-LE Q

Did he come?

b. ta lai-le meiyou?

3SG come-LE NEG

Did he come?

c. ta lai mei lai?

3SG come NEG come

Did he come?

d. ta you-mei-you lai?

3SG have-not-have come

Did he come?

As shown in the examples, to ask a Perfective question, one can use the A-not-A construction of the main verb/predicate (41c).¹⁰⁵ One could also attach the Perfective morpheme LE to the main predicate, in addition to using the question particle MA or

¹⁰⁵ The 'not' element is MEI in this case.

MEIYOU at sentence-final position (41a-b). These alternatives would be used to reword the clause in question: if the YOU-MEI-YOU+VP questions could be reworded by using some of these other formats, with their original meanings largely intact, they could be classified as Perfective.

The third test was similar to the second one in nature and involved the use of SHI BU SHI, the interrogative construction of the focus marker SHI in Standard Mandarin. This focus marker does not change the basic meaning of a question, but it does add a focal reading to the element that immediately follows it. Its interrogative form, SHI BU SHI, can be used to turn a simple declarative into a question and focus any constituent in the sentence, but it does not add any grammatical meaning to the element being focused or to the clause in which it appears. In other words, if a question was reworded by replacing YOU-MEI-YOU with SHI BU SHI, it would become apparent whether or not YOU-MEI-YOU conveyed some type of aspectual meaning in the original clause. If preverbal YOU-MEI-YOU did convey some type of aspectual meaning in the original clause, a new aspectual morpheme would be required for the transformed clause to be complete and undistorted in meaning. This can be illustrated by the sentences below:

(42) a. xiujian fangwu, zhengfu you-mei-you gei butie?

build house government have-not-have give subsidy

Did the government furnish any subsidy for building the houses?

b. xiujian fangwu, zhengfu shi bu shi gei-le butie?

build house government SHI BU SHI give-LE subsidy

Did the government furnish any subsidy for building houses?

The SHI BU SHI test was used to complement the other two tests, since it could be applied to subordinate clauses as well as matrix/main clauses.¹⁰⁶

It is worth mentioning that depending on the type of clause in which YOU-MEI-YOU appeared (matrix/main or subordinate), the third test could be applied to tease apart the function of YOU-MEI-YOU and that of any other aspectual morpheme that happened to be found in the same clause. It was found that, whenever YOU-MEI-YOU co-occurred with another aspectual marker, it was the latter that clearly conveyed the aspectual meaning of the clause, be it Perfective or not, and YOU-MEI-YOU seemed to be redundant in this respect, even though it still functioned as a question marker, as pointed out in Dong (2004, p. 3).

Along similar lines, in his analysis of YOU+VP and YOU-MEI-YOU+VP across four regional varieties of Standard Mandarin, Diao (2012) also pointed out the redundancy of YOU when used in the YOU+VP+GUO construction. It is not clear whether or to what extent YOU-MEI-YOU contributes to the aspectual meaning conveyed in this type of clause. Such instances probably could not be used as direct evidence to support the hypothesis, since only when the preverbal YOU-MEI-YOU conveyed the aspectual meaning, or marked the sentence type, of the clause, could it be established as an auxiliary, as claimed in this study. Nonetheless, these entries would still be included in order to present the broader picture of the emergence and development of a preverbal YOU-MEI-YOU.

The last step in coding clauses for aspectual meaning required that the aspectual meaning of the verb constellation, i.e. the main verb and its argument, in each clause be

¹⁰⁶ The second test is not always applicable in the case of subordinate clauses in which the VP is complex.

examined against the four ways in which an event could be bounded, as described in Li & Thompson (1989).¹⁰⁷ If the event denoted by the verb constellation could not be bounded, such clauses would not be coded as Perfective even if they contained preverbal YOU-MEI-YOU. This measure was taken to prevent incorrect labeling of clauses whose aspectual reading might have been changed by some “abnormal” or creative use of preverbal YOU-MEI-YOU.

3.4.2 Data Analysis to Identify the Developmental Stage of YOU-MEI-YOU+VP

The second research question was addressed by two types of investigations, one of which scrutinized the syntactic and semantic features of preverbal YOU-MEI-YOU in different historical periods, the other of which looked at the types of the VP complement the construction took. The first investigation was built on the parameters outlined in Heine (1993) and the second was inspired by cross-linguistic research on possessive perfects. Much overlap was found between these two lines of inquiry in terms of the linguistic features or criteria used. The following paragraphs touch on such overlapping features/criteria and challenges in applying them to the description of the development of preverbal YOU-MEI-YOU.

Following Heine (1993), the current study first looked at two parameters associated with the stages of development of auxiliaries: desemanticization and decategorialization.¹⁰⁸ The former is primarily concerned with the nature of the subject,

¹⁰⁷ See Section 2.3.2 in the previous chapter for more details.

¹⁰⁸ The other two parameters, cliticization and erosion, were found to be irrelevant or inapplicable in the case of this study: the former is concerned with morphology of a grammaticalized entity, but Standard Mandarin is a tenseless language, lacking morphological inflections; as for the parameter of

the concept expressed by the verb (lexical vs. grammatical), and what is expressed by the complement (object/location vs. dynamic situation). With regard to the latter, the focus is on the verbhood of the grammaticalized entity, the type of complement it takes (noun phrase/adverbial phrase vs. verb/verb-like phrase).

Similar parameters and diagnostic tests have been proposed in cross-linguistic studies on the possessive perfect (e.g. Coussé, 2013; Łęcki, 2010) measuring its stage of development, including the agentivity of the subject, agreement between the past-participle-verb and the Patient NP, the concreteness of the direct object, etc. However, most of these parameters or tests could not be applied to the data in this study. The reasons are as follows:

To begin with, any test that is contingent on the inflections of the verbs is simply not applicable in the case of Standard Mandarin, which has a severely impoverished morphological system. In addition, Chinese is a topic-prominent language in which the ordering of subject, verb and object can be very flexible, allowing for “double subjects” (a topic and a subject) and null expletives (as compared to “dummy subjects” in subject-prominent languages like English), making less distinction between the subject and the object and so on (Li & Thompson, 1976). All this means that it would be difficult to base any claim or conclusion on the identification of a certain type of subject in a clause, if the subject could be correctly identified in the first place.

Secondly, as mentioned in the previous chapter, it is difficult to test the verbhood of preverbal YOU-MEI-YOU due to its structure and meaning (but see the discussion in Section 2.4). This means that the loss of verbal properties—for example, its ability to

(phonological) erosion, it was not useful in this study that used data that were largely from written corpora.

form imperative, to passivize, to be negated (see Heine, 1993, p. 55), which was associated with the decategorialization parameter, would not be a useful indicator of developmental stages in the case of preverbal YOU-MEI-YOU.

As a direct consequence of the problems sketched above, I concerned myself only with the concept expressed by preverbal YOU-MEI-YOU (lexical vs. grammatical) in my investigation of the desemanticization parameter, and the type of complement it took (nominal/adverbial vs. verbal) in the examination of the decategorialization parameter. The data were grouped into 20-year periods and the analysis looked at two sets of features across time. The distribution of these clauses in different periods was compared to answer the second research question.

Since various filtering procedures have eliminated the majority of clauses in which YOU-MEI-YOU was followed by a clearly non-verbal complement, a finer-grained analysis of the type of complement following YOU-MEI-YOU in the data was conducted to supplement the above investigation.¹⁰⁹ The properties of the complement identified as characteristic of each stage of development of periphrastic constructions of perfect in various languages in previous studies would serve as the starting point and would be adapted in the current study for determining the developmental stage of preverbal YOU-MEI-YOU.

It was found in Heine & Kuteva (2006) and Łęcki (2010) that, as the verb of possession advances in grammaticalization, there are less restrictions on the kinds of verbs serving as the main verbs or the complement of the grammaticalized verb. For instance, in the case of the English *have*, at the initial stage of development, a lexical verb

¹⁰⁹ Note that some ambiguous cases, along with some that are difficult to label, remain after the two rounds of filtering.

meaning ‘have’ linked the NP in the object position (playing the thematic role of Possessee) with the NP at the subject position (playing the role of Possessor). In Stage II, there were two verbs: the possessive verb, and another verb, in the past participle form, which modified the Possessee NP. However, at the more advanced stages in which the verb meaning ‘have’ achieved auxiliariness, the other verb became the main verb of the sentence, taking a direct object NP. Later in the development, this main verb need not be transitive and could be intransitive. Therefore, the appearance of an intransitive main verb is taken as an indication of a more advanced stage of grammaticalization for languages with an auxiliary meaning ‘have’ developed from verbs of possession.

Similar to the European languages with possessive perfect, preverbal YOU-MEI-YOU in Standard Mandarin is also historically related to the possessive construction YOU MEI YOU+NP, which appeared much earlier in time, as argued in the literature mentioned previously (see Dong, 2004; Kuang, 2000; Ota, 1958/1987; Shi & Li, 2001; Xue, 2010). However, Chinese YOU MEI YOU can be used both as a verb of possession and an existential verb. It is not clear if YOU-MEI-YOU+VP has its origin in the possessive YOU MEI YOU or existential YOU MEI YOU. Hence the type of VP complement was used as a criterion for determining the developmental stage of preverbal YOU-MEI-YOU. The use of this criterion was inspired by cross-linguistic research on possessive perfect. However, this criterion was used in the analysis with the caveat that the findings might not truly reflect the development of YOU-MEI-YOU+VP since it could have originated from the non-possessive predecessor and so might have followed a different path of grammaticalization.

In summary, I looked at the types of VP complement in clauses produced or published in different 20-year periods to adduce (further) evidence for determining the stage of development of preverbal YOU-MEI-YOU. The analysis began with the elimination of clauses labeled as Imperfective & Other as well as the Perfective ones with (other) aspectual markers (e.g. GUO, LE). The remaining clauses were further categorized according to the types of VP complement YOU-MEI-YOU took.

In terms of the categorization of such VP complements, the current analysis drew on the findings and dominant theories about verb phrases and sentence structures of Standard Mandarin. Four major categories were identified and considered in the analysis of the data: simple transitive VPs, intransitive verbs (including the cases in which the meaning was ambiguous between an NP reading and an intransitive reading), complex VPs, and passive. Next, I explain the basis on which these categories were selected and how they were operationalized in this study.

Simple transitive VPs and intransitive verbs were selected based on the findings about periphrastic perfect in different languages in earlier research (e.g. Coussé, 2013; Łęcki, 2010). In this study, a simple transitive VP was defined as a simple verb taking one or more NP complements. An example of a simple transitive VP in Standard Mandarin is given below:

- (43) you-mei-you ganjue-dao yixie yali?
 have-not-have feel-RVC some pressure
 Did you feel any pressure?

The definition of simple transitive VPs was based on Li & Thompson (1989) and meant to include transitive verbs taking a topicalized direct object, and transitive verbs with their objects appearing before them.

The definition and identification of intransitive verbs was more complex and difficult because of the flexibility of the syntax and semantics of Standard Mandarin (e.g. null expletives, null subjects, null objects, topicalized objects vs. “double subjects”). The current analysis followed Wang (1985) and Zhang (2002) and defined an intransitive verb as a verb not capable of taking objects in normal contexts or a verb that did not take an object in a clause. Drawing on Li & Thompson (1989), adjectival verbs and copula verbs in simple copula sentences involving a referential subject NP and a non-referential NP (Li & Thompson, 1989, p. 149) were also considered intransitive. In addition, in the identification of intransitive verbs, studies on unaccusative verbs, including double unaccusative structure bearing surface resemblance to a simple transitive VP and unergative verbs, were also used as a frame of reference (e.g. Huang, 2007; Zhang, 2009).

Different from simple transitive VPs and intransitive verbs, complex VPs were included on the basis of previous research on YOU-MEI-YOU. In Diao’s (2012) investigation of corpora compiled for regional varieties of Mandarin Chinese, it was found that the structure of the VPs used in YOU-MEI-YOU+VP sentences was rather simple compared to the wide range of structurally complex VPs found in MEIYOU+VP sentences (Diao, 2012, p. 48). This, along with the observation that only a small number of sentences that contained YOU-MEI-YOU+VP were present in those regional data, led Diao to the conclusion that YOU-MEI-YOU+VP was still a relatively underdeveloped

form. In light of this, the category of complex VPs was proposed and was viewed as an indicator of the development of the target construction in this study.

In the operationalization of complex VPs, the pure length of the VPs, which was highlighted in Dong's (2004) discussion of the syntactic behavior of preverbal YOU-MEI-YOU, was not given much weight. The operationalization was instead based on the findings in Shi & Li (2001) and focused on the syntactic properties of the VPs.¹¹⁰ In the current analysis, complex VPs were operationalized in such a way that a verb taking a clausal object, a verb taking as its complement another VP (e.g. the BA construction,¹¹¹ which was analyzed as a type of serial verb construction in some studies, serial verb constructions and pivotal constructions); two or more VPs conjoined in some way would be included in this category.

The fourth category, passive constructions, refers to overtly marked passive sentences (see example 44), but does not include sentences without any syntactic passive marker (e.g. BEI) that have a passive (or unaccusative) reading (see example 45):

(44) ta bei (gongsi) chezhi le.¹¹²

3SG BEI (company) fire LE

He was fired (by his company).

¹¹⁰ Shi & Li (2001) pointed out that YOU-MEI-YOU+VP could not be reworded using other ways of question formation when the predicate following YOU-MEI-YOU was one of the following types: serial verb constructions, idiomatic expressions and a verb with a clausal complement.

¹¹¹ A typical example of BA construction will look like this: Subject+BA+NP+Verb [+XP]. This construction has been extensively discussed by many influential Chinese linguists, but "there has not been a clear consensus on how to best characterize the properties" (Huang, Li, & Li, 2009).

¹¹² Example adapted from Li & Thompson (1989, p. 494).

(45) jingzi shuai po le.¹¹³

mirror fall break LE

The mirror crashed.

Note that in (44), the semantic direct object NP which is affected by the action of the main verb appears in sentence-initial position and is followed by BEI.¹¹⁴ It was termed a “passive coverb” in Li & Thompson (1989, p. 492), which introduces the doer of the action.¹¹⁵

The passive constructions were placed in a separate group because they did not seem to fit in any of the other three categories: on one hand, the predicate of the sentence was a transitive verb, with its semantic object occupying the subject position in the sentence; on the other hand, the predicate could not take another NP as its syntactic object. In terms of syntactic complexity, this type of sentence, composed of a subject, a passive morpheme, and a main verb, could not be considered complex in structure.¹¹⁶ In addition, many of the passive morphemes (e.g. BEI) have been grammaticalized and have achieved some sort of auxiliariness in this kind of context, as argued in numerous studies (e.g. Li & Thompson, 1989; Wang, 1958). In view of this, such sentences were put in a separate group.

¹¹³ Example adapted from Deng (2004, p. 297).

¹¹⁴ There are other passive markers (e.g. RANG, GEI, ZAO, JIAO) and other possible forms in which the passive constructions can occur (see Li & Thompson, 1989; Wang, 1958, for a detailed discussion).

¹¹⁵ The introduction of the action doer by a passive coverb like BEI is syntactically optional.

¹¹⁶ The same argument could be made about the BA construction in Standard Mandarin, since BA was also analyzed as a coverb in Li & Thompson (1989, p. 357) and the BA construction was often analyzed side by side with the BEI construction in studies on Chinese syntax (e.g. Huang, Li, & Li, 2009; Li & Thompson, 1989). However, the BA construction differs from the BEI construction in one important manner. The predicate of the BA sentence can be followed another phrase of any type to form a serial verb construction (see Huang, Li, & Li, 2009, p. 163 for an overview). Hence, the passive constructions were not grouped together with the BA constructions.

To recapitulate, after analyzing the grammatical meaning(s) conveyed by preverbal YOU-MEI-YOU and the general types of complement it took in clauses found in different time periods, I further examined the data in terms of four types of VP complement that followed YOU-MEI-YOU. The following table presents these four categories with examples:

Table 6 Sub-categorization of VP complements

Type of VP Complement	Example Sentence
Simple Transitive VPs	you-mei-you jie-zhu liang jingjing have-not-have intercept-RVC Liang Jingjing Did (you) intercept Liang Jingjing?
Intransitive VPs	baoanqi you-mei-you shiling security.device have-not-have fail Did the security device fail?
Complex VPs	you-mei-you bang ren na dian dongxi have-not-have help person bring some thing Did (you) help others carry some stuff back?
Passive	liangshi you-mei-you bei hongshui jin food have-not-have BEI flood soak Was the food soaked in the flood?

3.4.3 Data Analysis to Present the Distribution of Preverbal YOU-MEI-YOU

The development of preverbal YOU-MEI-YOU could be investigated from a theory-internal perspective, by tracking certain linguistic features the construction exhibits across time, which are arguably connected with different developmental stages, as described above. Alternatively, one could study the historical development of the construction from a different point of view, and look at facts that are relatively independent of theories to glean more insights into its developmental trajectory. Frequency of use was found to provide such an empirical measure of the progression of grammaticalization. The present study looked at the distribution of preverbal YOU-MEI-YOU across time to answer the third research question. The analysis for this purpose would be conducted in the following manner:

After the clauses were coded according to the grammatical meaning preverbal YOU-MEI-YOU conveyed, the ones labeled as Perfective or Interrogative would be used in a cross-section comparison. First, the year of each such clause was determined by referring to the source information about the text in which the clause was found, if available in the databases, or by manually searching for the clause via Google, accompanied by follow-up verification and/or investigation to find out the year in which the source text was composed or published. The year of composition, whenever determinable, was used, and the year of publication was used in the absence of the former. Taking into consideration the possibility that clauses that contained preverbal YOU-MEI-YOU might not be found for some year(s), this study would make a broader comparison by looking at the clauses found in different 10-year periods (e.g. 1920-1929).

Next, the data were normalized to ensure comparability across time periods. Prior to the normalization/rescaling of data, the size of text for different time periods used in each corpus was calculated. Since no such clauses containing YOU-MEI-YOU+VP were found in the CNC-AC corpus, this corpus would simply be disregarded. Although one clause containing YOU-MEI-YOU+ VP was found in the CCL-AC corpus, it was treated as an outlier and was not considered in the normalization of data.

The texts for which the year of composition or publication could not be determined were excluded from the calculation. Some of the challenges encountered and the makeshift solutions adopted in this step were summarized below and a detailed discussion of the problems that remained was found in the next chapter.

One of the major problems with calculating the size of texts for different time periods was that the clauses containing preverbal YOU-MEI-YOU were found in two corpora, which are different in size and for which different kinds of information are available.¹¹⁷ According to the official introductory text, the CNC-CC corpus has over 100 million characters, but only about 19 million are searchable. Apart from that, judging from the list showing the quantity of characters used in each text, along with the year of composition and/or publication of these texts in the CNC-CC corpus, only the data from the period from 1919 to 1998 were available and searchable.¹¹⁸¹¹⁹ In addition, no texts are found for certain years in the list probably because they were part of the data that are not accessible to outside users. Differences were also found in the available information about the size of texts between the two corpora: for the CNC-CC corpus, the number of

¹¹⁷ Only the CCL-CC corpus and the CNC-CC corpus were considered in this type of analysis.

¹¹⁸ The names of the texts are not included in the list, and each text is assigned a number instead.

¹¹⁹ The list was obtained through my personal communication with the researcher who led the team in constructing this corpus.

characters for each text is listed.¹²⁰ In contrast, the CCL-CC corpus only provides introductory documents that specify the size of texts in bytes.

For the purpose of the current data analysis, it was decided that the two corpora should be combined and the sum of the texts from the same time period found in the two corpora would be used in the normalization of data. To eliminate the difference in unit of measurement between the corpora, the size of texts in the CCL-CC corpus, measured in bytes, would be divided by two. The number two was chosen because it is the ratio of bytes to characters in the common Chinese character encoding system used in Mainland China.¹²¹ This ratio was confirmed by a comparison between the size of a text (measured in bytes) and the actual character count of the text in several texts randomly selected from the CCL-CC corpus. Nevertheless, it should be kept in mind that discrepancies may exist between an estimate made using this method and the actual number of characters in a text. It should also be borne in mind that, since a large number of data in the CNC-CC corpus are not accessible, the findings about the distribution of preverbal YOU-MEI-YOU should be treated with caution.

The second problem worth noting here was that, in the case of the CCL-CC corpus, some clauses for which the year of composition or publication could be determined were taken from text documents for which such information could not be determined. This happened due to several reasons: in some cases, the name of an author was listed as a search result in the introductory document that provides source

¹²⁰ However, the total number of words as shown in the list is slightly greater than that shown in the official website.

¹²¹ According to Wikipedia (“Chinese character encoding”, 2016), Guobiao is one common encoding system used in Mainland China, and is usually displayed in simplified Chinese characters. The ratio was suggested in my personal communication with Professor Atsushi Fukada.

information about texts used. That means, the entry listed includes some or all the works by this author. However, it was difficult, if not impossible, to determine which works were actually included in this entry. In some other cases, “a collection of works by XXX”, XXX referring to one author or a group of writers, was listed as an entry, alongside the information about its size. Such entries were discarded from the calculation of sample size for a given time period because the specific work (and therefore composition/publication date) from which the data were taken could not be determined. The clauses from such entries were also excluded from the calculation conducted to normalize the data.

Aside from the above mentioned problems, some minor issues surfaced that required making a decision regarding how to determine the date of some entries found in the CCL-CC corpus. For instance, some books were listed whose writing spanned a couple of years or longer, but for the purpose of my data analysis, it would be necessary to pick a year. Some ad hoc decisions were made to deal with such problems: for example, if the number of years taken to write a book was an odd number, the middle year was chosen, and if it was an even number, the year later than the median was chosen.

It is worth mentioning that the entries in the CCL-CC corpus for which the year of composition or publication could not be determined account for less than 10% of the sum of all the text used in the CCL-CC corpus.¹²² This should be kept in mind when the results of data normalization are interpreted.

After the two corpora were combined, with the units of measurement made equal and the sum of texts in the same (10-year) period from these two corpora calculated, such

¹²² Such data account for about 9% of the total in the CCL-CC corpus.

sums would be used to normalize the data. The total number of clauses containing YOU-MEI-YOU+VP found in each time period was divided by the sum of the characters for the same period to get the frequency per million characters and the resulting number for each period would be shown in a time series plot, to present the frequency of use of preverbal YOU-MEI-YOU conveying a grammatical meaning. An upward trend in the time series plots would be interpreted as showing a real increase in the frequency of use of (auxiliary) preverbal YOU-MEI-YOU.

3.5 Chapter Summary

This chapter began with an introduction of the two databases, each with two corpora, from which the data were collected and of the procedures of data collection. This was followed by an explanation of how the data were processed prior to being analyzed to answer the research questions. I explained two rounds of filtering the data underwent, presented the reasons for such procedures and discuss the methods used in the filtering procedures and the merits and challenges of the methods. I also summarized the results of the filtering in tables.

Next, I detailed the various types of data analysis for answering different research questions. I broke down the discussion into three parts, dealing with how each research question is addressed respectively. I explained that the data were first coded according to the grammatical meaning (i.e. sentence type and aspect) preverbal YOU-MEI-YOU conveyed in the clauses. The theoretical constructs and methodological tools used in this process were also briefly discussed.

After introducing the process of data analysis for testing the hypothesis, I discussed the two types of investigation involved in addressing the second research questions. I gave the reasons for using certain parameters and linguistic features (i.e. grammatical meaning, types of complement) instead of other parameters and explained how the data would be compared in terms of these features. I also clarified why and how the complements of preverbal YOU-MEI-YOU were further divided into four groups.

In the last section of this chapter, I first made clear that the third research question was answered by looking at the distribution of preverbal YOU-MEI-YOU over time. Then I gave a detailed account of how the frequency of use of preverbal YOU-MEI-YOU over time was calculated, including the steps involved, the problems encountered in the process and the ad hoc solutions for the problems, and touched on the potential problems with those ad hoc measures.

In the next chapter, I will present the findings from the different kinds of analysis discussed above and how the findings can inform us of the status of preverbal YOU-MEI-YOU in the grammar of Standard Mandarin, and shed light on its grammaticalization process.

4. RESULTS AND DISCUSSION

In this chapter, I will discuss the results of the various types of data analysis mentioned earlier in relation to the three research questions. I will present findings from which evidence can be drawn to answer each of the three research questions. I will also discuss facts or observations that were unanticipated or fall beyond the consideration of this study and attempt to explain how these findings deepen our understanding of preverbal YOU-MEI-YOU in Standard Mandarin.

4.1 Status of Preverbal YOU-MEI-YOU in Standard Mandarin

In the current study, it is hypothesized that preverbal YOU-MEI-YOU is grammaticalizing into an auxiliary unit, which is defined as a (de)verbal¹²³ item taking a nonfinite (de)verbal complement that helps to form constructions that convey tense, aspect, modality and other “cross-linguistically recurrent meanings beyond person, number and case marking” (Krug, 2011, p. 158). The two types of grammatical meaning conveyed by YOU-MEI-YOU investigated in this study are sentence type and aspect. This claim is substantiated by the results of the analysis of this construction. The evidence is listed as follows:

¹²³ A deverbal item is an item that derives from a verb, according to Krug (2011).

First, preverbal YOU-MEI-YOU meets one of the descriptions of an auxiliary as a deverbal entity in Krug (2011). It is historically related to the A-not-A construction of the lexical verb YOU ('have', 'there is/exists') (Shi & Li, 2001; Dong, 2004), which appears before NPs and functions as a lexical/verbal construction. In addition, it behaves like auxiliaries in Chinese (see exx 25-27 in Section 2.4).

Second, YOU-MEI-YOU takes verbal phrases as objects, as the well-established auxiliaries in Standard Mandarin do (e.g. NENG, HUI) (Chao, 1968, p. 731). The question of finiteness, which is part of the description of the complement of auxiliaries in Krug (2011), is irrelevant in the context of Standard Mandarin, which does not make a distinction between finite and nonfinite verbs. It's worth mentioning that there are some instances in which YOU-MEI-YOU was found to precede phrases that were ambiguous between VPs/APs and NPs in reading, but this should not count as counter-evidence for the auxiliary status of YOU-MEI-YOU. In effect, this type of instance constitutes evidence for the grammaticalization YOU-MEI-YOU is undergoing. A detailed discussion of this will be carried out at the end of Section 4.2.2.

The third and the most important piece of evidence that could be adduced for the hypothesis comes from the findings of the data analysis that examined two grammatical functions (i.e. sentence type and aspect) this construction expressed in the data. The results, which will be presented later in this chapter, indicate that preverbal YOU-MEI-YOU helped to form grammatical constructions that conveyed aspect in the majority of the cases. In addition, it was found to help form constructions that marked sentence type in the majority of the instances in which it occurred, although the percentage for this function of YOU-MEI-YOU was slightly lower than that for the aspect marking function.

The following table shows the number of clauses in which YOU-MEI-YOU expressed some kind of grammatical meaning and the number of clauses in which YOU-MEI-YOU was not found to convey either of the grammatical meanings under investigation:

Table 7 Distribution of different uses of preverbal YOU-MEI-YOU

Aspect		Sentence Type		TOTAL
		Interrogative	Non Interrogative	
Perfective	YOU-MEI-YOU Only	463 (45.26%)	257 (25.12%)	720 (70.38%)
	YOU-MEI- YOU+GUO	205 (20.04%)	50 (4.89%)	255 (24.93%)
	YOU-MEI- YOU+LE	7 (0.68%)	5 (0.49%)	12 (1.17%)
Non Perfective (Other)		21 (2.05%)	15 (1.47%)	36 (3.52%)
TOTAL		696 (68.03%)	327 (31.97%)	1,023 (100%)

There were 696 clauses in which YOU-MEI-YOU helped to express sentence type, and 257 clauses coded as Non Interrogative in which YOU-MEI-YOU helped to convey aspect. In total, there were 953 (out of 1,023) clauses in which YOU-MEI-YOU helped to convey one or both of the hypothesized grammatical meanings, i.e. sentence

type or aspect.¹²⁴ This accounted for approximately 93.16% of the total usage of YOU-MEI-YOU+VP. Looking individually at the different type of grammatical meaning conveyed, there were 696 (out of 1,023) clauses in which YOU-MEI-YOU helped to mark the sentence type, accounting for about 68.04% of the total and, there were 720 YOU-MEI-YOU Perfective clauses, which took up the majority of the total (70.38%).

Next, I will discuss in more detail the results of the coding for sentence type and then discuss the results of the coding for viewpoint aspect.

4.1.1 The Function of YOU-MEI-YOU in the Marking of Sentence Type

With regard to sentence type, out of the 1,023 clauses coded, 696 were coded as Interrogative, or clauses marked as direct questions by YOU-MEI-YOU. The results are shown in the table below:

Table 8 Distribution of interrogative YOU-MEI-YOU

Sentence Type	# of Clauses	Percentage
Interrogative	696	68.04%
Non Interrogative	327	31.96%

Among the 696 clauses coded as Interrogative, 665 were canonical instances of direct questions. Among the 696 clauses coded as Interrogative, 665 were canonical instances of direct questions, such as the yes/no question given in (46).

¹²⁴ The 55 clauses coded as Perfective and Non-Interrogative (50 instances of YOU-MEI-YOU+GUO; 5 instances of YOU-MEI-YOU+LE) were not counted because it was not clear whether YOU-MEI-YOU helped to convey the viewpoint aspect in such clauses. A detailed discussion is in Section 4.1.2.

(46) you-mei-you jiaoshui?

have-not-have pay.taxes

Did (you) pay taxes?

In the canonical instances of direct questions, preverbal YOU-MEI-YOU helps to mark the sentence/clause type by encoding some semantic element that contributes to the force associated with a certain clause type.¹²⁵ This structure offers two options and expresses uncertainty (Shao, 1996, p. 105) (which is considered the key element of questions in Lyons, 1977, p. 754) and invites the addressee to resolve the uncertainty.

In contrast, in indirect questions, this meaning of preverbal YOU-MEI-YOU shifts or is overridden when the clause containing YOU-MEI-YOU+VP becomes a complement of another sentence. Depending on the main verb and the subject pronoun of the matrix clause, the embedded clause may state a fact, express some known but unsaid information, or express some uncertainty but does not request an answer to resolve it (Shao, 1996). In other words, the clause that is embedded may be turned into a relative clause. Therefore, when found in embedded clauses, preverbal YOU-MEI-YOU loses its role in marking the clause that immediately contains it as belonging to a certain type.

Another type of clause coded as Interrogative (25 out of 695) contained a direct question verb (Tang, 1981) or question verb (Shao, 1996). This accounted for 25 of the 695 Interrogatives, or 3.60%. The majority of such verbs were BUZHI(DAO)/BUXIAODE (22/25), most of which were in sentence-initial position. In a

¹²⁵ The force associated with a certain clause type is sentential force, according to Portner (2004), though in the syntactic literature, the concept of force often invoked is that of illocutionary force” (Portner, 2004, p. 1). I will offer the definitions of the two terms later. Drawing on Chierchia & McConnell-Ginet, (1990), Garzonio (2004) defined sentential force as “the conventional pragmatic force associated with a sentence type” (p. 3).

couple of these clauses, the direct question verb was found to immediately follow the subject, but the function remained the same. In addition, three clauses had three other such verbs/expressions: WENYIXIA (roughly translated as “I have a question”), NICAİ ‘you guess’, and DAJIA JUEDE (roughly translated as “in your opinion”). In such instances, the scope of the question only spanned the clause that immediately followed the direct question verbs,¹²⁶ which means that the addressee was expected to provide an answer in relation to this clause as opposed to the whole sentence beginning with such direct question verbs,¹²⁷ as in the following example, accompanied by the answer in the original text:

- (47) Question: bu zhidao qingqing zai jiaxiang, you-mei-you ding-guo
 NEG know Qingqing at hometown have-not-have engage-GUO
 qin?
 marriage
 Was Qingqing ever engaged in her hometown?
- Answer: mei... meiyou.
 NEG..NEG
 No... she was never engaged.

According to Shao (1996), this kind of question verb was used to introduce the actual question, and prompt the address to provide a response; most of such verbs/expressions could be seen as independent pragmatic elements and could be omitted without affecting the meaning of the actual question. Some of these verbs were analyzed

¹²⁶ If the “direct question verb” was found in the post-subject position, the scope of question applied to the clause without this verb.

¹²⁷ As one professor on my committee suggested, this type of question could also be seen as instances of indirect speech act.

as examples of interrogative marker by Chen & Chen (n.d.). Such markers are differentiated from expression of information (through lexical items, e.g. wh-morphemes in the case of interrogatives) and were believed to indicate what type of information should occur after it.

The remaining six Interrogative clauses behave differently from those discussed above in the sense that they were not pragmatically questions, though they did meet all of the criteria outlined in the previous chapter. This group was composed of independent clauses with a YOU-MEI-YOU+VP as the predicate, all ending with an exclamation mark. It was evident that they were not asking about anything but were instead expressing a certain attitude of the speaker, as shown in (48). The clause that preceded (48) in the original text also made it clear that the speaker was forced to play an unfavorable role and was not given a chance to choose. In another example, the clause seemed to be more a statement of what had happened, ending with a question mark as well as an exclamation mark, showing a kind of surprise (see 49):

(48) tamen you-mei-you gei wo yi ge xuanze de jihui!
 3PL have-not-have give 1SG one CLF choose DE¹²⁸ chance
 Was I really given any chance to choose?

(49) you-mei-you zheme kuai a?!
 have-not-have so quick A¹²⁹
 This/so quick?!

¹²⁸ The particle DE (的) marks XUANZE ‘choose’ as a relative clause, which modifies the noun it precedes—i.e. JIHUI ‘chance’. It is called Nominalizer DE in Li & Thompson (1989).

¹²⁹ The morpheme, A (啊), is a sentence-final particle; it expresses surprise or impatience when used at the end of an exclamatory clause; it can also be used in an interrogative.

Semantically, sentences like (48) and (49) are still interrogative sentences. However, from a pragmatic point of view, they were typical instances of indirect speech act. Morgan (1977) explained how an expression like “can you...”, which “has only the obvious literal meaning of a question about the hearer’s abilities”, “could have, via Grice’s maxims, the implicature of a request” (p. 23). This is how an indirect speech act is generated, according to Morgan. Similarly, if the person who heard (48) knew that the speaker did not have a choice, he or she could easily infer, via Grice’s maxims, that the sentence was intended to express an attitude (e.g. emphasis, condemnation). That could explain why this type of clause seemed to behave slightly differently from other interrogative clauses. Such clauses, however, were still coded as Interrogative.

Based on the findings presented above, it is reasonable to assume that preverbal YOU-MEI-YOU helps to mark sentence type in the majority of the clauses, although not in the absolute majority of the clauses. This suggests that the transition of YOU-MEI-YOU to an auxiliary unit may not be complete. Next, the results of data analysis showing the role of YOU-MEI-YOU in the expressions of viewpoint aspect are presented.

4.1.2 The Role of YOU-MEI-YOU in the Expression of Viewpoint Aspect

In this subsection, I present and discuss the results of coding for aspect. The following table encapsulates the results of coding clauses for viewpoint aspect.

I take the fact that YOU-MEI-YOU in 720 out of 1,023 clauses (70.38%) was the sole marker of perfectivity as evidence for my claim that YOU-MEI-YOU is grammaticalizing into an auxiliary unit. These 720 clauses all passed one or more tests

Table 9 Distribution of perfective YOU-MEI-YOU

Type of Viewpoint	# of Clauses	Percentage
YOU-MEI-YOU Perfective	720	70.38%
GUO Perfective	255	24.93%
LE Perfective	12	1.17%
Imperfective & Other	36	3.52%

for perfective aspect, as described in Chapter 3. For instance, for a typical example of YOU-MEI-YOU Perfective, the answer—actual or possible—would have the perfective morpheme, LE, attached to the (first morpheme of the) main verb, or be a simple MEI/MEIYOU.

However, caution was taken when the actual answer to the question being examined was not found in the original text and a possible answer was used to determine the aspectual meaning of the clause. If the answer was a simple MEI/MEIYOU or YOU, it was not considered as conclusive evidence for arguing that YOU-MEI-YOU in the original question conveyed Perfective viewpoint, because of the possibility that the answer did not correspond to the YOU-MEI-YOU+VP question in structure. To be more specific, the addressee used a YOU+NP or MEI(YOU)+NP structure in the answer, as shown in (50):

- (50) name fukena zuopin li you-mei-you xie-dao zhongguoren ne?
 then Faulkner works inside have-not-have write-RVC Chinese NE¹³⁰
 Did Faulkner ever write about Chinese people in his works?
 you de, danshi bu duo
 have DE¹³¹, but not many
 Yes, but not much.

Interestingly, the answer in the example above could also be taken as evidence for an auxiliary YOU since YOU in the answer could be seen as shorthand for YOU+VP (or XIE-DAO in this case). This might lead us to think that the addressee treated YOU-MEI-YOU as the interrogative form of an auxiliary YOU followed by a verb. A couple of other instances similar to this were found in the data which contained a YOU or YOU-DE, including one that was responding to a YOU-MEI-YOU+VP question containing the morpheme GUO. It is difficult to tell whether the addressee was using YOU as shorthand for YOU+VP or YOU+VP+GUO, or the sentence simply contained a lexical YOU and an NP. Thus, caution should be exercised when interpreting such responses.

Caution should also be taken when we look at an answer in the negative, which is much more common in the data. When an actual answer to a question could be found in the same text, most of the time, it consisted of a MEI/MEIYOU and maybe something else. If the answer was simply MEI/MEIYOU, without additional information, then we could not tell whether the answer had a different sentence or verb phrase+argument

¹³⁰ The morpheme, NE (呢), is a sentence final particle, often used in A-not-A questions.

¹³¹ The morpheme, DE (的), in this clause is used after the main verb YOU, for emphasis.

structure from the original question and whether it could be seen as evidence for a Perfective YOU-MEI-YOU.

The problem was further complicated by the observation that MEI/MEIYOU could be used to perform functions beyond marking perfectivity, as argued in Xiang (2014). That means, a verb constellation that is normally incompatible with Perfective LE can co-occur with MEI/MEIYOU. Therefore, if an answer to a YOU-MEI-YOU question was a simple MEI/MEIYOU, the addressee might be simply stating the (non)existentiality of the state or event encoded by the verb constellation. Since states are not compatible with Perfective viewpoint, the answer should not be treated as evidence for Perfective YOU-MEI-YOU.

The same problem was found with one type of paraphrasing test.¹³² When main or embedded clauses were transformed into an A-not-A questions, some stative verb constellations seemed to fit better into the V-MEI-V structure than the V-BU-V structure, if the original meaning was taken into account, although an answer in the affirmative would merely have the original verb in its bare form, without or not allowing for a Perfective LE attached to it. For such clauses, different tests were applied before a decision was made.

Apart from this, there were other challenges in using the paraphrasing technique, which were often associated with the structure of the VP, especially in the case of complex VPs. Sometimes, the challenge resulted from the semantics or pragmatics of the sentence. For instance, the A-not-A construction of the focus marker, SHI BU SHI, was likely to shift the meaning of the original sentence, which posed a challenge when

¹³² There are some other problems with this test, which I will briefly mention in the following paragraph.

deciding if and to what extent a clause reworded using this structure agreed with the original clause in terms of meaning.

Caution was also exercised when the transformation resulted in a LE after a verb phrase that was at the same time in the sentence-final position. As mentioned in the second chapter, the sentence-final particle LE is identical to a Perfective LE in form but conveys a different meaning: it signals the current relevance of some state of affairs to some particular situation (Li & Thompson, 1989, p. 240). A post-verb perfective LE and a sentence-final particle LE can co-exist in a sentence and it is hard to distinguish between these two forms in a sentence that is structurally ambiguous (e.g. Subject+VP+LE). Some instances were found in the data that had a LE in the post-verb position which was also the sentence-final position. In such cases, the results of other tests and technique were checked to find out if the LE in those clauses conveyed a perfective viewpoint.

Another point needs to be made about the results of coding for aspect: the relationship and interaction between YOU-MEI-YOU and GUO and LE. The data showed that YOU-MEI-YOU was compatible with LE and GUO in some clauses, although, the clauses that contained both YOU-MEI-YOU and LE did not constitute evidence for a Perfective YOU-MEI-YOU, which, in such cases, was redundant in terms of marking perfectivity. The relationship between YOU-MEI-YOU and GUO is not as straightforward. According to Smith (1997), GUO can be used as a (flexible) resultative

verb complement (abbreviated as RVC)¹³³ in Standard Mandarin, which indicates completion or termination, as in the following example:

(51) tamen xinshang-guo-le yi chang dianying.¹³⁴

3PL enjoy-RVC-LE one CLF movie

They enjoyed the movie.

Smith (1997) stated that “RVCs contribute to the perfective viewpoint and to the specification of situation type” and that they “may occur with the perfective morpheme” and may “also convey the perfective viewpoint alone” (p. 270). It is not clear to what extent such RVCs, when occurring in the same clause as a perfective morpheme, contributes to the perfective viewpoint or vice versa. Some clauses coded as GUO Perfective in the data seemed to contain a RVC GUO (as in 52) instead of an experiential GUO or the GUO presenting a prior closed situation whose final state no longer obtains (as in 53):

(52) fan you-mei-you chi-guo ne?

meal have-not-have eat-RVC NE

Did you already eat/Did you finish eating?

(53) you-mei-you dao-guo qianxian ne?

have-not-have go-GUO frontline NE

Have you ever been to the frontline?

The morpheme GUO was used as an RVC in (52) whereas (53) illustrates it experiential sense. For a couple of clauses, it was hard to tell, without enough contextual

¹³³ Two types of phase RVCs “affirming termination or completion” are discussed in Smith (1997): flexible and strictly completeive. Some examples of flexible RVCs listed include WAN, HAO, GUO.

¹³⁴ Example adapted from Smith (1997), Chapter 11, (19b).

clue, whether the question concerned the completion or termination of the verb constellation or the experiencing of the situation (“did it ever happen...”). The answer in the affirmative to a question containing RVC GUO would have a LE following GUO while the answer (in the affirmative or in the negative) to a GUO Perfective question would simply have the verb constellation followed by GUO. However, only a couple of such instances were found and the GUO in them was ambiguous between the perfective reading and the RVC reading. For a simpler analysis, they were both coded as GUO Perfective as the contribution of YOU-MEI-YOU to the perfective viewpoint in clauses like (52) remains unclear, while clauses containing other RVCs were still coded as YOU-MEI-YOU Perfective.¹³⁵

It is also worth mentioning that a single case was found where the clause that contained only YOU-MEI-YOU was given an answer composed of the verb with a GUO attached to it. The question is: was a GUO, which was phonetically null, in the original question? Another possibility is that, since the verb in the original question is a stative verb (XIANGNIANG ‘to miss’), which was normally only compatible with flexible RVCs like GUO, the addressee interpreted it as a GUO question, and responded with a GUO accordingly. This clause was coded as GUO Perfective in the data.

For these reasons, the results of such tests should always be treated with great caution since, despite the best effort, the semantic and/or pragmatic content of a clause would be twisted to some extent when it was rewritten using a different structure.¹³⁶

¹³⁵ Future research should consider grouping clauses containing RVCs and discussing them as a whole. The relationship or interaction between the perfective morpheme LE and YOU-MEI-YOU and RVCs is another topic worth exploring in the future.

¹³⁶ This should not invalidate the sentence-rewriting technique, because the change in meaning is minor and what’s really affected is the pragmatic properties rather than the propositional content. In the case of SHI

It should be pointed out that, even after these cases in which it is not clear whether YOU-MEI-YOU contributed to the aspect marking were excluded, along with those in which it clearly did not convey Perfective viewpoint, there were still a large number of clauses that could be considered evidence that YOU-MEI-YOU helped to convey a perfective viewpoint, and they accounted for about 70% of the total usage of preverbal YOU-MEI-YOU.

4.2 Developmental Stage of Auxiliary YOU-MEI-YOU: Evidence from Diachronic Comparisons

Having presented evidence for the auxiliary status of preverbal YOU-MEI-YOU, the next question to be considered is the developmental stage of preverbal YOU-MEI-YOU. The results of the second type of data analysis, i.e. diachronic comparison of the use of preverbal YOU-MEI-YOU in terms of different features, are presented in this section.

As mentioned in the previous chapter, for some years, the corpora did not have many or any clauses containing preverbal YOU-MEI-YOU. In order to consolidate the data, the comparison conducted to answer the second research question looked at how the same facets of preverbal YOU-MEI-YOU discussed above change across different 20-year periods. The earliest clause found in the corpora dates back to around 1808, and was the only one found in that period; the second earliest “eligible” clauses found in the data appeared over 100 years after this clause. Because of the large time gap, this one

BU SHI test, what is essentially changed is the degree of doubt. According to Shao (2002), SHI BU SHI+VP has, “semantically, obvious affirmative inclination” (p. 1)

example from the 19th century was excluded from the analysis pertaining to the second research question, leaving 1,022 tokens spanning from 1920 to approximately 2010.¹³⁷

4.2.1 Desemanticization: A Diachronic Comparison of Preverbal YOU-MEI-YOU

The investigation of the desemanticization parameter (Heine, 1993) focused on the concept conveyed by the construction in question. Heine (1993) proposed that grammaticalization would be accompanied by a transition in the concept, from lexical to grammatical. In this study, the grammatical concept was operationalized as sentence type and viewpoint aspect, as discussed above. Clauses in which YOU-MEI-YOU helped to mark aspect but not sentence type were coded as Perfective Only, and those in which it helped to mark sentence type but not aspect were coded as Interrogative Only.¹³⁸ When YOU-MEI-YOU expressed both meanings in the same clause, such clauses were coded as Both Perfective & Interrogative (abbreviated as Both P&I). Clauses in which YOU-MEI-YOU did not convey either grammatical meaning were coded as Neither Perfective Nor Interrogative (abbreviated as Neither P Nor I). The percentage of each of these four types was calculated based on the total number of clauses within the same time period and the results are shown in the tables below:

As shown in the table, preverbal YOU-MEI-YOU was used to express one or the other grammatical meaning in a consistent manner over time, with roughly comparable percentage of such uses in the different 20-year periods, since its earliest use that could

¹³⁷ The data used in this study were collected through a search in all the corpora performed at the end of 2013. It was noticed that the last update date was July 2009 shown in one of the official document listing the texts used in the CCL-CC corpus, although a few texts published after 2009 were also listed there. This raised the question as to how much data the CCL-CC corpus collected each year since 2009.

¹³⁸ Clauses that contained both preverbal YOU-MEI-YOU and another aspect morpheme (e.g. GUO, LE) were not coded as Perfective Only in the current analysis.

Table 10 Distribution of different uses of preverbal YOU-MEI-YOU in different time periods

Time Period	Perfective Only	Interrogative Only	Both P & I	Neither P Nor I	Total
1920-1939	13 (59.09%)	4 (18.18%)	4 (18.18%)	1 (4.55%)	22 (100%)
1940-1959	9 (50.00%)	3 (16.67%)	4 (22.22%)	2 (11.11%)	18 (100%)
1960-1979	15 (9.55%)	63 (40.13%)	76 (48.41%)	3 (1.91%)	157 (100%)
1980-1999	173 (27.33%)	112 (17.70%)	300 (47.39%)	48 (7.58%)	633 (100%)
2000- ¹³⁹	46 (23.96%)	51 (26.56%)	79 (41.15%)	16 (8.33%)	192 (100%)

be found in the databases. The percentage was slightly higher in some periods (e.g. the 1920-1939 period, the 1960-1979 period), but the differences seem small.¹⁴⁰ The same was roughly true of its use as a Perfective marker. The percentage of the use of Perfective

¹³⁹ The most recent clauses eligible for the data analysis were produced/published in 2006, although the data were collected in 2013. A rough estimate was made of the amount of data produced/published from 2007 to 2013, based on the texts for which the year could be determined. It was shown that the data collected in this period accounted for less than 2% of the total in the CCL-CC corpus (approximately from 1917-present). This indicates that the lack of YOU-MEI-YOU+VP might be partly attributed to the fact that not much data was collected after 2006 but other factor might be involved. I will come back to this problem when discussing the temporal distribution of preverbal YOU-MEI-YOU.

¹⁴⁰ No significance testing was conducting on the data.

YOU-MEI-YOU was comparable across different time periods, except for a dip (in the Perfective Only cases) in the 1960-1979 period.

What is worth pointing out is the rate of increase in the instances in which preverbal YOU-MEI-YOU helped express sentence type, with a spike in the 1960-1979 period, and an overall upward trend in the use of this construction as conveying both grammatical meanings in the same clause(s), despite the slight decrease in the last time period.

Upon closer look, the data in 1970-1979 were found to be responsible for the sharp increase in both cases: only about 66.67% of the 21 clauses found in the 1960-1969 period were coded as Interrogative, which was roughly equal to the figures in the last two 20-year periods. The percentage of preverbal YOU-MEI-YOU conveying both grammatical meanings in the same clause(s) was about 28.57% in the same period. In contrast, about 91.91 % of the 136 clauses found in the 1970-1979 period were coded as Interrogative and approximately 51.47% of the clauses with preverbal YOU-MEI-YOU were coded as Both P&I.

Since over 86% of the data for this 20-year period came from the 1970-1979 period, it should not be surprising that the distribution for this 20-year period was, in a sense, skewed. It was also noticed that the majority of the clauses found in the 1970-1979 period that contained preverbal YOU-MEI-YOU were taken from novels written by writers from Hong Kong and Taiwan.¹⁴¹ This may account for the abundance of different uses of preverbal YOU-MEI-YOU and a more detailed discussion will be carried out when the regional distribution of this construction is considered in Section 4.3.1.

¹⁴¹ Many clauses found in the 1980-1999 period were also taken from novels written by Hong Kong or Taiwan authors but they did not form the majority of the data for that particular time period.

According to the desemanticization parameter outlined in Heine (1993), who listed seven overall stages or focal points along the Verb-to-TAM (Tense, Aspect, Modality) chain, a lexical item undergoing grammaticalization when it acquires a grammatical function, which is a benchmark of Stage III. Relating the observations made of the grammatical use of preverbal YOU-MEI-YOU, it seems reasonable to argue that this construction had reached Stage III or beyond. The argument is strengthened by the fact that there was a noticeable increase in the use of Interrogative (or sentence-typing) YOU-MEI-YOU and a gradual rise in the instances of YOU-MEI-YOU performing a dual function over time, especially when the internal variation in distribution in the 1960-1979 period is taken into account.

4.2.2 Decategorialization: A Diachronic Comparison of the Complements of YOU-MEI-YOU

The second parameter for measuring the stage of development of preverbal YOU-MEI-YOU, adapted from Heine (1993), concerns the complement it takes. The data were coded as Verbal, Ambiguous and Other. This categorization was an adaptation of the nominal/adverbial versus verbal distinction in the decategorialization parameter proposed in Heine (1993). Since the data analyzed had gone through filtering procedures designed to eliminate instances in which YOU-MEI-YOU was followed by a complement that was clearly non-verbal, the ambiguous cases were used instead as a contrastive category.

The purpose of this kind of data analysis was to see if there was any pattern regarding the types of verbal complements that appeared after YOU-MEI-YOU and since not many clauses were found in the filtering that contained complements that were

ambiguous between NP and VP, the analysis focused more on the variety of verbal complement types in the data. The Other type was proposed because one clause was found in which YOU-MEI-YOU preceded a Modal+VP.¹⁴²¹⁴³ This clause was coded as Other since nothing better could be done within the scope of this study and it was felt necessary to do justice to its uniqueness.

The clauses coded as verbal were further divided into four groups: VP (verb phrases without a Perfective morpheme attached to it), VP+GUO (verb phrases with a Perfective GUO attached to it), VP+LE (verb phrases with a Perfective LE attached to it), and AP (adjective phrases). The reason for the sub-categorization was related to the functions of YOU-MEI-YOU in the clauses containing these types of complement: when YOU-MEI-YOU was found to co-occur with a verb phrase that was normally incompatible with it in a clause, it was often the case that this verb phrase was followed by GUO, and the aspectual reading of the clause was primarily expressed in the VP+GUO combination.

Similarly, YOU-MEI-YOU did not seem to play a role in conveying the viewpoint of the clause in the case of the VP+LE combination. As for the clauses in which an adjective phrase was found to follow YOU-MEI-YOU, the aspectual reading of the clause was different from the ones containing a VP complement: such clauses were usually coded as Imperfective & Other and YOU-MEI-YOU was not found to contribute

¹⁴² The modal found in the clause was KEYI meaning ability. This use was analyzed as a control verb in Lin & Tang (1995) and as a modal auxiliary in Huang (2009). It is normally not compatible with Perfective morphemes like LE and GUO in Standard Mandarin. The hierarchical relationship between KEYI and Perfective morphemes in a syntactic representation is beyond the scope of this study.

¹⁴³ The clause in which YOU-MEI-YOU preceded a Modal+VP is listed below:

(d) jintian, wo you-mei-you keyi gaowei ni?
 Today 1SG have-not-have can comfort 2SG
 Was I able to comfort you today?

to the aspect marking in them. Consequently, the clauses were put into sub-groups to better reflect the development of an auxiliary YOU-MEI-YOU.¹⁴⁴

The percentage of clauses belonging to the VP, VP+GUO, and VP+LE types in each of the five periods were also counted and listed in a separate column to give a more general idea of the presence of verb phrases in the post-YOU-MEI-YOU position. The results are shown in the Table 11.

At first glance, there seems to be no clear patterns in the distribution of complements of YOU-MEI-YOU in different time periods. Over time, the percentage of the clauses of the VP type, and that of the clauses containing a verb phrase complement seems to be comparable across (time) sections, fluctuating only to a certain degree. No recognizable pattern was found in the distribution of clauses labeled as the VP+GUO type or that of the VP+LE clauses, except for the observation that the latter appeared later in time and was quite rare in the data.

When the distribution of the different types of complement was looked at through a holistic lens, one interesting observation was made: the range of complement type YOU-MEI-YOU took seemed to grow wider, roughly speaking, without considering the Ambiguous type. During the first time period, there were only instances belonging to the VP and VP+GUO types. The second period saw the appearance of AP complements, and instances of the VP+LE type were found in the third time period. The peculiar clause coded as Other appeared in the fourth time period, though no such clauses were found in

¹⁴⁴ Since YOU-MEI-YOU in these sub-types of clauses may or may not help mark the sentence type, it is important to present clear cases of Perfective YOU-MEI-YOU to show the development of this construction. Presenting the sentence type coding of these clauses may help make the argument stronger but there is not enough data to break it down to even further subcategories. This may be undertaken in future research.

the most recent period. This pattern is interesting and if more data could be found showing a similar pattern, it could lend support to the argument about the development stage of preverbal YOU-MEI-YOU over time.

Table 11 Distribution of complements of preverbal YOU-MEI-YOU in different time periods

	VP Complements				Total VPs	Non-VP Complements		Total # of Clauses
	VP	VP+GUO	VP+LE	AP		Ambiguous	Other	
1920-1939	16 (72.73%)	4 (18.18%)	0 (NA)	0 (NA)	20 (90.91%)	2 (9.09%)	0 (NA)	22 (100%)
1940-1959	11 (61.11%)	4 (22.22%)	0 (NA)	1 (5.56%)	16 (88.89%)	2 (11.11%)	0 (NA)	18 (100%)
1960-1979	90 (57.33%)	62 (39.49%)	2 (1.27%)	1 (0.64%)	155 (98.73%)	2 (1.27%)	0 (NA)	157 (100%)
1980-1999	482 (76.15%)	129 (20.38%)	8 (1.26%)	5 (0.79%)	624 (98.58%)	8 (1.26%)	1 (0.16%)	633 (100%)
2000- ¹⁴⁵	132 (68.75%)	54 (28.13%)	2 (1.04%)	3 (1.56%)	191 (99.48%)	1 (0.52%)	0 (NA)	192 (100%)

When the distribution of the different types of complement was looked at through a holistic lens, one interesting observation was made: the range of complement type

¹⁴⁵ The most recent clauses eligible for the data analysis were produced/published in 2006, although the data were collected in 2013. A rough estimate was made of the amount of data produced/published from 2007 to 2013, based on the texts for which the year could be determined. It was shown that the data collected in this period accounted for less than 2% of the total in the CCL-CC corpus (approximately from 1917-Present). This indicates that the lack of YOU-MEI-YOU+VP might be partly attributed to the fact that not much data was collected after 2006 but other factor might be involved. I will come back to this problem when discussing the temporal distribution of preverbal YOU-MEI-YOU.

YOU-MEI-YOU took seemed to grow wider, roughly speaking, without considering the Ambiguous type. During the first time period, there were only instances belonging to the VP and VP+GUO types. The second period saw the appearance of AP complements, and instances of the VP+LE type were found in the third time period. The peculiar clause coded as Other appeared in the fourth time period, though no such clauses were found in the most recent period. This pattern is interesting and if more data could be found showing a similar pattern, it could lend support to the argument about the development stage of preverbal YOU-MEI-YOU over time.

Another point I want to make about the distribution of complements following YOU-MEI-YOU is that ambiguity seemed to persist over time in the development of YOU-MEI-YOU. Overall, there seemed to be fewer ambiguous cases as time passed, in terms of the percentage of such cases in each of the different time periods, but this could well be the result of the way the sections or time periods were selected. It is tempting to treat such ambiguous cases as equivalent to the nominal or adverbial complements that appear after a grammaticalized entity, and see its decrease and eventual disappearance or transformation into a verb as an indicator of a more advanced or later stage of grammaticalization, as suggested in Heine (1993).

Putting the above observation into a developmental perspective, I would argue that the discovery of ambiguous cases and of its persistence cast new light on the development of an auxiliary YOU-MEI-YOU. These instances could be seen as “critical contexts” (Diewald, 2006) or “bridging contexts” in Heine's terms (Heine, 2002, p. 84). These terms refer to situations in which multiple readings of the construction in question are possible, including the new grammatical meaning. Seen from this perspective, the

auxiliary YOU-MEI-YOU seemed to be at an early stage of development, since there still exist clauses in which the old and new meanings are both possible. However, given that YOU-MEI-YOU in historically earlier contexts normally took NP complements, and abundant instances of “isolating contexts” in which YOU-MEI-YOU could be unambiguously labeled as a grammaticalized construction have been found in the data,¹⁴⁶ there is good reason to believe that these entries merely reflect the persistence of ambiguous/critical contexts. In fact, it is not uncommon to find such persistence of critical contexts in the process of grammaticalization, a view entertained by Traugott (2012). This is particularly true in Standard Mandarin, a language in which the syntactic categories of words are hard to pin down.

I would also argue that that such ambiguous cases could potentially drive the grammaticalization of YOU-MEI-YOU, since the listeners might interpret the complement one way or the other and carry this over to (re)analyze similar cases. For instance, the example sentences (31a) (repeated here as 54a) could be read as a question asking whether something, referred to by the complement of YOU-MEI-YOU, existed or was found in a place or in a situation (as in 54b):

- (54) a. fushipin gongying qingkuang you-mei-you gaishan?
 nonstaple.food supply situation have-not-have improve
 Has the supply of non-staple foods improved?

¹⁴⁶ This refers to instances in which YOU-MEI-YOU takes VPs as complements and expresses some grammatical meaning, as those presented in first section of this chapter.

- b. fushipin gongying qingkuang you mei- you gaishan?
 nonstaple.food supply situation have NEG have improvement
 Is there any improvement in the supply of non-staple foods?

This interpretation is completely acceptable and natural, given the commonplace structure of NP (referring to a place or a situation)+YOU MEI YOU+NP in Standard Mandarin and the dual categorial status of the phrase GAISHAN, which could be used as an NP or a VP in other contexts. It is impossible to find out what was intended by the speaker, a character in a novel, or the author or whether s/he was actually using GAISHAN as an NP when producing this sentence.¹⁴⁷ An imaginary addressee could interpret GAISHAN as an NP and give an answer like (55) or (56); Or the addressee could read it as a VP and respond in a different way (as in 57):

- (55) cong qunian yilai you-le hen da de gaishan.
 from last.year since have-LE very big DE¹⁴⁸ improvement
 There has been a lot of improvement since last year.
- (56) meiyou shenme gaishan.
 NEG-have any improvement
 There is not much improvement.
- (57) gaishan-le (hen duo).
 improve-LE (very many)
 It has improved (a lot).

¹⁴⁷ The original text in which this clause was found was not available in the corpus, but one could easily think of similar sentences that have GAISHAN, or a synonymous intransitive verb following YOU-MEI-YOU, with an ambiguous reading

¹⁴⁸ The morpheme DE (的) in this sentence connects the adjective HENDA ‘very big’ with the noun it modifies, GAISHAN ‘improvement’.

Based on the findings from the investigation of the types of complement YOU-MEI-YOU took in the data, it should be reasonable to speculate that YOU-MEI-YOU has reached Stage IV outlined in Heine (1993), in which the complement loses nominal properties (p. 55). It might have even reached Stage V, in which the complement turns into a verbal phrase. In the following paragraphs, the results of a finer-grained analysis of the complement types YOU-MEI-YOU took in the data will be discussed in relation to this speculation.

4.2.3 Types of VP Complement of Auxiliary YOU-MEI-YOU in Different Time

Periods

The second type of analysis of the VP complement involved a total of 719 clauses, with the ones coded as GUO Perfective, LE Perfective, and Imperfective & Other excluded from the analysis.¹⁴⁹ These clauses were put into four groups, namely, Simple transitive, Intransitive, Complex and Passive. A detailed discussion about the rationale for the subcategorization and what are subsumed under each subgroup can be found in the previous chapter. The results of sub-categorization are shown in the table below:

The results did not produce anything insightful. The period of 1960-1979 saw an increase in the use of simple transitive VPs as the complement of YOU-MEI-YOU from the period of 1920-1939, but the percentage has remained roughly comparable since then. The use of intransitive VPs appeared at the very beginning and was found across the time periods. As for the use of complex VP complements, some instances were found from the very beginning when YOU-MEI-YOU started to take VP complements and with the

¹⁴⁹ The ancient clause, produced in around 1808, was excluded and the clauses coded as Ambiguous were included in the analysis.

exception of the 1940-1959 period, this type of complement has been in use over time. Passive VP complements were not found at the beginning of the use of preverbal YOU-MEI-YOU, had some presence in the 1940-1959 period, but did not appear again until the 1990s. The overall percentage of Passive complements remained relatively low.

Table 12 Distribution of VP complements of YOU-MEI-YOU in different time periods

Period	Simple Transitive	Intransitive	Complex	Passive	Total # of Clauses
1920-1939	3 (17.65%)	10 (58.82%)	4 (23.53%)	0 (NA)	17 (100%)
1940-1959	6 (46.15%)	5 (38.46%)	0 (NA)	2 (15.38%)	13 (100%)
1960-1979	43 (47.25%)	15 (16.48%)	33 (36.26%)	0 (NA)	91 (100%)
1980-1999	189 (39.96%)	146 (30.87%)	130 (27.48%)	8 (1.69%)	473 (100%)
2000	53 (42.40%)	30 (24.00 %)	41 (32.80%)	1 (0.80%)	125 (100%)

The results show that YOU-MEI-YOU took different types of VP complements from the beginning of its documented use and achieved a more balanced presence in various contexts (with different complements), so to speak, along the path of its

development. The pattern observed of the possessive perfect constructions in various European languages regarding the type of VP complement found in these constructions (Heine & Kuteva, 2006; Łęcki, 2010) did not seem to apply in the case of preverbal YOU-MEI-YOU, which took intransitive VPs as its complement in its earliest as well as latest use.

The results also contradict the finding in Diao's (2012) corpus-based study that the VPs found in the YOU-MEI-YOU+VP construction were rather simple. In the current study, the Complex category was operationalized in such a way as to include a wide variety of structurally complex VPs,¹⁵⁰ such as multiple verbs conjoined in a clause, serial verb constructions, pivotal constructions, Chinese idioms that were inseparable, transitive VPs taking clausal complements, etc. Such VP complements were not common (in terms of absolute count, not in terms of percentage in each time period) in the early use of preverbal YOU-MEI-YOU, but they eventually found their niche in the YOU-MEI-YOU construction later and have had a consistent and visible presence since then.

It is reasonable to conclude that the detailed analysis of the VP complements YOU-MEI-YOU took in different time periods did not produce evidence that could help us (better) determine the developmental stage that preverbal YOU-MEI-YOU has achieved. The theoretical and methodological constructs proposed on the basis of cross-linguistic studies on what seemed to be a functional equivalent to preverbal YOU-MEI-YOU did not prove to be very useful in investigating the construction in question. This suggests that a different kind of diagnostic tool and perhaps a different approach are needed to obtain a good grasp of the developmental stage of preverbal YOU-MEI-YOU.

¹⁵⁰ Many VPs were found to have lengthy but were not complex in terms of structure: for example, some clause contained a transitive verb and an NP object modified by a lengthy relative clause.

Based on the findings from the different diachronic comparisons discussed above, I postulate that YOU-MEI-YOU has reached Stage III (out of the seven stages of auxiliaries outlined in Heine (1993)). During Stage I, it was the main verb in the A-not-A construction taking NPs as complements. It was used to express a lexical concept. During Stage II, the complement expressed a dynamic situation, (cf. Section 1.4, Heine, 1993), as exemplified by GAIBIAN in (34). GAIBIAN in (34) can also be seen as a nominalized verb functioning as the complement of YOU-MEI-YOU. Examples such as this exemplify the second stage of the decategorialization chain in Heine (1993). In such sentences, however, the verbhood of YOU-MEI-YOU becomes difficult to test because two interpretations are possible (exx 6, 31, 34).¹⁵¹ During Stage III, YOU-MEI-YOU takes VPs as complements and expresses a grammatical function—i.e. sentence type or aspect—in the majority of the data. Furthermore, it exhibits features like auxiliaries in Chinese (ex. 25).

In the following section, I will present the results of another kind of diachronic comparison, that of the frequency of use of preverbal YOU-MEI-YOU, in different 10-year periods, as well as the results of comparison of different uses of the construction in different regions in different 20-year periods.

4.3 Temporal Distribution of Auxiliary YOU-MEI-YOU

A total of 912 (out of 1,022) clauses were used in the calculation of the percentage representing the frequency of use of auxiliary preverbal YOU-YOU-

¹⁵¹ See Dong (2004) for an alternative hypothesis about the intermediate stage.

YOU.¹⁵²¹⁵³ The number of clauses found in each 10-year period was used as the numerator and the number of characters in the two corpora in the corresponding period combined was used as the denominator. The following table shows the raw numbers of the base size and of the clauses in each time period:

Table 13 Base size and number of clauses in different time periods

Period	# of Clauses	# of Characters (CCL)	# of Characters (CNC)	Total # Characters	Tokens per Million Characters
1920-1929	3	230,136	40,976	271,112	11.07
1930-1939	13	548,817	1,171,734	1,720,551	7.56
1940-1949	5	1,613,722	623,226	2,236,948	2.24
1950-1959	8	5,846,452	1,406,291	7,252,743	1.10
1960-1969	20	5,419,514	1,185,680	6,605,194	3.03
1970-1979	134	7,414,179	1,246,769	8,660,948	15.47
1980-1989	148	29,842,592	10,952,883	40,795,475	3.63
1990-1999	408	223,872,812	4,881,547	228,754,359	1.78
2000-	173	216,324,174	0	216,324,174	0.80

¹⁵² A total of 1023 clauses contained YOU-MEI-YOU. One clause was excluded because it was produced in around 1808.

¹⁵³ A total of 952 clauses produced in the 20th century were found to contain an auxiliary YOU-MEI-YOU, but 40 were found in text documents for which the year of composition or publication could not be determined, although the information for the clauses could be determined (see Section 3.4.3 for a detailed explanation). These 40 clauses were all from the CCL-CC corpus and were not included in the calculation.

The results are shown in the following figure, with the time periods on the x-axis and the count per ten million characters on the y-axis:

In Figure 1, a downward trend was observed from the first time period (1920-1929) to and the fourth period (1950-1959). This was followed by an upward trend in the next two time periods (1960-1969 & 1970-1979). After that, a downward trend resumed. Two things are noticeable in the figure: a plunge on the leftmost side and a spike in the middle. First, the frequency of use as reflected in the percentage seemed high in the first two time periods, probably due to the small number of clauses found, and plummeted

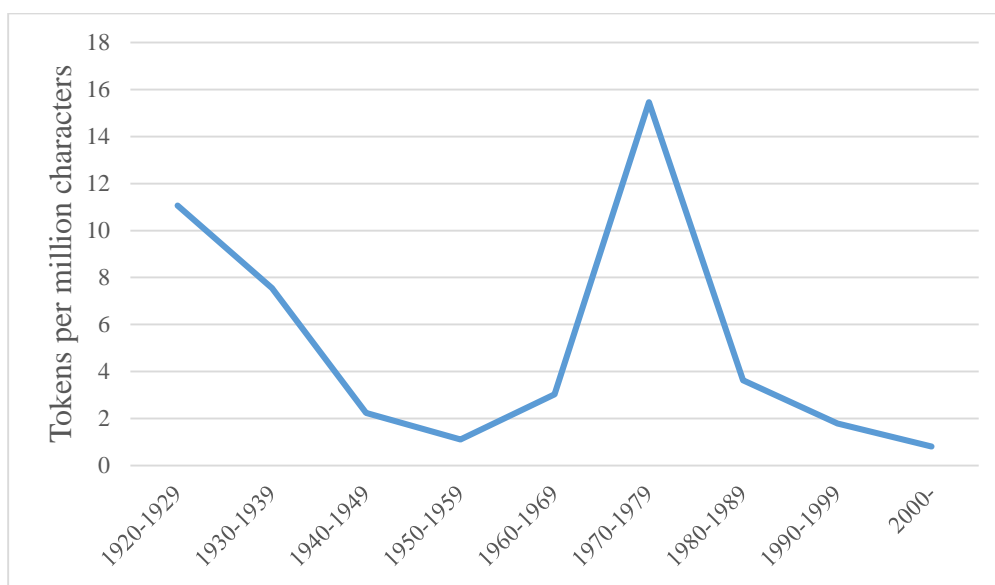


Figure 1 Relative frequency of use of auxiliary YOU-MEI-YOU over time

afterwards in the following two time periods. Second, the frequency dropped from the period of 1970-1979) to the next time period, as sharply as it rose from the 1960-1960 period to the 1970-1979 period.

In order to find out if the distribution was skewed by the differences between the two corpora that were combined, the data from the CCL-CC corpus, which accounted for the majority of the clauses containing auxiliary YOU-MEI-YOU (869 out of 912) were looked at separately and presented in a different figure.¹⁵⁴ Again, the data-points representing the first two time periods in the CCL-CC corpus were removed, as the base or the denominator for the percentage calculations for them was very small compared with that of other periods.¹⁵⁵ The following figure shows the results:

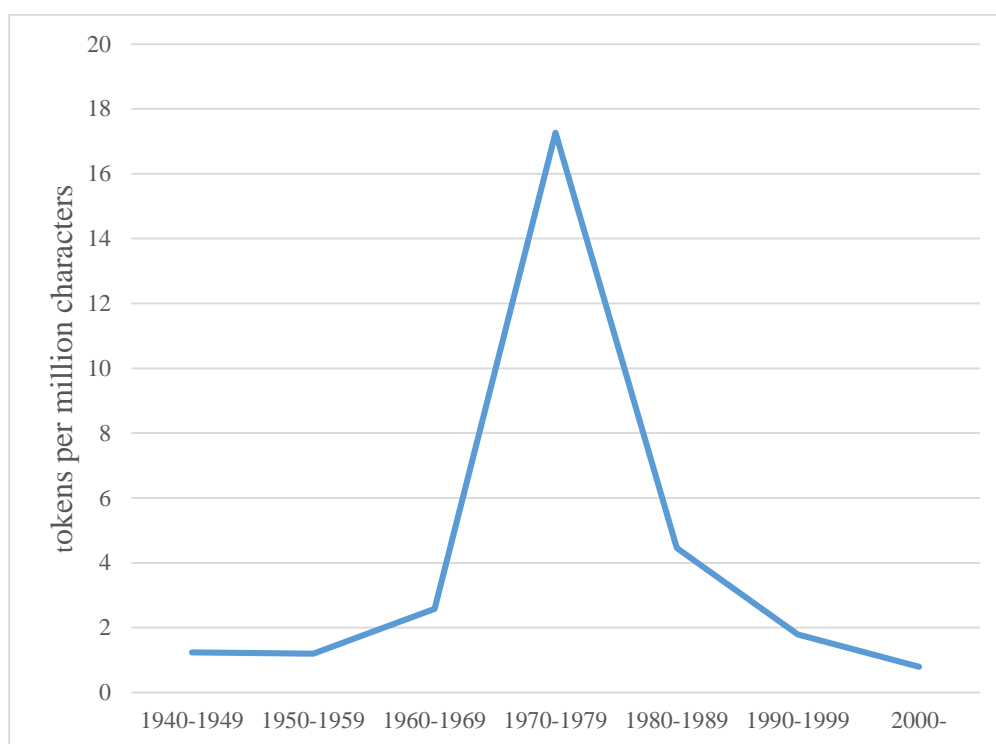


Figure 2 Frequency of use of auxiliary YOU-MEI-YOU in the CCL-CC corpus (all regions)

¹⁵⁴ For example, no texts produced or published after 1998 were found or available for search in the CNC-CNC-CC corpus. In addition, there did not seem to be much overlap in the source texts used between the CNC-CC corpus and the CCL-CC corpus, judging from the results of the search for YOU-MEI-YOU+VP clauses.

¹⁵⁵ The base number for the first time period in the CCL-CC corpus was less than half of that for the second time period, which had the second smallest base. The base number for the second time period in turn was less than a third of the base number for the third time period, which had the third smallest base.

If the data-points for the 1920-1929 and 1930-1939 in Figure 1 are ignored, Figure 1 and Figure 2 bear much resemblance to each other: an overall low percentage for different time periods, and a spike in the 1970-1979 period. In summary, no evidence was found that showed an increase in the frequency of use of preverbal YOU-MEI-YOU conveying some grammatical meaning over time. Since the exclusion of clauses taken from texts for which the year of composition or publication from the calculation produced a similar distribution for data in the CCL-CC corpus, the results of such modifications are not shown here.

Two questions remain about the frequency of use of this construction: (i) What does the sudden and sharp increase in the frequency (represented by the percentage) in the 1970-1979 period tell us about the use of auxiliary YOU-MEI-YOU? (ii) Is this construction on its way out, since a lower percentage was observed of the latest time period?

4.3.1 Distribution of Auxiliary YOU-MEI-YOU in Mainland China Chinese

In order to address the first question, I looked at the source information of the clauses found in the 1970-1979 period and of those found in other periods. As mentioned earlier, the majority of the clauses containing auxiliary YOU-MEI-YOU found in this period were taken from novels written by Hong Kong and Taiwan writers (121 out of 128).

When such clauses were excluded from the calculation, the percentage of auxiliary YOU-MEI-YOU for the 1970-1979 period dropped to about 0.94 tokens per

million characters for the 1970-1979 period. In other words, the peak in this time period disappeared when the clauses produced by speakers from Hong Kong or Taiwan were not taken into account. For the following time period, i.e. the 1980-1989 period, a sizable, though not a majority, of the clauses were also from such source texts and if such clauses were taken off the numerator, the percentage goes down to about 1.68 tokens per million characters.

This new percentage rate for the 1980-1989 period was higher than the “modified”¹⁵⁶ rate for the 1970-1979 period and was also higher than the “modified” rate of the 1960-1969 period, in which clauses containing auxiliary YOU-MEI-YOU produced by writers from Hong Kong or Taiwan were first found.¹⁵⁷ In summary, a new peak emerged in the 1980-1989 period when only clauses produced by Mainland Standard Mandarin speakers were considered.

Interestingly, if we did the same modification, i.e. excluding clauses produced by speakers or writers from Hong Kong or Taiwan from the calculation of the percentage, the last two time periods would not change to a great extent. Figure 3 presents the distribution of the data for Mainland China in the CCL-CC corpus:¹⁵⁸

To sum up, there were a large number of clauses containing auxiliary YOU-MEI-YOU produced by Hong Kong or Taiwan speakers in the 1970-1979 period, creating a

¹⁵⁶ The “modified” rate was based on the calculation that excluded the clauses containing auxiliary YOU-MEI-YOU produced by writers from Hong Kong or Taiwan.

¹⁵⁷ If this type of “regional” clause was taken off the calculation, the number for the 1960-1969 period was about 1.11 tokens per million characters instead of 2.58 tokens per million characters and the number for the 1980-1989 period would change from 4.46 to 1.68 tokens per million characters for the CCL-CC corpus.

¹⁵⁸ It should be noted that the source texts from the Hong Kong and Taiwan areas were not excluded from the calculation of the base number (used as the denominator in the calculation of the percentage for each time period due to the limited time. A rough estimate showed that the texts produced by speakers or writers from Hong Kong and Taiwan account for only around 1% of the total.

peak in Figure 2. However, when such clauses were removed, the 1980-1989 period saw the largest percentage of use of auxiliary YOU-MEI-YOU in Standard Mandarin used in Mainland China

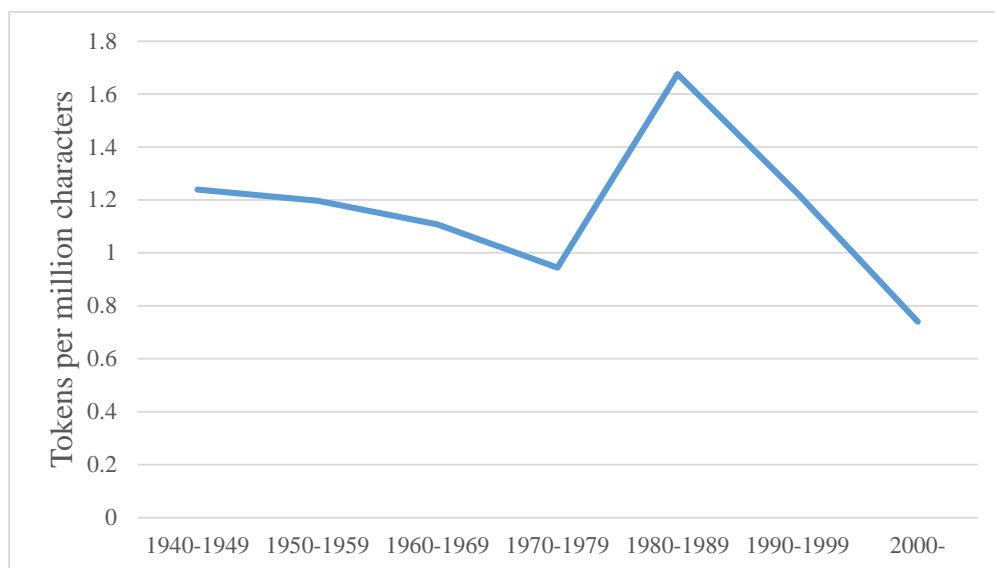


Figure 3 Frequency of use of auxiliary YOU-MEI-YOU in the CCL-CC corpus (Mainland China)

Previous research has speculated on the influence of Southern varieties of Standard Mandarin and dialects on the origin and development of YOU-MEI-YOU+VP in Mainland Standard Mandarin. This figure seemed to produce evidence supporting this speculation, though indirectly. It shows that the significant increase in the frequency of use of auxiliary YOU-MEI-YOU occurred in the 1980-1989 period, and this happened after the influx of a number of writings by famous or important writers from Hong Kong and Taiwan into Mainland China. Most of these books or novels were written or published in Hong Kong or Taiwan mostly in the 1960s and 1970s, some in the 1980s, but were not introduced to Mainland Chinese speakers until after the Reform and Opening-up Policy, proposed and implemented at the end of the 1970s, was in effect.

In fact, Hong Kong and Taiwan pop cultures introduced to Mainland speakers in the form of novels, TV shows and movies in the early 1980s started to gain momentum in the 1980s and 1990s.¹⁵⁹ This could well explain why the high rate of frequency of use of auxiliary YOU-MEI-YOU in Mainland China in the 1980-1989 period. As for the percentage of the 1990-1999 period, it is much lower than that of the previous period but still higher than that of other periods, probably because the once strong influence from the Hong Kong and Taiwan was diminishing in this period. Given the historical facts and the patterns shown in Figure 4, it is reasonable to postulate some kind of correlation between the use of YOU-MEI-YOU+VP in Hong Kong and Taiwan and the (recent) development of this construction in Mainland Standard Mandarin.

4.3.2 Distribution of Auxiliary YOU-MEI-YOU in the 2000s

The second question that follows from what can be observed in the figures above concerns the current relevance of preverbal YOU-MEI-YOU. It seems that the frequency of use of auxiliary YOU-MEI-YOU has decreased since the turn of the 21th century. Judging from the data from the two corpora, not only did this period witness a comparatively lower frequency, but there were no instances of auxiliary YOU-MEI-YOU found that were produced after 2006. Do these observations mean that this construction is on its way out? In order to put this question in the right perspective, different factors, need to be considered, including the nature of the corpora and of the data used in this study, some facts and observations from outside the data analyzed here, and so on.

¹⁵⁹ A discussion on this topic could be found in Tian (2006).

First and foremost, it should be kept in mind that, in the CNC-CC corpus, no texts published after 1998 were available for search or accessible to outside users, though such texts may be present in the unabridged corpus (the abridged corpus contains approximately 20% of the unabridged version). The texts in the CCL corpus only go up through 2013, but even if the trend from 2000-2013 were to continue, there would still be a decline in the number of tokens per million characters from the 1970s onward.¹⁶⁰ However, it should be remembered that both corpora focus heavily on written new media. Since Standard Mandarin is somewhat regulated (recall the grammars cited in Chapter 1), it would be interesting to find out if the data reflect a history in which YOU-MEI-YOU+VP became popular enough in the 1970s to be explicitly flagged as ungrammatical in Mainland Standard Mandarin.

Whether or not this turns out to be the explanation for the trend found in the CCL and CNC corpora, there is evidence that YOU-MEI-YOU+VP has not lost ground in other genres. For instance, the Media Language Corpus (National Language Resources Monitoring and Research Center, 2013) contained transcriptions of episodes from different TV and/or radio programs in national and regional media played between 2008 and 2013. In a class project I conducted in December 2013, I used the same searching, filtering and analytical procedures as used in this dissertation to investigate the distribution of preverbal YOU-MEI-YOU in the Media Language Corpus. In a representative sample of 150 clauses analyzed for that project, 56 instances were found that contained YOU-MEI-YOU+VP (53) or YOU-MEI-YOU+Ambiguous (3).

¹⁶⁰ As stated in the official introductory text, about 10,000,000 characters from texts published after 2002 were collected, at the time when the introductory text was written (2009). It is also stated that “about 3 million characters are added to the corpus every year”, though it is not clear whether this still applies.

Although a detailed analysis of data from the Media Language Corpus is not possible within the scope of this study, and further investigation into the data from the CCL-CC corpus and the CNC-CC corpus is needed, at this point, one might not want to jump to the conclusion that preverbal YOU-MEI-YOU is fading away, or even losing momentum.

4.4 Chapter Summary

In this chapter, I summarized the results from various types of data analysis and discussed how they informed us of the grammatical status and development of preverbal YOU-MEI-YOU.

I started the chapter by presenting the percentage of preverbal YOU-MEI-YOU conveying a certain grammatical meaning, i.e. sentence type and aspect. I first showed that preverbal YOU-MEI-YOU not only helped mark clauses as belonging to the regular Interrogative category in the majority of the clauses. It also exhibited some unique characteristics and helped to form constructions that mark clauses as “non-canonical” interrogatives. I also presented results showing the active role that preverbal YOU-MEI-YOU plays in expressing viewpoint aspect in clauses. I also touched on the interaction and relationship between this construction and the resultative verb morpheme GUO. I argued that the hypothesis of this study, i.e. preverbal YOU-MEI-YOU is grammaticalizing into an auxiliary unit, was confirmed by the results of the data analysis.

In the second section of this chapter, I drew on the results from three types of diachronic comparisons of various linguistic features of preverbal YOU-MEI-YOU between different 20-year periods. These three types included the following: the

grammatical meaning conveyed by preverbal YOU-MEI-YOU, the general types of complement it took, and four sub-categories of VP complement following YOU-MEI-YOU. Based on the results of the comparison of the grammatical meaning and general types of complements of this construction, I postulated that this construction has reached Stage III described in Heine (1993). I further argued that it might have reached Stage IV or Stage V or an intermediate stage. I also briefly discussed the ambiguous cases, namely, clauses in which the complement of YOU-MEI-YOU could be interpreted as an NP or a VP. I suggested that such cases could help us better understand the grammaticalization of YOU-MEI-YOU. After that, I discussed the findings from the third type of comparison, that of the VP complement YOU-MEI-YOU took, which did not produce evidence useful for identifying its stage of development. I argued that a different kind of diagnostic tool or approach is needed to investigate the development of preverbal YOU-MEI-YOU.

In the last section, I summarized the results of the data analysis that looked at the frequency of use of preverbal YOU-MEI-YOU in different 10-year periods in time series plots, discussed how the “regional” data, data from Hong Kong and Taiwan texts, shaped the distribution of this construction in different periods and what that informs us of the development of the construction under investigation. I concluded the discussion by warning readers against drawing the conclusion that preverbal YOU-MEI-YOU was disappearing in the most recent time periods, by adducing evidence from different facts and observations, some of which are outside the scope of this study.

5. CONCLUSION

In this section, I will summarize the findings of the data analysis, discuss the limitations of the study and directions for future research on the YOU-MEI-YOU+VP structure.

5.1 Findings about the Use and Development of YOU-MEI-YOU+VP

This dissertation claimed that preverbal YOU-MEI-YOU grammaticalizing into an auxiliary unit, helping to form constructions that conveyed cross-linguistically recurrent grammatical meanings. The claim was based on studies that have shown a historical relationship between YOU MEI YOU+NP and YOU-MEI-YOU+VP (Dong, 2004; Kuang, 2000; Ota, 1958/1987; Shi & Li, 2001) and inspired by the approach to linguistic categorization in the framework of grammaticalization, in particular the approach to defining auxiliaries in Heine (1993) and Krug (2011). The investigation of two types of grammatical meaning, i.e. sentence type and aspect, conveyed by YOU-MEI-YOU+VP in the data collected from different corpora, produced evidence for the claim about the auxiliarihood preverbal YOU-MEI-YOU was achieving in Standard Mandarin: preverbal YOU-MEI-YOU was found to help form constructions that conveyed either of the grammatical meanings in the majority of the data, although such data did not account for the absolute majority of the total.

The question about the developmental stage of preverbal YOU-MEI-YOU was also taken up in this dissertation. It was addressed by conducting diachronic comparisons of three features of this construction, i.e. grammatical meanings conveyed, types of complements (e.g. verbal vs. ambiguous), types of VP complements (e.g. transitive vs. intransitive VPs), following the analysis of the grammaticalization of auxiliaries in Heine (1993) and the approach to identifying stages of development of possessive perfects in cross-linguistic studies (e.g. Heine & Kuteva, 2006; Łęcki, 2010). The results of the comparisons of the grammatical meanings and the general types of complements of preverbal YOU-MEI-YOU in different 20-year periods suggest that this construction has reached Stage III (out of the seven stages outlined by Heine, 1993). The last type of diachronic comparison, however, did not produce findings that could inform us of the developmental stage of preverbal YOU-MEI-YOU and I argued for the use of a different linguistic feature (other than the transitivity of VPs) or a different approach to analyzing the development of preverbal YOU-MEI-YOU.

The frequency of use of YOU-MEI-YOU+VP conveying a certain grammatical meaning, was examined in order to identify trends in its development. The results of a different set of diachronic comparison, one that looked at the frequency of use of auxiliary YOU-MEI-YOU in different 10-year periods, showed that there was not an upward trend in its use, except for a spike in the 1970-1979 period. A separate analysis was conducted to eliminate the effect of the imbalance in the amount of data from different corpora and three major observations were made in this analysis: an overall low percentage for different time periods, a sharp increase in the percentage in the 1970s, and a downward trend towards the most recent time period. Based on the first observation, it

was concluded that no evidence was found for the increase in the frequency of use of YOU-MEI-YOU+VP conveying a grammatical meaning over time.

With respect to the second observation made of the distribution, a further examination of the origin of the data (i.e. the regions in which the data were produced) revealed that the spike observed of the 1970-1979 period could be ascribed to the inclusion of a large amount of data produced by speakers from Hong Kong or Taiwan in the CCL-CC corpus for that particular time period. This finding lends some support to the speculation that some correlations existed between the use of YOU-MEI-YOU+VP in certain dialects or varieties and the development of this construction in Mainland Standard Mandarin. The decline in the percentage from the beginning of the 21st century was examined in relation to other factors, including the nature of the corpora, the type of data used in the study, and observations from outside this dissertation. It was argued that further investigation was needed before any conclusion could be drawn about the current relevance of YOU-MEI-YOU+VP.

5.2 Theoretical Implications

This study has shown that approaches to analyzing the development of grammaticalized items in previous research on grammaticalization were not as effective with Chinese as they were or are in the analysis of languages with morphological distinctions. Due to the typological differences between Standard Mandarin and such languages based on and for which the majority of theories of grammaticalization have been proposed, the analysis of YOU-MEI-YOU+VP drawing on theoretical constructs from grammaticalization studies (e.g. the parameter of decategorialization), which were

grounded on morphological features to some extent, did not reveal much about the developmental stage of this construction. In addition, the distinct characteristics of verb phrases (e.g. coverbs, serial verb constructions, pivotal constructions) and the flexible word order also add to the difficulty of identifying certain types of VPs as characteristic of a certain stage of development. A modification of theories and methodologies in the framework of grammaticalization is needed to accommodate Standard Mandarin so that future analyses could produce real insight into the development of YOU-MEI-YOU+VP and other constructions.

5.3 Limitations

The use of large corpora that span a long period of time has made it possible, to a certain extent, to investigate the emergence and development of YOU-MEI-YOU+VP, but the characteristics of each corpus used have been the source of major issues of concern for this dissertation. One such issue is that the data collected from the corpora are not balanced across time periods in the two corpora used, i.e. the CCL-CC corpus and the CNC-CC corpus. This means, when the distribution of instances of YOU-MEI-YOU+VP in different time periods is compared on the basis of such corpora data, the results may not truly reflect the frequency of use of this construction over time. Furthermore, the ratio of conversational texts to other genres is not uniform across time periods in either of these two corpora. Given that YOU-MEI-YOU questions are much less likely to appear in non-conversational texts, the normalization of data found in these corpora (i.e. using the target-clause-per-character ratio for each time period) may still fall short of reflecting the frequency of use of this construction over time.

Another issue that stems from the types of data used in the corpora concerns one vital piece of information needed for the data analysis in the current study: the year of composition or publication. This piece of information was simply not available for a large amount of data in the CCL-CC corpus. As discussed in Chapter 3, this problem was only partly solved by a manual search using the title of each text listed in the official introductory document, since for some entries, the year of composition or publication could not be determined, because they contained a collection of writings by a certain author or different authors and the titles of the writings that were part of such entries were not specified. This type of data was excluded from the calculation of the base size for a given time period. Since they accounted for about 9% of the total of the texts in the CCL-CC corpus, the inclusion of such data might have an effect on the data normalization based on which the frequency of use of YOU-MEI-YOU+VP over time has been plotted.

The problem with the CNC-CC corpus is different from those with the CCL-CC corpus in nature: the former provides the time information for the clauses found in the search and specifies the year of composition or publication for the texts included; however, only about a fifth of the data in the corpus is accessible to the outside users and the data for certain years, including the data collected after 1998, are not available for search. This may explain the small number of clauses containing YOU-MEI-YOU+VP found in this corpus. Since only a portion of the data is accessible, and it is not clear which part of the data in this corpus is inaccessible, it is difficult to measure the degree to which the patterns of use and distribution of YOU-MEI-YOU+VP observed on the basis of the available data, and the claims made on the basis of such observations in the

previous chapter would be affected if such data in the CNC-CC corpus were added to the picture.

Apart from the challenges associated with the types of data used in the study, the coding of data also needs to be considered when the claims and arguments made in this dissertation are evaluated. The analysis of the types of complements, in particular the types of VP complements, YOU-MEI-YOU took in the data was heavily dependent on the coding of the complements. For example, the criterion of complexity was based on findings of previous studies on YOU-MEI-YOU+VP and was not clearly defined. One type of complex VP was a verb taking a clausal complement, which was considered structurally more complex than a transitive VP with a lengthy NP complement. However, this type of VP could be subsumed under the simple transitive VP category, if their similarity was taken into account. Re-categorizing such instances would not necessarily change the overall results of the pertinent analysis, but it might have an effect on the distribution of the complex VPs in different time periods.

Another issue with the coding of complements lies in the definition of intransitive VPs in Standard Mandarin. Some verbs that could be used as transitive verbs were coded as Intransitive when found in a clause in which they did not take any object: for example, a verb would be coded as Intransitive if it was found in a clause without any passive marker in which the subject underwent the action denoted by the verb and the doer of the action was not expressed. This type of clause was distinguished from clauses in which the VP had an overt subject which was the doer of the action, and a preposed or topicalized object. The VPs in the latter case were coded as Simple Transitive or Complex VPs depending on the type of complement the verb took. It was, however, not always easy to

determine if the agent or doer of the action was simply implied in the context and phonetically null in the former type of clause. Furthermore, this type of clause whose subject expressed the theme or patient of the verb, could not be classified as Passive, since there was no overt passive marker, or middle voice, as the subject did not perform the action denoted by the verb upon itself.¹⁶¹ Consequently, the classification of the VPs in such clauses as intransitive should be evaluated against theoretical arguments and development concerning the syntax and semantics of verbs.

Additional limitations of this study are discussed in the next section on directions for future research projects, which could provide further theoretical and methodological underpinning for the research conducted for this study, and strengthen the arguments and claims made in this dissertation.

5.4 Future Research

This dissertation has been built on the findings of previous theoretical and empirical work on Chinese linguistics and grammaticalization. The assumption has been based on findings from prominent work on YOU-MEI-YOU, and the research questions as well as approaches to analyzing data have been informed and inspired by previous research on grammaticalization. In the meantime, the findings of this dissertation have revealed the inadequacy of theoretical and/or methodological constructs taken from earlier work in these areas and point to the need to examine the distribution of YOU-MEI-YOU+VP in different kinds of data and to look at the development of YOU-MEI-

¹⁶¹ The definition of middle voice was based on Nida (1952, p. 168) and Bybee (1985, pp. 20-21).

YOU+VP in relation to that of other linguistic constructions in order to glean insight into what drives its grammaticalization.

First and foremost, this dissertation did not investigate the relationship between the development of YOU-MEI-YOU+VP and that of YOU+VP and instead chose to focus on the former. It should be acknowledged that the reality may be much more complex. Although the first instances of YOU-MEI-YOU+VP documented in previous research and found in the data analysis in this study were earlier than the first appearance of YOU+VP in Standard Mandarin used in Mainland China as documented or indicated in some studies (e.g. Chao, 1968; Wang, 2012), no systematic research effort has been carried out to study the emergence and development of YOU+VP in Mainland Standard Mandarin over time. It would be interesting to compare the distribution of YOU+VP and that of YOU-MEI-YOU+VP in large corpora and see if there is any parallel between the development of YOU-MEI-YOU+VP and that of YOU+VP in Standard Mandarin. This kind of corpus-based study would be a tremendous undertaking, but it might shed light on whether the development of one has contributed to or fueled that of the other in Standard Mandarin.

By a similar token, a thorough and theory-based investigation of YOU MEI YOU+NP that traces its development in large corpora that span a long historical period is necessary for a better understanding of the relationship between YOU MEI YOU+NP and YOU-MEI-YOU+VP. This line of inquiry should focus on locating or identifying instances that (seem to) characterize the intermediate stage, as posited in previous work, along the development of YOU-MEI-YOU (from the A-not-A construction of a lexical verb to a preverbal, construction conveying some grammatical meaning). Finding from

such investigations could lend further support to the arguments about the historical relationships between YOU MEI YOU+NP and YOU-MEI-YOU+VP.

The above-mentioned suggestions for future research are primarily concerned with what has been assumed about YOU-MEI-YOU+VP in this current study regarding the origin and the non-compositional nature of preverbal YOU-MEI-YOU. With respect to data analysis, an expanded project should include data from spoken corpora in which YOU-MEI-YOU+VP is more likely to be found, and examine the use of this construction in spoken Chinese. Ideally, such spoken corpora should cover the latest time period so that it will be possible to determine the current relevance of the construction in question. The Media Language Corpus, with data from different TV and radio programs broadcast in China from 2008 to 2013 is a good example of such corpora. An investigation into the frequency of use of this construction in this kind of data could produce evidence for its presence in the language and its recent development. Furthermore, such data contain a lot of information about the social contexts in which YOU-MEI-YOU+VP questions were produced (e.g. the regions in which the conversation occurred, the gender and age of the speaker and of the addressee) and they are also likely to contain the responses to YOU-MEI-YOU+VP questions, and studying the use of the construction in relation to such contextual clues would give a fuller picture of its development and integration in Standard Mandarin.

REFERENCES

REFERENCES

- Barth-Weingarten, D., & Couper-Kuhlen, E. (2002). On the development of final *though*: A case of grammaticalization? *Typological Studies in Language*, 49, 345-362.
- Benveniste, E. (1968). Mutations of linguistic categories. In W. P. Lehmann & Y. Malkiel (Eds.), *Directions for historical linguistics. A symposium* (pp. 83-94). Austin, TX & London: University of Texas Press.
- Bolinger, D. (1980). *Wanna* and the gradience of auxiliaries. *Wege zur Universalienforschung: Sprachwissenschaftliche Beiträge zum*, 60, 292-299.
- Bybee, J. L. (1985). *Morphology: A study of the relation between meaning and form*. Amsterdam: John Benjamins Publishing.
- Bybee, J. (2003). Mechanisms of change in grammaticization: The role of frequency. In B. D. Joseph & R. D. Janda (Eds.), *The handbook of historical linguistics*, (pp. 602-623). Malden, MA: Blackwell Publishing Company.
- Bybee, J., Perkins, R., & Pagliuca, W. (1994). *The evolution of grammar: Tense, aspect, and modality in the languages of the world*. Chicago: University of Chicago Press.
- Cai, T. (2009). 蔡琪. 上海高校学生“有+VP”句使用情况调查分析. *语言教学与研究*, 6, 79-86.

- Ceong, H. H., & Saxon, L. (2013). Features of questions and interrogatives. *Proceedings of the 2013 Annual Conference of the Canadian Linguistic Association*. Toronto, Canada: University of Toronto.
- Center for Chinese Linguistics PKU. (2009). *Center for Chinese Linguistics Corpus*.
http://ccl.pku.edu.cn:8080/ccl_corpus
- Chao, Y. R. (1968). *A grammar of spoken Chinese*. Berkeley: University of California Press.
- Character (computing). (April 28, 2016). In *Wikipedia*. Retrieved June 20, 2016, from [https://en.wikipedia.org/wiki/Character_\(computing\)](https://en.wikipedia.org/wiki/Character_(computing))
- Chen, P. (1999). *Modern Chinese: History and sociolinguistics*. Cambridge, UK: Cambridge University Press.
- Cheng, L. L. S., Huang, C. T. J., & Tang, C. C. J. (1997). Negative particle questions: A dialectal perspective. *Journal of Chinese Linguistics Monograph Series*, 10, 65-112.
- Chen, Z., & Chen, Z. (n.d.). 陈振宇& 陈振宁. 疑问形式、疑问标记和命题结构. Retrieved from <attach3.bdwm.net/attach/.../疑问形式、疑问标记和命题结构 1.pdf>
- Cheung, H. N. S., Liu, S. Y., & Shih, L. L. (1994). *A practical Chinese grammar*. Hong Kong: Chinese University Press.
- Chierchia, G., & McConnel-Ginet, S. (1990). *Meaning and grammar: An introduction to semantics*. Cambridge, MA: MIT Press.
- Chinese character encoding. (July 19, 2016). In *Wikipedia*. Retrieved July 24, 2016, from https://en.wikipedia.org/wiki/Chinese_character_encoding

- Comrie, B. (1976). *Aspect: An introduction to the study of verbal aspect and related problems*. Cambridge, UK: Cambridge University Press.
- Coussé, E. (2013). The grammaticalization of the have perfect in Dutch: A corpus study of contextual extension and semantic generalization. *Language Sciences*, 36, 103–112.
- de Condillac, E. B. (1973 [1746]). *Essai sur l'origine des connaissances humaines*. Paris: Galilée.
- Deng, S. (2004) 邓思颖. 作格化和汉语被动句. *中国语文*, 4, 291-301.
- de Swart, H. (2011). Verbal aspect across languages. In R. I. Binnick (Ed.), *The Oxford handbook of tense and aspect*, (pp. 752–780). Oxford, UK: Oxford University Press.
- Diao, Y. (2012). 刁晏斌. 两岸四地“有+VP”考察. *励耘学刊*, 1, 27-51.
- Diessel, H. (2007). Frequency effects in language acquisition, language use, and diachronic change. *New Ideas in Psychology*, 25(2), 108-127.
- Diewald, G. (2006). Context types in grammaticalization as constructions. *Constructions* SV1-9/2006.
- Ding, S. (1961). 丁声树. *现代汉语语法讲话*. 北京: 商务印书馆.
- Dong, X. (2004). 董秀芳. 现代汉语中的助动词“有没有.” *语言教学与研究*, 2, 1–8.
- Dowty, D. R. (1979). *Word meaning and Montague grammar: The semantics of verbs and times in Generative Semantics and in Montague's PTQ*. Dordrecht, Netherlands: D. Reidel.
- Ernst, T. (1995). Negation in Mandarin Chinese. *Natural Language & Linguistic Theory*, 1, 665–707.

- Filip, H. (2012). Lexical aspect. In R. I. Binnick (Ed.), *The Oxford handbook of tense and aspect*, (pp. 721-751). Oxford, UK: Oxford University Press.
- Garcia, E. C. (1967). Auxiliaries and the criterion of simplicity. *Language*, 43(4), 853-870.
- Garzonio, J. (2004). Interrogative types and left periphery: Some data from the Fiorentino dialect. *Quaderni di lavoro dell'ASIS*, 4, 1-19.
- Givón, T. (1979). *On understanding grammar*. New York: Academic Press.
- Guo, R. (2002). 郭锐. *现代汉语词类研究*. 北京: 商务印书馆.
- Heine, B. (1993). *Auxiliaries: Cognitive forces and grammaticalization*. New York: Oxford University Press.
- Heine, B., Claudi, U., & Hünnemeyer, F. (1991). *Grammaticalization: A conceptual framework*. Chicago: University of Chicago Press.
- Heine, B., & Kuteva, T. (2002). *World lexicon of grammaticalization*. Cambridge, UK; New York: Cambridge University Press.
- Heine, B., & Kuteva, T. (2006). *The changing languages of Europe*. Oxford, UK: Oxford University Press.
- Heine, B., & Kuteva, T. (2007). *The genesis of grammar: A reconstruction*. Oxford, UK: Oxford University Press.
- Heine, B., & Reh, M. (1984). *Grammaticalization and reanalysis in African languages*. Hamburg, Germany: Helmut Buske.
- Hidalgo-Downing, L. (2000). *Negation, text worlds, and discourse: The pragmatics of fiction*. New York: Greenwood Publishing Group.

- Hopper, P. J., & Traugott, E. C. (2003). *Grammaticalization*. Cambridge, UK: Cambridge University Press.
- Huang, C.T. J. (2007). 黄正德. 汉语动词的题元结构与其句法表现. *语言科学*, 6(4), 3-21.
- Huang, C.T. J., & Li, Y. H. A., & Li, Y. (2009). *The syntax of Chinese*. Cambridge, UK: Cambridge University Press.
- Huang, Y. (2015). 黄雅思. 否定词“没(没有)”的历史演变探析. *新课程研究*, 1, 113-115.
- Jakobson, R. (1959). Boas' view of grammatical meaning. *American Anthropologist*, 61(5), 170-178.
- Jespersen, O. (1917 [1966]). *Negation in English and other languages*. Copenhagen, Denmark: A.F. Høst.
- Kaisse, E. M. (1981). Luiseño particles and the universal behavior of clitics. *Linguistic Inquiry*, 12(3), 424-434.
- Kiparsky, P. (2012). Grammaticalization as optimization. In D. Jonas, J. Whitman, & A. Garrett (Eds.), *Grammatical change: Origins, nature, outcomes*, (pp. 15-51). Oxford, UK; New York: Oxford University Press.
- Krug, M. (2011). Auxiliaries and grammaticalization. In H. Narrog, & B. Heine (Eds.), *The Oxford handbook of grammaticalization*, (pp. 547-558). Oxford, UK: Oxford University Press.
- Kuang, X. (2000). 邝霞. “有没有”反复问句的定量研究——对经典作家白话文作品的定量研究. *汉语学习*, 3, 69-72.

- Kuteva, T. (1995). The auxiliarization constraint and reference. In R. A. Geiger (Ed.), *Reference in multidisciplinary perspective: Philosophical object cognitive subject, intersubjective process*, (pp. 374-386). Hildesheim, Germany; New York: G. Olms Verlag.
- Kuteva, T. (2001). *Auxiliation: An enquiry into the nature of grammaticalization*. New York: Oxford University Press.
- Larsen-Freeman, D., & Cameron, L. (2008). *Complex systems and applied linguistics*. New York: Oxford University Press.
- Laury, R. (1997). *Demonstratives in interaction: The emergence of a definite article in Finnish*. Amsterdam; Philadelphia: John Benjamins Publishing.
- Łęcki, A. M. (2010). *Grammaticalisation paths of 'have' in English*. Frankfurt am Main, Germany: Peter Lang Publishing.
- Lehmann, C. (2002 [1982]). *Thoughts on grammaticalization. Arbeitspapiere des Seminars für Sprachwissenschaft der Universität Erfurt* (Vol. 9). Erfurt, Germany: Seminar für Sprachwiss. der Universität.
- Li, C. N., & Thompson, S. A. (1976). Subject and topic: A new typology of language in subject and topic. In C. Li (Ed.), *Subject and Topic*, (pp. 297-329). Amsterdam: John Benjamins Publishing.
- Li, C. N., & Thompson, S. A. (1989). *Mandarin Chinese: A functional reference grammar*. Berkeley: University of California Press.
- Li, Y., & Tang, Z. (1991). 李宇明&唐志东. 汉族儿童问句系统习得探微. 武汉: 华中师范大学出版社.

- Lin, J. W. (2012). Tenselessness. In R. I. Binnick (Ed.), *The Oxford handbook of tense and aspect*, (pp. 669-695). Oxford, UK: Oxford University Press.
- Lin, J. W., & Tang, C. C. J. (1995). Modals as verbs in Chinese: A GB perspective. *中央研究院歷史語言研究所集刊*, 66(1), 53-105.
- Lin, T. H. J. (2012). Multiple-modal constructions in Mandarin Chinese and their finiteness properties. *Journal of Linguistics*, 48(1), 151-186.
- Liu, L. (2008). 刘丽春. “有没有 VP” 句法格式的发展历程及其对“有 VP” 在普通话中地位的影响. *文教资料*, 28, 36-38.
- Liu, M. (2010). 刘敏. 汉语否定副词来源与历时演变研究 (Unpublished master's thesis). 湖南师范大学, Changsha, China.
- Liu, Z. (1998). 刘照熊. *普通话水平测试大纲 (修订本)*. 长春: 吉林人民出版社.
- Lü, J. (2000). 吕冀平. *当前我国语言文字的规范化问题*. 上海: 上海教育出版社.
- Lü, S. (1982). 吕叔湘. *中国文法要略*. 北京: 商务印书馆.
- Lü, S. (1985). 吕叔湘. 疑问·否定·肯定. *中国语文*, 4, 241-250.
- Lu, K. (2010). 鲁克伟. 论新兴语法标记“有”——语言接触促使的语法化. *理论语言学研究*, 4, 78-90.
- Lyons, J. (1977). *Semantics*. Cambridge, UK: Cambridge University Press.
- Ma, L. (2010). The syntactic status of Chinese negators—mei and meiyou. *Working Papers of the Linguistics Circle*, 15, 1-11.
- McCawley, J. D. (1992). Justifying part-of-speech assignments in Mandarin Chinese / 為劃分漢語詞類提供理論依據. *Journal of Chinese Linguistics*, 20(2), 211-246.
- Meillet, A. (1912). L'évolution des formes grammaticales. *Revue internationale de synthese scientifique*, 12, 130-158.

- Morgan, J. L. (1977). Two types of convention in indirect speech acts. *Center for the Study of Reading Technical Report; no. 052*.
- National Language Resources Monitoring and Research Center, Broadcast Media Language Branch. (2013). *Media Language Corpus*.
<http://ling.cuc.edu.cn/RawPub/>
- Nichols, J., & Timberlake, A. (1991). Grammaticalization as retextualization. In E.C. Traugott, & B. Heine, B. (Eds.), *Approaches to grammaticalization, Vol. 1: Theoretical and methodological issues*, (pp. 129-146). Amsterdam/Philadelphia: John Benjamins Publishing.
- Nida, E. A. (1952). *Morphology: The descriptive analysis of words*. Ann Arbor: University of Michigan Press.
- Ota, T. (1958/1987). 太田辰夫. *中国语历史文法*. 北京: 北京大学出版社.
- Peng, R. (2011). 彭睿. 临界频率和非临界频率——频率和语法化关系的重新审视. *中国语文*, 1, 3-18.
- Portner, P. (2004). The semantics of imperatives within a theory of clause types. *Proceedings of Semantics and Linguistic Theory*, 14, 235-252.
- Reis, M. (1976). Zum grammatischen Status der Hilfsverben. *Beiträge zur Geschichte der deutschen Sprache und Literatur (PBB)*, 98, 64-82.
- Reuland, E. J. (1983). Government and the search for AUXes: A case study in cross-linguistic category identification. In F. Henry, & B. Richards (Eds.), *Linguistic categories: Auxiliaries and related puzzles*, (pp. 99-168). Dordrecht, Netherlands: Kluwer Academic Publisher.

- Rittenhouse, R. (2014). *Verbal periphrasis in two early Germanic languages: A comparative study of the passive and perfect in the Old High German Evangelienbuch and the Old Saxon Heliand* (Unpublished doctoral dissertation). University of Wisconsin-Madison, Madison, WI.
- Shao, J. (1996). 邵敬敏. 现代汉语疑问句研究. 上海: 华东师范大学出版社.
- Shen, L. (2007). 沈莉娜. 现代汉语中的“没有”作为疑问语气词的思考. *景德镇高专学报*, 22(1), 41-43.
- Shen, Y. (2006). 申云玲. “有没有”正反问句答句的对称与不对称. *忻州师范学院学报*, 22(1), 78-80.
- Shi, Y. (2002). *The establishment of modern Chinese grammar: The formation of the resultative construction and its effects*. Amsterdam: John Benjamins Publishing.
- Shi, Y., & Li, N. C. (2001). 石毓智 & 李讷. 汉语语法化的历程: 形态句法发展的动因和机制. 北京: 北京大学出版社.
- Smith, C. S. (1997). *The parameter of aspect* (2nd edition). *Studies in Linguistics and Philosophy* (Vol. 43). Dordrecht, Netherlands: Kluwer Academic Publishers.
- State Language Commission. (1993-2012). *State Language Commission Chinese National Corpus*. <http://www.cncorpus.org/>
- Sun, Q. (2003). 孙琴. “有”能否作为“有没有VP”句的肯定回答?. *语文学刊: 高等教育版*, 5, 76-77.
- Sweetser, E. E. (1988). Grammaticalization and semantic bleaching. *Annual Meeting of the Berkeley Linguistics Society*, 14, 389-405.

- Tang, H. (2008). 汤洪丽. 汉语正反问句的历时演变考察 (Master's thesis). Retrieved from <http://cdmd.cnki.com.cn/Article/CDMD-10542-2010126819.htm>
- Tang, Y. (1981). 湯廷池. 國語疑問句的研究. *師大學報*, 26, 219-276.
- Tang, Y. (1984). 湯廷池. 國語疑問句研究續論. *師大學報*, 29, 381-435.
- Tian, Z. (2006). 田志凌. 新派武侠小说传入内地 南风窗开新武侠长驱直入. *南方都市报*. Retrieved from http://www.gzlib.gov.cn:8080/shequ_info/ndgz/NDGZDetail.do?id=18344
- Traugott, E. C., & Heine, B. (Eds.). (1991). *Approaches to grammaticalization, Vol. 1: Theoretical and methodological issues*. Amsterdam; Philadelphia: John Benjamins Publishing.
- Traugott, E. C. (2012). On the persistence of ambiguous linguistic contexts over time: Implications for corpus research on micro-changes. In J. Mukherjee, & M. Huber (Eds.), *Corpus linguistics and variation in English: Theory and description*, (pp. 231-246). Amsterdam: Rodopi.
- Vendler, Z. (1957). Verbs and times. *The Philosophical Review*, 66(2), 143-160.
- Verkuyl, H. J. (1972). *On the compositional nature of the aspects*. *Foundations of Language Supplementary Series* (Vol. 15). Dordrecht, Netherlands: D. Reidel Publishing Company.
- von Schlegel, A. W. (1818). *Observations sur la langue et la littérature provençales*. Paris: Librairie grecque-latine-allemande.
- Wang, L. (1958). 王力. *汉语史稿* (Vol. 2). 北京: 科学出版社.
- Wang, L. (1985). 王力. *中国现代语法*. 北京: 商务印书馆.

- Wang, S., Wang, Y., & Jiang, L. (2006). 王森, 王毅, & 姜丽. “有没有/有/没有+ VP”句. *中国语文*, 1, 10-18.
- Wang, S.Y. (1965). Two aspect markers in Mandarin. *Language*, 41, 457-470.
- Wang, Z. (2012) 王周明. 现代汉语普通话“有+ VP”用法的现状及其与方言的关系试析. *言语文化*, 15(1), 97-117.
- Wu, F. (1995). 吴福祥. 否定副词“没”始见于南宋. *中国语文*, 2, 153-153.
- Xiang, Y. (2014). Obligatory neg-raising in Mandarin: Negatives and aspects. Unpublished manuscript, Department of Linguistics, Harvard University, Boston.
- Xing, F. (1990). 邢福义. “有没有 VP”疑问句式. *华中师范大学学报(人文社会科学版)*, 1, 86-91.
- Xu, S. (2003). 徐时仪. 否定词“没”“没有”的来源和语法化过程. *湖州师范学院学报*, 25(1), 1-6.
- Xue, H. (2010). 薛宏武. 现代汉语“有没有”结构的形成. *宁夏大学学报(人文社会科学版)*, 32(2), 46-50.
- Yang, R. (1999). 杨荣祥. 近代汉语否定副词及相关语法现象略论. *语言研究*, 1, 20-28
- Yuan, Y. (2010). 袁毓林. *汉语词类的认知研究和模糊划分*. 上海: 上海教育出版社.
- Zhang, B. (2002). 张斌. *新编现代汉语*. 上海: 复旦大学出版社.
- Zhang, H. (2009). 张慧. 现代汉语非宾格结构的内部句法差异. *Proceedings of the 21st North American Conference on Chinese Linguistics (NACCL-21)*, 2, 417-435.
- Zheng, Y. (2001). 郑燕萍. *普通话训练测试教程*. 北京: 中国广播电视出版社.
- Zhou, L. (2012) 周莉. 两岸汉语的正反问“有没有 VP”的对比研究. *华夏文化论坛*, 7: 158-164.
- Zhu, D. (1982). 朱德熙. *语法讲义*. 北京: 商务印书馆.

VITA

VITA

I am from a small city in China, and grew up speaking Hakka at home, Standard Mandarin in school and Cantonese with friends. I was an English education major in my undergraduate program and over time, I developed an interest in how language is acquired, learned and used. I spent two and a half years studying (applied) linguistics in my master's program in Beijing before coming to the United States to pursue a doctorate in linguistics.

My research interests cover a wide range of areas including sociolinguistics, historical linguistics, semantics, phonology, discourse analysis and World Englishes. I am interested in how language use affects and is affected by social and interactional contexts. In one of my projects, I looked at how Chinese speakers expressed disagreement in real-time online communication and gained a better understanding of how the use of certain linguistic strategies was constantly negotiated in the conversation.

I am also intrigued by how language develops and changes over time. My current research focuses on one new syntactic construction in Mandarin Chinese. The goal is to gain some insight into the social and linguistic forces that shape how language develops by investigating the development and usage of this new construction. Another research project I have in mind on language change is related to the study of World Englishes.

More specifically, I want to examine how English is used in Chinese media to find out whether a Chinese variety of English is in the process of formation.