SENTENCE-FINAL PARTICLES (SFPS):

A USAGE-BASED CONSTRUCTIONIST APPROACH

by

LIN ZHU

A DISSERTATION

Presented to the Department of East Asian Languages and Literatures
and the Division of Graduate Studies of the University of Oregon
in partial fulfillment of the requirements
for the degree of
Doctor of Philosophy

June 2021

DISSERTATION APPROVAL PAGE

Student: Lin Zhu

Title: Sentence-Final Particles (SFPs): A Usage-based Constructionist Approach

This dissertation has been accepted and approved in partial fulfillment of the requirements for the Doctor of Philosophy degree in the Department of East Asian Languages and Literatures by:

Zhuo Jing-Schmidt	Chairperson
Kaori Idemaru	Core Member
Nayoung Kwon	Core Member
Spike Gildea	Institutional Representative

and

Andrew Karduna	Interim Vice Provost for Graduate Studies

Original approval signatures are on file with the University of Oregon Division of Graduate Studies.

Degree awarded June 2021

DISSERTATION ABSTRACT

Lin Zhu

Doctor of Philosophy

The Department of East Asian Languages and Literatures

June 2021

Title: Sentence-Final Particles (SFPs): A Usage-based Constructionist Approach

This dissertation adopts the usage-based constructionist approach to investigate three sentence-final particles (SFPs) in Mandarin Chinese, *bei*, *ne*, and *a*, with the help of British corpus linguistics techniques. The study is based on data collected from the BCC corpus and the instant communication tool WeChat. The results show that their functions are best understood holistically as part of those constructions, instead of associating functions with SFPs as morphemes. This study therefore demystifies the long-held belief that SFPs are elusive.

CURRICULUM VITAE

NAME OF AUTHOR: Lin Zhu

GRADUATE AND UNDERGRADUATE SCHOOLS ATTENDED:

University of Oregon, Eugene
Shanghai International Studies University, Shanghai, China
Henan University, Kaifeng, China

DEGREES AWARDED:

Doctor of Philosophy, 2021, University of Oregon
Master of Arts, 2017, University of Oregon
Master of Arts, 2015, Shanghai International Studies University
Bachelor of Arts, 2012, Henan University

AREAS OF SPECIAL INTEREST:

Chinese Linguistics
English Language and Literature

PROFESSIONAL EXPERIENCE:

Teaching Assistant, University of Oregon, 2015-2020

Tutor and Domain Coach of Chinese Language Flagship Program, University of Oregon, 2016-2020

GRANTS, AWARDS, AND HONORS:

Graduate Teaching Fellowship, East Asian Languages and Literatures, 2015 to 2020

PUBLICATIONS:

Jing-Schmidt, Z., Lang, J., Shi, H.H., Hung, S.H. & Lin, Z. (2022) Aspect construal in Mandarin: A usage-based constructionist perspective on LE. Linguistics. (To appear).

Zhu, L. (2019). Kinship metaphors in the Chinese construction A shi B zhi fu/mu: Biology and culture as conceptual basis. In Y. Xiao & L. Tsung (Eds.), *Current Studies in Chinese Language and Discourse: Global context and diverse perspectives*. John Benjamins Publishing Company.

ACKNOWLEDGMENTS

My special thanks go to Professor Zhuo Jing-Schmidt, who is the lighthouse and the accelerator of my academic progress in the whole PhD program. She offers me valuable guidance during the early planning phase of this dissertation and supervises me throughout. She teaches me “Excellence cannot be acquired without constant practice.” She will be my role model for my career in academia. Thank you again, Jing Laoshi, for giving me this opportunity to study in the states. Thank you so much for being the chair of my dissertation committee.

I thank other EALL faculty members for your great patience in helping me to adapt to student life in US, especially in my early years at University of Oregon (UO). I couldn’t survive till today without Professor Lucien Brown and Professor Kaori Idemaru’s generous help. The two professors witnessed every step of my progress in the PhD program. I couldn’t survive without the help of Denise Gigliotti, Fengjun Mao, Dr. Jean Wu, and Weijun Chen in fulfilling the Teaching Assistant (TA) duties at UO. Special thanks go to Denise Gigliotti for offering me the building block of teaching Mandarin Chinese.

I am grateful for the whole faculty members of the Language Teaching Studies (LTS) program in the Department of Linguistics at the University of Oregon. They teach me how to teach languages and this plays a decisive role in my work as a TA. My gratitude especially goes to Dr. Keli Yerian, the Director of the LTS program, who showed considerable patience and provided considerable support for me, without whom

I would not be able to complete the task at all.

I thank Professor Kaori Idemaru, Professor Nayoung Kwon, and Professor Spike Gildea, as my dissertation committee members, for their great patience and valuable assistance in the preparation and revision of this dissertation. I also cordially thank them for their active participation in the oral defense, during which they provided me with considerably constructive questions and comments.

My sincere gratitude is due to University of Oregon for offering me the fellowships as TA. This financial support makes this dissertation and my years here possible. Special thanks go to all students in my teaching sessions, whose active participation provides me with great encouragement.

I am also thankful for my peers, friends, and classmates. Steffi Hung recommended the 2019 new book on construction grammar by Goldberg. Jun Lang sent me quite a few WeChat screenshots. I also thank my buddies Hao Mao and Ge Song for their valuable input and thoughts on this dissertation, and thank Yi Ren, Zhongyi Qiu, Shuangting Xiong, Yunzhu Tao, Xiang Li, Dan Li, Kuo Xie, Xiao Tan, and Fei Pei for their active cooperation in preliminary data collection. Apart from these folks, thank you all for accompanying, cheering up, and joking with me. I could not survive without you.

Last but not least, my gratitude goes to my parents, who provide great emotional and financial support over the years. Without you, I would not be able to be a dreamer.

To my parents.

TABLE OF CONTENTS

LIST OF FIGURES

Figure Page

Figure Page

LIST OF TABLES

Table Page

Table Page

Table Page

GLOSSING TERMS

A *a* (an SFP)
CL classifier
DE *de* (a particle)
EXP This is the glossing for *guo*, an experiential marker in Mandarin.
IMP This is the glossing for *zhe*, an imperfective marker in Mandarin.
LE *le* (an SFP)
MA *ma* (an SFP)
ME *me* (an SFP)
MIR mirative marker
NE *ne* (an SFP)
NEG negation
NP noun phrase
SG/sg singular form
PFV perfective
PL plural form
PRT particle
VOC vocative
VP verb phrase
1 first person
2 second person
3 third person
n. noun
v. verb

CHAPTER I

INTRODUCTION

1.1. Overview

Sentence Final Particles (SFPs) in Mandarin Chinese are defined by Li and Thompson (1981) as:

> "All of them (SFPs) are destressed and have the neutral tone. They typically occur in speech or in writings that reflect or recount conversations. Their semantic and pragmatic functions are elusive, and linguists have had considerable difficulty in arriving at a general characterization of each of them". (p. 238)

Goddard (2005) defines them as "little words which express a speaker's immediate 'here-and-now' emotions, thoughts, and desires towards what he or she is saying" (p. 144). Enfield (2005) thinks the use of SFPs is "a basic mode of distinguishing illocutionary force at the utterance level" (p. 190).

SFPs are an Asian areal feature, especially common in East Asian and Mainland Southeast Asian languages (Enfield, 2005; Erbaugh, 1985). This areal feature is considered a typological correlate of the tonal feature of these languages (Feng 2015). Hancil (2015) states that, "East and Southeast Asian languages have a complex system of particles, which constitute one of the 'hallmarks of natural conversation' in these languages (Luke [1990: 11] for Cantonese)" (p. 13). SFPs are occasionally given other names in the literature, such as "utterance particle" (e.g. Luke, 1990; Tang, 2015), "utterance-final particle" (e.g. Chu, 2006; Li, 2013), and "modal particle" (e.g. Chappell, 1991). "SFP" is the most commonly deployed term in studies of Mandarin and it is therefore used in this dissertation.

All regional varieties of Chinese have SFPs (e.g. Chao, 1926; Jing-Schmidt, 2019). Cantonese and Mandarin are the most researched with regard to SPFs (e.g. Chan, 2002; Jing-Schmidt, 2019). Researchers on Chinese SFPs generally agree that in Chinese SFPs are a colloquial feature that conveys a host of discourse organizational and

interactional functions (Chao, 1968; Chappell, 1991; Jing-Schmidt, 2019; Simpson, 2014; Xiang, 2011). However, when it comes to specific SFPs, there is little consensus on their functions and the only agreement seems to be the recognition that their meanings are elusive (Jing-Schmidt, 2021; Li & Thompson, 1981, p. 238; Qi, 2002, inter alia). The "elusiveness" is an obstacle for linguistic description and analysis of SFPs (Lee-Wong, 1998, p. 388). It is also maintained that the patterns of SFPs' distribution in discourse is *fenfanfuza* 'numerous and complicated' (Zhao & Sun, 2015). Callier (2007) used the expression "highly polyvalent" when discussing the meanings and functions of SFPs (p. 5). Moreover, but at the same time understandably, SFPs pose a challenge for second language learners (Han, 1988, p. 64).

To get a sense of the numerosity and complexity of Mandarin SFPs, we only need to open a Chinese dictionary where you will find that every SFP has a dozen entries. Let's take the dictionary explanation for the SFP *a* as an example. I use the 7th digital edition of 现代汉语词典 *xiandai hanyu cidian* 'The Dictionary of Modern Chinese' (CASS, 2019) as my reference book. For the sake of convenience, I will use capitalized "The DMC Dictionary" to just refer to this dictionary hereafter. The DMC Dictionary uses the following entries to explain *a*:

- 用在感叹句末，表示增强语气 'Used at the end of an exclamative sentence to strengthen the overtone'
- 用在陈述句末，使句子带上一层感情色彩 'Used at the end of a declarative sentence to add a flavor of emotion to the sentence'
- 用在祈使句末，使句子带有敦促或提醒意味 'Used at the end of an directive sentence to add a hortatory tone or a sense of reminder'
- 用在疑问句末，使疑问语气舒缓些 'Used at the end of an interrogative sentence to soften the tone'
- 用在句中稍做停顿，让人注意下面的话 'Used to mark a pause in sentence so as to remind the listener to pay attention to following utterances'
- 用在列举的事项之后 'Used after listed items'
- 用在重复的动词后面，表示过程长 'Used after repeated verbs to indicate the

lengthiness of action'

As it reflects, just one SFP can express so various meanings. Likewise, the SFP *ba* is believed by Chinese linguists to have the following functions, according to Han (1988, p. 1):

- indicating interrogative mood
- marking yes-no questions
- a similar function to that of tag-questions
- a function comparable to rising intonation in English
- changing declarative sentences into interrogatives
- indicating imperative[1] mood
- making a sentence advisive
- making a sentence into a mild command or suggestion
- expressing uncertainty in doubtful posed statements
- occur in pleas
- occur in dilemmas
- etc.

This list means that SFP *ba* can be deployed to fulfill more than 11 functions in communication. This phenomenon is expressed as "the functional diversity of individual particles" by Hancil, Post, and Haselow (2015, p. 15). SPFs are ubiquitous in Chinese conversation and are among the most frequently used morphemes in the language. For example, the SFP *ne* ranks 84, *ba* ranks 135, and *a* ranks 288, on a list of the most frequently used 14629 words in Mandarin Chinese, provided by Institute of Applied Linguistics at Ministry of Education of China[2], henceforth IALMEC.

In what follows, I provide a list of the prototypical[3] SFPs in the literature (the list was summarized by Chappell (1991) and Callier (2007)), with examples that characterize their usages and illustrate the "elusive" nature of their meanings.

---

[1] I follow Zhan and Bai (2016) in terming "imperative" as "directive." "Imperative" is maintained in citations if it is used by scholars of previous studies.

[2] IALMEC, (n.d.), 语料库在线 *yuliaoku zaixian* [Corpus Online]. http://corpus.zhonghuayuwen.org/index.aspx

[3] There are other less typical SFPs, such as *ha* (Yang & Wiltschko, 2016), *la*, which is believed to be the fusion of *le* and *a* (Jing-Schmidt, 2019, p. 4; Zhu, 1984, p. 207), *ei* (Zhu, 1984, pp. 212-213) and compound SFPs, such as *laizhe* (Sun, 2006, p. 81; Zhu, 1984, p. 209), *haole*, *suanle*, *dele*, *chengle*, and *xingle* (Yap et al., 2010).

- *le* (e.g. Fang, 2018; Jing-Schmidt et al., 2021; C. Li & Thompson, 1981)

SFP *le* can occur after a verbal predicate, an adjectival predicate, or nominal predicate, such as in Example (1), Example (2), and Example (3).

(1) Ta chu qu mai dongxi le
she exit go buy thing LE
'She's gone shopping.' (adapted from Li & Thompson, 1981, p. 240)

(2) ni zhongyu hao le, tai hao le!
2SG finally well LE too good LE
'You are finally well, that's so great!' (adapted from Jing-Schmidt et al., 2021, p. 2)

(3) Ta liang sui le
he two year LE
'He is (already) two years old.' (Jing-Schmidt, personal communication)

As "arguably the most controversial grammatical element in Chinese linguistics and the most confusing grammar point for Chinese second language learners" (Jing-Schmidt et al., 2021, p. 1), this SFP "mysteriously" invites a two-fold interpretation of completion and commencement. For instance, in Example (1), *Ta chu qu mai dongxi le* 'She's gone shopping' tells an addressee that a person has completed all decision-making and preparatory processes of shopping and has left the place and accordingly is now in a state of doing shopping and therefore not present at the place where the conversation happens. The presence of *le* in this utterance enables such a heavy semantic load. Furthermore, Fang (2018) argues that *le* can actually functions as a mirative marker to mark newsworthiness or surprise for both speakers and addressees, which goes against the traditional belief that Mandarin does not have mirative markers. One of Fang's examples is as follows,

(4) ta gao le
3SG tall LE.MIR
'He has gotten tall!' (Li & Thompson (1981, p. 250), adapted and cited from Fang (2018, p. 591))

- *ba*

SFP *ba* can occur in declarative sentences, as in Example (5), directive sentences, as in Example (6), and interrogative sentences, as in Example (7).

(5) ta bu hui zuo zhe-yang-de shi ba

3SG not will do this-manner-DE things BA

'S/He wouldn't do such things, don't you agree?' (Li & Thompson, 1981, p. 309)

(This example is a declarative sentence, although the free English translation in Line 3 ends with a tag question. This translation only serves the purpose to illustrate the message conveyed by the original sentence.)

(6) women zou ba

1PL go BA

'Let's go!' (Li & Thompson, 1981, p. 307)

(7) nimen shi yi ge cun de ni yinggai zhidao ba

2PL be one CL village DE 2SG should know BA

'You guys come from the village. You know that, right?' (sentence obtained from the BCC corpus, http://bcc.blcu.edu.cn/ , a 15-billion-word corpus)

What is intriguing is that some if not the majority of scholars maintain that *ba* renders its host utterance less confident (e.g. Ljungqvist, 2010; Sun, 2006; Tantucci, 2017; Zhao & Sun, 2015), Tantucci (2017) tells us that this SFP does occur in utterances that sound rather confident. One of his examples is:

(8) wo shuo, ni-zuotian-wanshang-lai-guo? bu-keneng-ba, zuotian-wanshang-xia-name da-yu, ni-zenme-hui-lai? ni-jiushi-meilai-ba.

I say, you-yesterday-evening-come-EXP? Not-possible-BA, yesterday-evening-down-so-big-rain, you-how-possibly-come? You-definitely-did not come-BA.

'Let me tell you, did you really come back yesterday evening? This is impossible, come on, yesterday evening there was such a heavy rain, how could you possibly return?' You definitely did not come back, let's be frank! (adapted from Tantucci, 2017, p. 40)

- *ou*

According to Wu (2004), *ou* has two variants: unmarked *ou*, which "is produced with a flat, low pitch, and exhibits prosodic characteristics closer to what has been described for final particles in the literature" (p. 50), and marked *ou*, which is "produced either with a markedly high pitch, or with some kind of dynamic pitch movement, such as a rising or a falling-rising pitch contour" (p. 50). Example (9) is an instance of unmarked *ou* and Example (10) is an instance of marked *ou*.

(9) ni da guo ou.

you play EXP OU

'You played (ball in the league) OU?/!' (adapted from Wu, 2004, pp. 51-52)

(A girl produced this utterance when she realized for the first time that her boyfriend had played baseball in professional games, in an ongoing conversation about early experience of sports among a group of friends.)

(10) Na ge xiaohai hen piaoliang ou

that CL kid very beautiful OU

'That kid- is (very-) pretty OU' (adapted from Wu, 2004, p. 91)

(One conversationalist produced this utterance when he or she was part of an ongoing conversation about a mutual friend's kid. At that time, both the friend and the kid were absent.)

Wu (2004) maintains that the two function of unmarked *ou* are "confirmation-eliciting and news-indexing" (p. 53). For marked *ou*, which is further pursued in Wu (2005), when it is used in the first position of a sequence in a conversation, it is "commonly used to register a heightened sense of newsworthiness of the event being reported", and in such cases final *ou* usually "turns on the juxtaposition of a background situation" (p. 983). In a few cases where information of a background situation is not present in conversations, final *ou* can also be used to emphasize or intensify some points, such as in commercial ads. When it is used in the responsive position in a conversation, it occurs in utterances "to accomplish potentially negatively-valenced interactional work", such as to disagree, to decline a request, to deny a proposition, and to issue a

warning (p. 991).

Mixed results have arisen when scholars discuss the speaker-hearer relationship of the practice of using *ou*. Li and Thompson (1981) believes that *ou* often occurs when an adult speaks to a child. Their example is:

(11) wo yao da ni ou

1.SG will hit you OU

'Let me tell you, (if you do this,) I will hit you' (adapted, p. 312)

However, both studies conducted by Wu (2004) and Wu (2005) are based on conversation excerpts among adults.

- *ma*

Ma occurs at the end of interrogative sentences. Such sentence can be a sincere[4] question, as is shown in Example (12), or a rhetorical question, as is shown in Example (13).

(12) ni chi riben cai ma?

2SG eat Japan food MA

'Do you eat Japanese food?' (adapted from Sun, 2006, p. 76)

(13) Ni zheme ying-le guangcai ma?

2sg this.way win-LE.PFV[5] look-good MA

'Do you think you look good to win this way?' (adapted from Lu, 2005, p. 33)

(The scenario provided in the original text is that a father produced this utterance to his son in their badminton play when he realized his son won by cheating. This is a rhetorical question because of course the father knew win-by-cheating was not good.)

A sincere question is one used in a neutral context. A neutral context is defined by Li and Thompson (1981) as a context "in which the questioner has no assumptions concerning the proposition that is being questioned and wishes to know whether it is true" while a nonneutral context is one "whenever the questioner brings to the speech situation an assumption about either the truth or the falsity of the proposition s/he is

[4] The word "sincere" for a question is borrowed from Lu (2005, p. 30).

[5] This particle le denotes a perfective aspect and is thus different from SFP le. They are two different morphemes, sharing the same Chinese character.

asking about, then that context is nonneutral with respect to that question" (p. 550).

The two examples have revealed a controversial issue of *ma*: its use condition. On the one hand, some scholars (e.g. Callier, 2007; Han, 1988; Sun, 2006) are convinced that interrogative utterances ending with *ma* are used in neutral contexts, as is exemplified by Example (12) and the following Example (14) adapted from Han (1988),

(14) Zhangsan shi laoshi ma?
person name copula teacher MA
'Is Zhangsan a teacher?' (p. 92)

On the other hand, Lu (2005) maintains that:

> "The use of the particle *ma* indicates the speaker suspects the statement preceding the particle not to be true. The speaker is uncertain and therefore *ma*-ending utterances have the effect of questions. Since the general characterization of the particle ma indicates that the speaker has a supposition, an utterance ending with the particle *ma* occurs only in a nonneutral context". (p. 39)

Further complicating the picture, Li and Thompson (1981) believes that *ma* can be used in both neutral and nonneutral contexts (p. 550).

- *me*

SFP me can be used in declarative sentences and directive sentences, shown in Example (15) and Example (16) respectively.

(15)

B: Xianzai shou zhei zhong chuguochao yingxiang de ren tai duo le
now suffer this kind go:abroad:trend influence DE people too many LE
'There are far too many people being influenced by the trend to go abroad'

C: Zhe ye shi hao shi me!
this also be good matter ME
'That's something good too!' (adapted from Chappell, 1991, pp. 54-55)

(B and C were talking about "Chinese Education System". B thought a lot of Chinese people were influenced by the trend of studying overseas. C replied with a me-tagged utterance and showed his or her attitude that this was actually a good thing. C's

assumption was that many people studying abroad should be a self-evident good thing and he or she uses *me* to indicate disagreement and impatience with B in questioning the positiveness of this.)

(16) ni zou kuai-dian me

2SG walk fast-bit ME

'Can you just walk fast a little bit!' (an example from The DMC Dictionary[6])

Example (15) is abstracted from the Chappell (1991)'s seminal work on *me*. This example supports Chappell's claim that *me* can be used to convey disagreement on the part of the speaker based on a self-evident situation. This use of *me* has more emotional element in it than its another usage, where a speaker uses *me* to point out the self-evident logical cause-and-effect connection in the discourse.

The emotionally charged usage shown in Example (16) suggests me is intriguingly intertwined with extralinguistic factors. As another example in point, Callier (2007)'s corpus driven sociolinguistic study found that *me* has a significantly higher frequency in speeches made by female speakers, especially young women. He also found that *me* is also markedly frequently used by Mandarin speakers with a southern accent and Taiwanese Mandarin speakers. The author therefore concludes that "the social distribution of *me* can be explained in part by the increased influence of Taiwan and Hong Kong on worldwide Chinese-speaking communities" (p. 2). That is to say, *me* is closely related to speaker gender and birthplace.

- *ne* (e.g. Lee-Wong, 1998; Boya Li, 2006; Lin, 1984), and *a* (e.g. Li & Thompson, 1981; Sun, 2006; Wu, 2004; Zhan & Bai, 2016, p. 412), which will be elaborated in later sections as two of the three SFPs are the focus of this dissertation. *bei* (e.g. Liu et al., 2001; Zhao & Shi, 2015) actually also receives some scholarly attention, although it is not on the prototypical list provided before. It is another focused SFP in this dissertation.

Note that I have just provided you with representative examples discussed in

[6] This is the 7th digital edition of 现代汉语词典 Xiandai Hanyu Cidian 'Dictionary of Modern Chinese' (CASS, 2019), as my reference book. For convenience sake, I will use "The DMC Dictionary" to just refer to this dictionary hereafter.

previous accounts. The whole picture is not that simple. As Simpson (2014) puts it:

> "They (SFPs, addition mine) are used to communicate a range of discourse-sensitive meaning relating to speaker attitude and "emotional coloring" (Matthews and Yip 1994), force of assertion, evidentiality and clause type, along with various other semantic and pragmatic factors that are sometimes **difficult to pin down** (bolding mine)". (p. 157)

Simpson (2014) also points out that linguists have more agreement in some SFPs, such as *ba* and *ma*, than some controversial cases, such as *ne* and *a*.

From a functional perspective, SFPs belong in the broad functional category of pragmatic particles (PPs) (Xiang, 2011). One notable member of PPs is discourse particles. Aijmer (2002a) points out that for discourse particles (defined later in this section), which exist at least in English and are near functional equivalents of Mandarin SFPs, speakers can "know what a particle means and be able to use it appropriately in different contexts"; "Native speakers use discourse particles with great precision as signposts in the interaction, for example to make it easier for the hearer to understand how the different parts of the text are related" (p. 3). That is to say, as with discourse particles in English, Mandarin SFPs should not be deemed elusive, both as members of PPs. Against this backdrop, this dissertation enters the conversation.

I argue that multifunctionality indeed underlies the difficulty to describe and explain SPFs, but SFPs are not elusive. My argument is inspired by Aijmer (2002a)'s discussion of the "flexibility and multifunctionality" (p. 55) of discourse particles. Discourse particles, as well as SFPs, are specific names of members belonging to the pragmatic particle family, to which I now turn.

1.2. Pragmatic Particles

The major functions of pragmatic particles are summarized by Xiang (2011) as follows:

> "Pragmatic particles across diverse languages are broadly comparable in their ability to accomplish a complex range of interactional functions and convey

> speakers' nuanced stances vis-a`-vis the propositional content of the utterance, toward the addressee, and other elements of the interactional context. Occurring pervasively in natural discourse, pragmatic particles also play a significant role in creating and maintaining textual cohesion, highlighting discourse relationships, facilitating conversational tasks, and, on the more macro-levels, indexing sociocultural identities". (pp. 1377-1378)

Pragmatic particles have other various titles, as listed in Xiang (2011, p. 1377), which include "discourse markers", "discourse particles", "modal particles", "interactional particles", "utterance-particles", and "sentence particles". The list is actually longer. I found from the literature that pragmatic particles were also called "sentential particles" and "final particles." In the following, I will present a brief summary of relevant studies under each of the names given to pragmatic particles, based on literature available to me, with the hope to provide a synopsis of their structural and functional properties.

- Pragmatic Particle

The case study of pragmatic particles offered by Xiang (2011) is SFP *lāh* in Shishan dialect in China. The study shows that the particle has a notion of restrictivity. This notion gives rise to several pragmatic extensions including "marking suggestions necessitated by external circumstances, assertion of 'obviousness,' negative politeness strategies, and various emotive stances toward the situation in focus and/or toward the addressee" (p. 1377). The data comes from transcription of naturally occurring conversations collected by the author.

- Final Particle

Haselow (2012) provides an overview of final particles in English with five case studies: then, though, anyway, actually, and even. His data comes from the British component of the International Corpus of English (ICE-GB). The function of final particles in English is summarized by the author as "to post-modify a pair of utterances as being linked to each other in a specific semantic way after their production in unprepared speech, without previous planning or anticipation of this relation" (p. 202).

The author uses detailed analyses to show that all the five final particles can be used to negotiate common ground (CG) and express (inter)subjective meanings in real-time spoken communications. In addition, the author believes that "final particles are economic devices for the service of the interactive, dynamic character of negotiating particular states of affairs in real-time interactions" (p. 203).

Sylvie Hancil, Margje Post and Alexander Haselow (2015) is a book volume that contains a typological study of what they call "final particles (FPs)" in sixteen languages. They point out that a study of FPs in Asian languages is particularly difficult because of "the functional diversity of individual particles, and the complex interplay of clause-type, pitch and particle meaning" (p. 15). This means that an Asian FP has multiple functions and its meaning is influenced by what kind of sentence it occurs in and the intonation of that sentence.

Sato (2017) discusses features of "I think" in English as a final particle (FP) in English and compares it with two Japanese SFPs: *yo* and *ne*. The data comes from "the spoken language section of the Corpus of Contemporary American English (COCA) (Davies, 2008), the Michigan Corpus of Academic Spoken English (MICASE) (Simpson et al., 2002), and the Santa Barbara Corpus of Spoken American English (SBCSAE) (Du Bois et al., 2000--2005)" (p. 85). The findings are "I think" in the right periphery (RP) has both subjective and intersubjective meaning components. "I think" as an FP shares common features with the two Japanese SFPs.

- Utterance Final Particle

Li (2013) used the term Utterance Final Particles (UFPs) in his study of *a* and *ne*. He thinks that host utterances for a suggest "surprise arising from lack of background information" and such utterances "may trigger responses from the listener who may provide more information" (p. 165). Host utterances for ne reveal "a gap between what the speaker expects and what s/he actually sees" and such utterances "may be followed by the speaker's own remarks of reflection on their own subject knowledge" (p. 165). His phonetic perception analysis compared utterances with the two target UFPs and those without. The finding shows that "only when UFPs are not used, prosodic features

will be promoted to take over the responsibility of a mood/meaning messenger" (p. 165).

- Discourse Marker

Fraser (1999)'s discussion of discourse markers (DMs) has a thorough review of previous accounts. The author proposes a strict definition of DMs in that he thinks a necessary property of DMs is to relate two separate messages (p. 940) and thus segment-initial expressions such as *frankly, obviously, stupidly* and interjections are excluded because these expressions do not relate two separate messages (pp. 942-943). The author divides DMs into two classes: discourse markers which relate messages and discourse markers which relate topics. The first class has four subclasses, which include:

a. Contrastive markers. e.g. *(al)though, but, contrary to this/that*

b. Elaborative markers. e.g. *above all, also, analogously*

c. Inferential markers. e.g. *accordingly, all things considered, as a (logical) consequence/conclusion*

d. Other infrequently used DMs. e.g. *after all, because, for this/that reason, since*

The second class are exemplified by *back to my original point, before I forget, by the way*, etc.

Based on data obtained from the spoken component of the International Corpus of English (ICE-GB), Haselow (2011) discusses two functions of utterance-final *then* in spoken English. The first function is utterance-final *then* used as a discourse marker (DM), where "retrospective marking of the relation between two propositional units" is created (p. 3608). In such case, the first propositional unit is the protasis and the second unit the apodosis. For example,

B: uhm I'll have a vodka and lemonade if you've got any

Z: well have vodka and lime *then* [ICE-GB s1a-047]

The second function is *then* used as a modal particle, where *then* links the utterance in which it appears to a non-verbalized "pragmatic pretext" outside the discourse. For example,

A: oh he's fairly happy (.)

uhm (.)

B: why do (-) why do you think he doesn't write *then*

does he not have the time [ICE-GB s1a-015]

Heine (2013) provides a list exemplar DMs in English: *after all, anyway, as it were, besides, however, indeed, in fact, instead, I mean, now, okay, so, then, I think, well, what else, you know, you see* (p. 1208). She summarizes that DMs have three functions: textual (or text planning) functions, interactional ("intersubjective") functions, and attitudinal (or "subjective") functions (p. 1239).

- Discourse Particle

Aijmer provides a list of what he calls "discourse particles (DPs)" in English: *actually, ah, all right, anyway, God, goodness, gosh, I mean, I see, ...* (Aijmer, 2002a, p. 2) His description of the functions of DPs is as follows:

> "An important property of discourse particles is their flexibility and multifunctionality. It follows that they need to be analysed from many different perspectives. They indexically point to the speaker's epistemic attitude to the utterance and affective attitude to the hearer as well as to the preceding and following discourse. This flexibility explains their enormous usefulness and frequency in discourse: they are used to grease the relations between speakers, to create coherence, to avoid conversational 'bumps', simplify on-line planning or simply to fill a pause". (p. 55)

As a case study, Lam (2009) investigates the use of *well* used as a discourse particle in the English-speaking community in Hong Kong. He compares two sources of data: data from the Hong Kong Corpus of Spoken English (HKCSE thereafter) and data a textbook database. He identifies three function domains of *well* as a discourse particle: textual functions where *well* is used as a frame "to signal transitions in topic and discourse stage" and a link to introduce additional information to the discourse (p. 270), interpersonal functions where *well* is used in responsive tokens[7] and in expressions of

[7] "Token" vs "type" are two linguistic concepts. Here is a definition: "An instance of a unit, as distinct from the unit that is instanced. E.g. in *fluffy* there are three tokens of the letter 'f' and one each of 'l', 'u', and 'y'. Distinguished from type: thus in the same word six successive letter tokens are instances of four types ('f', 'l', 'u', and 'y')." (https://www.oxfordreference.com/view/10.1093/acref/9780199675128.001.0001/acref-9780199675128-e-3434?rskey=Fhl6el&result=1 , 09/28/2020.) Loosely speaking, a token is one instance of a type.

feelings (pp. 270-271), and interactional functions where *well* is "used for processing purposes and turn management" (p. 271-272). Importantly, the author explains the functions of discourse particles in the following way:

> "As items which occur frequently in speech, discourse particles are characteristic of the spoken language. Their importance in everyday talk can hardly be overstated. Discourse particles facilitate the processes of interpretation and social involvement in spoken interaction (Watts 1988). They are essential to the maintenance of conversational cooperation (Leech and Svartvik 2002). Acting as a discourse lubricant, they ensure interactions go on smoothly". (p. 261)

- Pragmatic Marker

Based on a database of five corpora, Norrick (2009) comes up with his list the most frequently used what he calls "pragmatic marker" in English: *yeah, oh, and, well, okay, so, but, mhm, y'know, mm, um, uh, (be)cause, I mean, like, huh, or, hey, hm, uh-huh, wow, ah, ooh, anyway* (p. 867). Norrick's study focuses on interjections as pragmatic markers. He summarizes that "interjections are often grouped with exclamatives as items which signal both surprise and either positive or negative emotional involvement" (p. 867).

- Interactional Particle

Morita (2015) discusses two what she calls "interactional particles" (IPs) in Japanese: *ne* and *yo*. She argues that the two IPs are just one kind of resources Japanese speakers can deploy to build stance *in situ*. Her evidence comes from two detailed analyses of transcript excerpts in which the two IPs occur. Her theory thus enriches previous accounts that, functionally speaking, *ne* is a discrete marker of shared information or agreement and *yo* is a discrete marker of strong assertion.

- Modal Particle

In her study of Mandarin Chinese, Chappell (1991) describes two different but related uses of *me*, a modal particle as she calls. The first use is that *me* can be used by the speaker to point out the self-evident logical cause-and-effect connection. The second use is that *me* can be used to convey disagreement on the part of the speaker

based on a self-evident situation and this use of *me* has more emotional element in it.

- Utterance Particle

Tang (2015) used the term "utterance particle" (UP) for SFPs by adopting the generative approach. He proposes a generalized syntactic schema in which UPs are internal conjuncts in a phrase headed by a functional category F. He shows that this schema applies to both predicative and non-predicative utterance particles in Chinese. Being "predicative" means UPs share the status of a predicate in a syntactical sense.

***

The brief literature review of studies on pragmatic particles suggest that they are called "particles" because they are not "wordy words" like nouns, verb, or adjectives. Consider what has been mentioned as PPs in English: *yeah, oh, and, well, okay, so, but, mhm, y'know, mm, um, uh, (be)cause, I mean, like, huh, or, hey, hm, uh-huh, wow, ah, ooh, anyway.* In terms of "wordhood", such English PPs are nothing like the exemplar lexical words such as "chair" and "table" which have referential meanings, nor are they similar to the exemplar grammatical words such as "of" and "by", which have relational meanings as part of larger phrasal structures. For Mandarin SFPs, Goddard (2005) called them "little words" (p. 144). These little words can only be used in utterances and they have no referential meanings, nor relational meanings. Their meanings are exclusively at the utterance and discourse level.

It is still difficult to define "pragmatic particles" on the basis of the previously reviewed literature, because this notion has a wide coverage and has been assigned a myriad of titles. Therefore, in order to see the core properties of pragmatic particles more clearly, I used a corpus linguistic technique, i.e. keyword and term extraction, by a mini-corpus I constructed via a corpus software called Sketch Engine (Kilgarriff & Rychlý, 2003). This corpus software generated a keyword list and a term list.

The mini-corpus consists of 34 English-written papers or book chapters on pragmatic particles, either case studies or overviews, which were accessible to me up to the date February 4th, 2021. A list of all the studies compiled in this corpus is available in Appendix B. All files were in editable pdf. format before I uploaded them to the

sketchengine.eu website to build my own corpus. For all book-length treatments in this corpus, I only compiled the introductory chapters. The corpus has 360,068 words.

Keywords are "words (single-token items), that appear more frequently in the focus corpus than in the reference corpus" and "they can be used to define or understand the main topic of the corpus"[8]. Therefore, the keywords can provide information of what are the main topic(s) of pragmatic particle studies. The focus corpus in this case is my own mini-corpus and the reference corpus is The English Web Corpus (enTenTen)[9] by default.

Table 1: Keywords extracted from my mini-corpus of pragmatic particle studies and their statistics

Keyword	Frequency (focus)	Frequency (reference)	Relative frequency (focus)	Relative frequency (reference)	Score
utterance	1190	25266	2471.5	1.6	936.8
grammaticalization	278	638	577.4	0	555.4
sentence-final	255	108	529.6	0	526.9
hearer	380	10802	789.2	0.7	464.6
pragmatics	327	7261	679.1	0.5	462.3
lāh	215	4	446.5	0	447.4
interrogative	237	2315	492.2	0.2	428.8
illocutionary	208	368	432	0	422.9
interactional	214	3237	444.4	0.2	368.1
interjection	218	3685	452.8	0.2	366.2

The keyword list contains 1,000 keywords. Limited by space, only the top 10 most frequently used words on the list are discussed, as is shown in Table 1. The results provided in this table are calculated by the "simple maths" method, the default method adopted by the Sketch Engine[10]. The first column presents the keywords. The second and the third column present the raw frequency of a keyword in the focus corpus and the reference corpus respectively. The fourth and the fifth column provides the relative frequencies respectively. A relative frequency is a normalized frequency, which is the frequency of that keyword per million words in either corpus. The last column has the

[8] Lexical Computing Limited. (n.d.) *Keyword and term extraction*. Sketch Engine. Retrieved February 5th, 2021 from https://www.sketchengine.eu/quick-start-guide/keywords-and-terms-lesson/

[9] Detailed information of this website can be accessed on this website: https://www.sketchengine.eu/ententen-english-corpus/

[10] Lexical Computing Limited. (n.d.) *Simple Maths*. Sketch Engine. Retrieved February 5th, 2021 from https://www.sketchengine.eu/documentation/simple-maths/

keyness score of each keyword. The keyness score is the quotient of the sum of relative frequency of focus corpus and the constant 1, i.e. $f_{focus} + 1$, divided by the sum of relative frequency of reference corpus and the constant 1, i.e. $f_{reference} + 1$.

I looked up the Concise Oxford Dictionary of Linguistics (3 ed.) (Matthews, 2014) and other sources to find definitions for the ten keywords. The Concise Oxford Dictionary of Linguistics is my primary source book. The following definitions are therein, unless otherwise clarified:

- "Utterance" is defined as:

  Anything spoken on a specific occasion. Often distinguished from a sentence: e.g. the words 'Come here!', spoken by a specific speaker at a specific time, form an utterance which would be an instance of a sentence *Come here!*

- "grammaticalization" is defined as

  The process by which, in the history of a language, a unit with lexical meaning changes into one with grammatical meaning. E.g., in Italian *ho mangiato* 'I-have eaten', a form that was in Latin a full verb ('to have, possess') has been grammaticalized as an auxiliary (*ho*). In *mangerò* 'I will eat', the same form, first combined as an auxiliary with an infinitive (lit. 'to-eat I-have'), has further changed to an inflectional ending (*-ò*).

- "hearer" is defined as "strictly, anyone who hears an utterance, whether addressed to them (as an addressee) or not".

- "pragmatics" is defined as

A branch of linguistics conceived as dealing, separately from others, with the meanings that a sentence has in a particular context in which it is uttered. Distinguished in that spirit from semantics, conceived as studying meaning independently of contexts. E.g. *There's a car coming* would have the meaning, out of context, of a statement that a car is coming. But on a specific occasion it might be a warning to a pedestrian not to step onto a road, an expression of hope that people invited to a dinner are at last arriving, and so on. Hence, in particular, pragmatics includes the study of implicatures as opposed to 'literal meanings' or

truth conditions of sentences.

- "interrogative" is defined as

(Construction etc.) whose primary role is in asking questions: e.g. that of *Is he here?* as distinguished from the declarative (1) *He is here*. An interrogative particle or interrogative inflection is one which marks an interrogative: e.g. in the equivalent sentence in Latin, a clitic *-ne* (*adestne?* 'be-present-3sg=interrog'). An interrogative pronoun, adverb, etc. is one that represents a focus of questioning: e.g. *who* or *what* in *Who did this?, What have they done?*; *where* in *Where are they taking us?*

The terms 'question' and 'interrogative' are often interchanged. E.g. *Is he here?* is a question or is an interrogative sentence or 'an interrogative'. But a distinction can and sometimes must be made. Thus *Can't you shut up?* has the construction of an interrogative, but its usual role would not be as a question but as a request or order.

- "illocutionary" is defined as

Applied in the theory of speech acts to the force that an expression of some specific form will have when it is uttered. E.g. a speaker might stop someone and say 'Please, can you help me?' By virtue of its form (interrogative preceded by *please*) this would have the illocutionary force of a request for assistance.

Compare locutionary; perlocutionary. In the theory developed by Austin and his successors, the simple act of uttering this sentence is a locutionary act; the illocutionary act is that of uttering it as a request; the perlocutionary act is what is accomplished by uttering it (e.g. the addressee might ignore the request, or might in fact help). But what is 'illocutionary' and what is 'perlocutionary' plainly depends on how much is judged to flow conventionally from the form of an utterance. E.g. if the chairman of a meeting says 'This meeting is now closed', this may be seen as a formula which has the illocutionary force of closing it. But its form is more generally that of a statement, and, as made by the chairman, its effect might instead be claimed as perlocutionary.

- "Interactional" is define as "capable of acting on or influencing each other"[11].

[11] Vocabulary.com (n.d.). *interactional*. https://www.vocabulary.com/dictionary/interactional

- The remaining three keywords do not need a definition. “lāh” is a specific SFP of a Chinese dialect, hence a pragmatic particle. It was discussed as a case study in the literature. “sentence-final” appears on the keyword list because my mini-corpus contains many studies on SFPs, as the focus of this dissertation. “interjection” is a category of words that is believed to belong to pragmatic particles (Norrick, 2009).

Then I will explain the significance of each keyword’s occurrence on the list, based on the literature compiled in the corpus. The keyword “utterance” indicates that PPs are mainly used in spoken communication. All 34 studies are based on corpus data of spoken language. Haselow (2012), in his study final particles in English, maintains that “final particles are characteristic for the oral medium” (p. 183).

The keyword “grammaticalization” is related to studies that argue PPs are grammaticalized from other lexical categories, such as verb, adverbs or conjunctions. One notable case of grammaticalization of SFPs is *le*, which is the grammaticalization of a verb *liao* meaning “finish” (Dong, 2014, pp. 121-122). The verb “developed an aspectual use around the Tang Dynasty (618-907)” (Fang, 2018, p. 590). The topic of grammaticalization of PPs will not be elaborated on in this dissertation because it is not my focus here. Readers can start from Hancil et al. (2015) to get a sense if interested in further pursue.

The keyword “hearer” is on the list because the literature shows that PPs are “hearer-oriented”, in the words of Sato (2017, p. 85). PPs provide a hearer with information on how to process an utterance. For example, Haselow (2012)’s corpus-based study of five PPs in English, i.e. *then*, *though*, *anyway*, *actually*, and *even*, shows that they are hearer-oriented as speakers produce them when establishing a common ground among conversationalists. He maintains that:

> “They (the five English PPs, addition mine) serve as instructions for the hearer as to the role of the utterance they accompany in terms of the common ground shared by the participants, and indicate subjective and intersubjective meaning components, such as degree of certainty or self- and other-correction”. (p. 182)

The keyword "pragmatics" refers to a huge discipline in linguistics. It is impossible for me to fully explain this term[12]. However, the brief definition offered by the Concise Oxford Dictionary of Linguistics (3 ed.), presented earlier in this section, has already pointed out several key aspects, which can support my discussion of the pragmatics of PPs. The definition explains that pragmatics concerns an utterance's meaning in contexts. Contextual information of a PP's use has been a necessary part in all exemplar utterances cited and elaborated in my mini-corpus's 34 studies. This amounts to say that to understand a PP, one needs to know the context. In the 34 studies, such contextual information has to include pretext and/or post-text of an occurrence of a PP, so that enough can be known of the role played by that PP in that conversation.

The keyword "interrogative" appears on the list because studies compiled in my mini-corpus have discussed a lot about PPs used in interrogatives. I think the reason is that interrogatives are basically questions which involve at least one hearer (except for self-directed questions, but I think such kind of questions are not the norm). Since PPs are Hearer-oriented, as discussed earlier, they are legitimate resources for speakers who ask questions in conversations to deliver extra pragmatic imports to hearers, on top of interrogative messages as propositional contents. For example, Heine discussed the example of *what else* "as an elliptical interrogative clause that underwent grammaticalization and acquired subjective and intersubjective meanings" (p. 1224), as in "*Of course, on Monday nights they settle down to watch "Murphy Brown" – what else*" (p. 1225). Haselow (2011) shows that utterance-final *then*, as a pragmatic marker in English, can strengthen "the expectation of the speaker for the addressee to answer the question" (p. 3610) in interrogatives. Haselow provides the following example:

139 A: oh he's fairly happy (.)

140 A: uhm (.)

141 B: why do you think he doesn't write ***then***? [ICE-GB s1a-015] (obtained from Haselow (2011, p. 3610))

In a study of Mandarin, Ljungqvist (2010) argues that SFP *ba* as a pragmatic

[12] Huang (2016) is an excellent textbook that I recommend if you are interested in reading more about this discipline.

marker can indicate “hearer-oriented weak commitment” (p. 261) towards the proposition expressed in interrogatives, which is exemplified as follows:

(17) 他 是 学生 吧？

ta shi xuesheng ba?

he be student BA

'He is a student, right?/I guess he is a student’ (adapted from Ljungqvist (2010, p. 262))

The keyword “illocutionary” (a topic discussed in pragmatics) will be discussed with the term “illocutionary force” later in this section.

The keyword “interactional” appears on the list because PPs generally occur in interactional and dialogic settings. As Aijmer (2002a) reported in his study discourse particles in English:

> “Stenström (1990:149f) found, not surprisingly, that nearly all categories of discourse particles were less frequent in monologue (narrative) than in dialogue (in the London-Lund Corpus) and that the categories found in the monologue were restricted to the area of planning and organization. But discourse particles with an organizing function were also less frequent in monologue than in dialogue, an obvious reason being that the monologue had been preplanned (Stenström 1990:152)”. (p. 34)

Also, Haselow (2012)’s study of final particles in English (*then, though, anyway, actually,* and *even* are five cases investigated in the study) is based on corpus data of “online production of speech in interaction with at least one other participant” (p. 183), excluding the parts for preplanned speeches and monologues in his corpus. Sato shows that English utterance-final particle *I think* “implicitly invites a responsive act from the recipient and reminds the recipient of the reciprocal nature of interaction” (p. 94), which is a functional parallel with Japanese SFP *yo* and *ne*.

For a better understanding, I also used the same mini-corpus to generate a term list. A term is similar to a keyword, but a term is a multi-word unit, instead of a single word as a keyword. A term can serve the similar function as a keyword in that a term list can

show what are the typical multi-word expressions in a corpus, which indicates main topics therein[13]. Unfortunately, up to the date February 9th, 2021, I still could not solve an unexpected technological issue which prevented me from downloading the whole term list in .xlsx or .csv format (while I could somehow download the keyword list). Instead, I used the copy-and-paste function of the webpage and copied the top 50 terms on the list.

The complete list of 50 terms can be accessed in Appendix B. Only the top 10 is shown in Table 2.

Table 2: Top 10 terms extracted from my own mini-corpus of pragmatic particle studies

Ranking	Term
1	modal particle
2	use subject
3	final particle
4	illocutionary force
5	final ou
6	final position
7	speech act
8	propositional content
9	discourse particle
10	english discourse

Again, the Concise Oxford Dictionary of Linguistics is my primary source book for definitions of the 10 terms. Other sources will be clarified when cited.

- "Illocutionary force" is related to the keyword "illocutionary", as defined earlier. Here I provide the definition of the linguistic term "force" I found:

  Sometimes synonymous with 'meaning'. Usually distinguished, however, in

[13] Lexical Computing Limited. (n.d.) *Keyword and term extraction*. Sketch Engine. Retrieved February 5th, 2021 from https://www.sketchengine.eu/quick-start-guide/keywords-and-terms-lesson/

theories of pragmatics as the meaning an utterance will have, in practice, in a particular context, as opposed to a meaning it might be assigned independently of any context. E.g. *You had better leave now* has the form of a statement; therefore, it would be said to have the meaning, or a semantic representation of, a statement. But its force, in a likely context, would be not as a statement, but as an order or request to someone to leave.

- "Speech act" is defined as

  An utterance conceived as an act by which the speaker does something. Originally of performatives: e.g. by saying 'I name this ship the Queen Elizabeth' a speaker will, in the appropriate circumstances, perform the act of naming it. Thence of utterances generally. E.g. in saying 'I will be there tomorrow' one makes a promise or a prediction: i.e. one performs an act of promising or predicting. If one says 'Stephen is my brother' the act is that of making a statement, if 'Is Stephen your brother?' that of asking a question, and so on.

- "Propositional content" is defined as

  (A proposition is, addition mine) Whatever is seen as expressed by a sentence which makes a statement. Hence, for example, the same proposition might be said to be expressed by both *I understand French* and, in Italian, *Capisco il francese*. It is a property of propositions that they have truth values. Thus this proposition would have the value 'true' if the speaker did understand French and the value 'false' if the speaker did not.

  The 'propositional content' of a sentence is a part of its meaning seen, in some accounts, as reducible to a proposition. E.g. *The porters had shut the gates*, *The gates had been shut by the porters, Had the porters shut the gates?*, *If only the porters had shut the gates!* might be said to have the same propositional content, though in other respects their meanings differ.

- "English particle" appears on the list because my literature shows that pragmatic particles in English have been extensively researched. For example, Aijmer

(2002a), Norrick (2009), Fraser (1999, 2015), Haselow (2012), Heine (2013) are comprehensive studies. Lam (2009), Haselow (2011), Schourup (2011), Sato (2017) are case studies.

- The remaining terms do not need a definition. “modal particle”, “final particle”, “discourse particle” are titles given to pragmatic particles by previous studies. “use subject” seems a technical failure by the software because this term does not make sense. “final position” regards pragmatic particles in the final position of an utterance. It appears on the term list because my literature review revolves around SFPs. Therefore “final *ou*” is also on the list, as a case study of SFPs, which is extensively discussed in Wu (2004).

The term “illocutionary force”, “speech act”, and “propositional content” are related to the Speech Act Theory in pragmatics. According to Huang (2014), the theory is attributed to British philosopher J. L. Austin (p. 118), and

> “The (capitalization mine) central tenet of speech act theory is that the uttering of a sentence is, or is part of, an action within the framework of social institutions and conventions. Put in slogan form, saying is (part of) doing, or words are (part of) deeds”. (p. 119)

The term “illocutionary force” is otherwise known as “illocutionary act”, which refers to “the type of function the speaker intends to fulfil or the type of action the speaker intends to accomplish in the course of producing an utterance” (Huang, 2014, p. 128). It’s one of the three facets of a speech act distinguished by Austin, the other two being locutionary force/act and perlocutionary force/act (Huang, 2014, 4.3.), and the concept of speech act in a narrow sense just refers to the illocutionary force/act (p. 128).

Apart from the general definition offered by the Concise Oxford Dictionary of Linguistics, a definition in pragmatics of the term “propositional content” is offered by linguist John Searle as a notion “that has to do with specifying the restrictions on the content of what remains as the ‘core’ of the utterance (i.e. Searle’s propositional act) after the illocutionary act part is removed” (Huang, 2014, p. 131). For example, the

propositional content of a promise is "to predicate some future act of the speaker"; the propositional content of a request is "to predicate some future act of the addressee" (Huang, 2014, p. 131).

As for the propositional content of PPs, my mini-corpus reveals a consensus reached by scholars that PPs do not have propositional contents. For example, a typical property of pragmatic particles is believed to be to "resist clear lexical specification and be propositionally empty (i.e., it would not be part of the propositional content of the sentence)" (Östman, 1982, cited from Aijmer, 2002a, p. 28). Hancil et al (2015) maintain that final particles (FPs) "have no propositional content and do not effect the truth conditions of the unit they accompany" (p. 5). Haselow (2012) thinks that discourse particles in English "do not contribute propositional information, i.e. information which alters or specifies the propositional content of an utterance" (p. 185).

As an example regarding PPs in terms of the three terms, SFP ba-tagged sentences in Mandarin performed the following speech acts,

a. Given a declarative sentence S, S-ba means: I tell you with uncertainty that S.

b. Given an interrogative sentence S, S-ba means: I request from you a direct answer to the question S.

c. Given an imperative sentence S, S-ba means: I advise you that (you) S. (Deng, 2015, p. 38)

Here "S" stands for "sentence", and "S-ba" stands for a *ba*-tagged sentence. Note, for the three sentence types, namely declarative, interrogative, and imperative, a *ba*-tagged sentence type can not only express the core meaning of that sentence type, but can express specific speaker attitudes. These attitudes, as Deng (2015) further explains, rest in the corresponding illocutionary forces/acts of the three *ba*-tagged sentence types:

a. telling the hearer some information with uncertainty when attached to a declarative (the information is provided by the declarative);

b. requesting the hearer to provide a direct answer to a question when attached to an interrogative (the question is provided by the interrogative);

c. advising the hearer with a piece of advice when attached to an imperative (the

advice is provided by the imperative) (p. 39)

Note further that SFP *ba* does not alter or contribute to the propositional content of each sentence type. They index speaker attitudes. This observation of the relationship between PPs and utterance propositional contents echoes Chao (1926)'s observation in his study of Mandarin SFPs that, 说话有所说的内容，说话有说话的口气 'There are contents in people's words. There are also attitudes in people's words' (p. 133). He thinks that SFPs are one of the means to express people's attitudes in Chinese.

To sum up at this point, those 10 keywords and 10 terms showing up on the lists means that they occur significantly more frequently in my mini-corpus, which consists of previous accounts of pragmatic particles, than they do in the general corpus, i.e. The English Web Corpus (enTenTen). In other words, those keywords and terms are the main aspects from which previous scholars have discussed PPs. From these aspects, I summarize the inherent properties of PPs as follows:

- They predominantly occur in dialogues[14] in spoken language.
- They do not change or contribute to the propositional content of an utterance they accompany.
- They accompany utterances that have illocutionary forces/acts in particular contexts. Stated in another way, such utterances are produced by speakers to do something for a certain interpersonal effect in specific conversations.
- They are speaker-initiated (as Deng (2015) coins it) and they accompany utterances that are hearer-oriented.

As one member of the PP family, Mandarin SFPs arguably also possess those properties. That is to say, SFP-accompanying utterances mainly occur in dialogic exchanges. Such utterances have illocutionary forces/acts, which are deployed by speakers and geared towards hearers. As is echoed in Han (1988):

> "Particles (the term used by the author for SFPs, addition mine) are a class of morphemes in Mandarin Chinese which, unlike the majority of the linguistic items in Mandarin, do not possess any distinct lexical meaning, yet carry a very

[14] "Dialogue" and "conversation" are used interchangeably in this dissertation unless otherwise clarified.

important pragmatic load (such as indicating the attitude of the speaker towards the context of his/her utterance). Their behaviour in linguistic communication may provide a rich source for our understanding of human communicative behaviour generally". (p. 1)

The big picture set, the next section offers a brief typological discussion of SFPs.

1.3. SFPs as an Areal Feature of Several Asian Languages

SFPs are one member of what Hancil et al. (2015) term "final particles (FPs)" in world languages because SFPs share both structural and functional commonalities with FPs in other languages. They present us with FP examples in several languages (p. 4), many of which are genetically distant (FPs are in bold):

a. English: I wouldn't care **actually/anyway/but/even/so/then/though**.

b. Dutch: Die avond moest 't gebeuren **dus/immers/maar/misschien**.

'That evening it had to happen thus/after all/but/perhaps.'

c. Northern Russian

Ona davno ne robotat. Bol'na **dak**.

she:NOM;SG long NEG work:PRS;3SG ill:NOM;F;SG PRT

'She hasn't been working for a long time. (Because) she is ill.'

d. Cantonese[15]

Neih sik keuih **maa.3/me.5**[16]

you know him/her PRT

'Do you know her PRT (neutral question)/ PRT (surprise, dismay)?'

e. Venetican Italian

Dove valo, **ti**?

Where goes.he PRT

'Where on earth is he going?'

[15] Clearly here Hancil and his colleagues regard Cantonese as a language, instead of a regional variety of Chinese, as upheld by other linguists (e.g. Sun (2006)). I see Cantonese as a regional dialect in this dissertation.

[16] As explained by Chan (2002), these numbers represent tones of SFPs in Cantonese, with 5 being the highest pitch and 1 being the lowest (p. 59).

Hancil et al (2015)'s examples show that FPs occur at the end of utterances. In addition, FPs are "non-inflecting and usually monomorphemic" (p. 9), which means FPs often do not occur with affixes and FPs are made by single morphemes[17]. Mandarin SFPs are also monomorphemic in most cases, as has been shown by the examples in the Overview section.

Like pragmatic particles in all other languages, Mandarin SFPs occur in conversations (e.g. Chappell, 1990; Goddard, 2005; Jing-Schmidt, 2019; Simpson, 2014). SFPs are rare in scientific texts (Alleton, 1981, p. 95). Nevertheless, Mandarin SFPs are special in that they are noticeably connected to previous sounds in natural speeches. As depicted by Simpson (2014):

> "It has been noted that such elements (SFPs, addition mine) are frequently de-stressed and neutral in tone (Li and Thompson 1981), but also prosodically integrated into the sentence they combine with, and are not set off from the main sentence by any pause intonation, unlike question particle "tags" in languages such as English (e.g., "John saw you, right?"; Matthews and Yip 1994)". (p. 159)

Apart from the structural commonalities, FPs also share functional commonalities. Hancil et al. (2015) point out that FPs "have little or no lexical or conceptual, but predominantly procedural meaning in terms of Blakemore (1987) in that they provide an interpretive cue to the hearer as to how to understand the sentence or utterance they accompany" (p. 4).

Similarly, in Enfield (2005)'s introduction to what the author calls "MSEA" (standing for Mainland Southeast Asia) languages, which include "Vietnam, Laos, Cambodia, and Thailand, with some extension west into Burma, south into Peninsular Malaysia, and north into southern China" (p. 182), the author points out that SFPs are one of the three clausal/sentential organization features shared by MSEA languages, the other two being verb-object order and topic-comment structure. (pp. 189-190) He particularly explains that the major function of SFPs in these languages is to express

[17] The definition offered in the Concise Oxford Dictionary of Linguistics is as follows: "A minimal unit of grammar into which a sentence or a word within a sentence can be divided. E.g. *Come inside* can be divided into the minimal units *come*, *in-*, and *-side*; *distasteful* into *dis-*, *taste*, and *-ful*".

illocutionary force (in this citation "PCL" stands for "particle"):

> "A third feature of sentential organization in MSEA is the use of sentence-final particles as a basic mode of distinguishing illocutionary force at the utterance level (Crisfield 1974, Luke 1990). A basic proposition such as Lao *man2 kin3 nam4* (3sg drink water) can be made into questions or statements of various kinds by adding one of a large set of monosyllabic final particles at the right border of the clausal core. For example, *man2 kin3 nam4 b`o`o3* (3sg drink water PCL) 'Will he drink water?'; *man2 kin3 nam4 vaa3* (3sg drink water PCL); 'Oh, he'll drink water, will he?'; *man2 kin3 nam4 dˆej2* (3sg drink water PCL) 'He'll drink water, you know'; *man2 kin3 nam4 dee4* (3sg drink water PCL) 'He'll drink water, y'hear!'" (p. 190)

As for Mandarin SFPs, Erbaugh (1985) explains that SFPs are "an areal feature of the languages of the East Asian mainland", including Chinese and Vietnamese, and SFPs are also an areal feature of other Asian languages including Lahu, Burmese, Thai, and Cambodian (p. 84). Functionally, they serve as discourse signal (p. 86) with various specific functions and they are used to elicit hearers' response (p. 85). Functionally, the author believes that SFPs are equivalent to intonation in English and other Indo-European languages (p. 85).

Erbaugh's observation of SFPs' typological traits compared with intonation is echoed in Goddard (2005) and Feng (2015). Goddard summarizes that SFPs mainly serve two functions: one is to "distinguish different kinds of speech-acts (requesting, questioning, persuading, advising, reminding, instructing, and so on)" and the other is to "express the speaker's emotional responses (surprise, doubt, impatience, reluctance, hesitation, and so on)." (p. 144) The two functions, maintained by Goddard, are likely to be achieved by intonation in English. In a similar vein, Feng notices that Indo-European languages do not have lexical tones (they have intonation) and they do not have SFPs. At the same time, languages such as Chinese (including Mandarin and other regional dialects including Cantonese and Hakka), Thai, Vietnamese, and Japanese have lexical tones and SFPs. He goes on to hypothesize that in languages such as Mandarin that lack intonation but have lexical tones, functions accomplished by

intonation in Indo-European languages are accomplished by SFPs.

Intonation "is used to signal how a speaker intends his or her utterances to be interpreted" (Ashby & Maidment, 2005, p. 154) and it is used in all languages, to the best of linguists' knowledge (p. 165). A clearer description of its role in Mandarin Chinese is offered by Lin (2007):

> "SC (the author uses "SC" to refer to Standard Chinese, another term for Mandarin Chinese) still has intonation, but with the use of particles (the author is referring to SFPs here) the change of pitch contour for intonation purposes is minimized to some extent". (p. 229)

Lin (2007) also provides an explanation of the functions of SFPs:

> "To avoid the potential conflict between tone and intonation, SC, like many other tone languages (especially Asian tone languages), makes use of sentence-final PARTICLES (capitalization the author's) to indicate certain groups of syntactic and contextual meanings expressed by intonation in a non-tone language". (p. 228)

"The potential conflict" exist because both tone and intonation are produced by pitch variation (Ashby & Maidment, 2005, p. 166). In addition, Lin's insight on SFPs also parallels with Östman's observation (cited from Lee-Wong, 1998, p. 390):

> "In languages like Chinese, which have so many other uses for pitch variation that they need to use particles to express many phenomena that are covered by intonation in English (1981: 43 & 84)".

Let's see an example of how SFPs play the part of intonation, as discussed in Feng (2015). Feng provides us with a scenario where a professor played tennis with two of his friends for several rounds. Then the professor asked his friend whether to play a few more rounds or not. The two friends both replied "I'm good," but in different intonations. The first friend's reply had a rising pitch and he or she meant "yes" while the second friend's reply had a falling pitch and the meaning was "no". Feng explains that whereas the two replies are distinguished by intonation in English, this distinction can only be expressed by SFPs in Mandarin. The Mandarin version of the two friends' replies will respectively be:

Reply 1: 好　　哇！

hao　　wa (a phonological variant of SFP *a*[18])

good　A

'I am good. (with a rising intonation)'

Reply 2: 我 好　　啦！

wo hao　　la

I　good　LE.A

'I am good. (with a falling intonation)'

In Feng's example, *a* and *la* (the fusion of *le* and *a*) are both SFPs.

SFPs are also attested in other East Asian languages. In Japanese, for example, *no* is used "when the speaker gives, or asks for, an explanatory or clarifying comment with regard to a certain situation in the discourse context"; *koto* is used to "express exclamation" or "to give an order or direction"; *to* can only occur in the informal register to convey "defiance", "self-affirmation", or a casual declaration; *tte* is used to "report a proposition expressed by someone else" or "express the speaker's insistence on a proposition" (Okamoto, 1996). *ne* is believed to be a marker of shared information/shared feeling or a marker of agreement; *yo* as a marker of strong assertion (Morita, 2015). In Korean, for example, SFP *-nikka* can express meanings of contingency, contrast, adversative, reassertion, and emphasis (Rhee, 2012).

1.4. SFPs in Chinese

1.4.1. Previous Accounts of non-Mandarin SFPs

Cantonese also has SFPs (Simpson, 2014). Chan (2002) tells us that Cantonese SFPs appears frequently in colloquial Cantonese. Their major function is described as "reflecting the attitude or emotion of the speaker" (p. 57). She also tells us that Cantonese has a far larger inventory of SFPs than Mandarin, with Cantonese having 30

[18] For a detailed description of the phonological rules for SFP *a*, please refer to Appendix B in Han (1988). For a recent empirical treatment on this topic, which was written in Chinese, please refer to Xu (2018).

basic forms of SFPs and Mandarin only having 7 to 17 (pp. 58-59). Sybesma and Li (2007) maintain that the number of Cantonese SFPs is at least 40 (p. 1739).

Some examples of Cantonese SFP case studies include: Chan (1996) and Chan (1998) for *je* vs *jek*, Chan (2002) for *la.33* vs *la.55*, and *a.33* vs *a.55* ("33" means a mid-level tone and "55" means a high tone), Luke (1990) for *la*, *lo*, and *wo*, etc. Since this dissertation does not focus on Cantonese SFPs, I will not go into more details. I will start a brief review of Mandarin SFPs in the next section.

1.4.2. Previous Accounts of Mandarin SFPs

Reference Grammars

Li Wang's seminal book *Zhongguo Yufa Lilun* 'Theory of Chinese Grammar' is the earliest study that touches upon SFPs, which is accessible to me. This book was first published in 1940's and its 2015 reprinted version is my reference source. Wang calls SFPs as 语气词 *yuqici* 'emotional particles'[19], which express 语气 *yuqi* 'emotional moods' (p. 169). The following list includes what Wang believes as the emotional moods expressed by SFPs, with representative SFPs for each emotional moods included in brackets: determination (*le*), explanation (*de*), emphasis (*ne* and *bale*), interrogation (*ma* and *ne*), rhetorical question (*ma* and *ne*), hypothesis (*ne*), conjecture (*ba*), command (*ba*), urgency (*a*), resignation (*yeba* and *bale*), indignation (*me*), and persuasion (*ma*).

Chao (1968) is a classical study on SFPs. His study is based on a solid dataset comprising six sources: (1) Made-up examples by Chao; (2) Written records of daily conversations by Chao; (3) Disc and tape recordings of Mandarin conversations; (4) Literary texts that represent spoken Chinese at that time; (5) Literary texts that represent spoken Chinese at other periods; (6) *Baihua*, i.e. written vernacular Chinese at that time. (pp. 15-16) In this book, Section 8.5 Particles (pp. 795-814) is devoted to a lengthy list of 26 SFPs, where he lays out communicative functions for each particle. Rich in width,

[19] The translation of "emotional particles" and "emotional moods" are provided by the author Wang himself.

Chao's study of each SFP lacks depth, but his preliminary insights are referred to by almost every later study as starting point.

Lyu (1974) also provides a definition of 語氣 *yuqi* 'emotional moods', i.e. 概念的内容相同的語句，因使用的目的不同所生的分別 '(emotional moods are) nuances generated by sentences with the same proposition but used for different purposes' (p. 261). He also thinks that *yuqi* is expressed by 兼用語調與語氣詞 'the interplay between intonation and SFPs' (p. 261). A problem of his discussion is that the study is based on either introspective example sentences or sentences abstracted from literature works, which might not reflect the authentic usages of SFPs in naturally occurring conversations.

Li and Thompson (1981) is another widely cited source. The authors believe that SFPs have elusive semantic and pragmatic functions, which poses great difficulty for linguists (p. 238). "They typically occur in speech or in writings that reflect or recount conversations" (p. 238). Their case studies include:

Table 3: Li and Thompson's capturing of the six SFPs' functions

SFP	**Function**
le	Currently Relevant State
ne	Response to Expectation
ba	Solicit Agreement
ou	Friendly Warning
a/ya[20]	Reduce Forcefulness
ma	Question

A problem of Li and Thompson (1981)'s study relates to the explanatory power. For instance, when they explain the function of "Response to Expectation" of *ne*, they use the example,

(203) (This is the numbering for the example in the original text.)

a. tamen you san tiao niu

they exist three CL cattle

[20] *ya* is a phonological variant of *a*.

'They have three cattle.'

b. tamen you san tiao niu *ne*

they exist three CL cattle *NE*

'Listen, they have three cattle.' (adapted from Li and Thompson (1981, p. 301))

Then the authors go on to elaborate that:

> "For example, one context for which (203) b would be a perfect response would be one in which A has just stated that 'they' don't have any money and are very poor, and B then challenges A's claim, ending the challenge with (203) b". (p. 301)

Following this, they give us another legitimate scenario by saying that:

> "For example, another speech context for (203) b could be that A had been describing how rich these people are; B could then utter (203) b in support of A's claim. In this context, (203) b would be more appropriate if the adverb *hai* 'even, still, also' were placed in front of the verb *you* 'exist' so that the complete meaning of the sentence would be: 'You know, they even have three cattle.'" (p. 301)

They use two modal verbs, "would" and "could", to indicate that there can be countless such legitimate scenarios in which *ne* fulfills the "Response to Expectation" function. The protean nature of such imagined example sentences shows the difficulty to exactly pinpoint the *ne* semantics. In addition, the use of modal verbs in the authors' explanation indicate that the utterances are not authentic ones but constructed ones. Stated it in another way, they are not naturally occurring utterances. The same is true for all other SFPs investigated by the authors. My point is not to demonize Li and Thompson's study. I am arguing that their choice of non-naturally occurring examples limited the explanatory power of their theory because it may not be generalized to other scenarios.

Zhu (1984) maintains that SFPs can be divided into three groups (p. 208): 1) SFPs that mark aspect, including *le*, *ne*, and *laizhe*; 2) SFPs that express a tone of interrogation and directiveness, including[21] ne_2, *ma*, ba_1, and ba_2. Ne_1 and ne_2 are two variants of *ne* and ba_1 and ba_2 are two variants of *ba*. These variants have different

[21] Note here that I used subscripts for variants of these Mandarin SFPs, instead of numbers in the format of body texts, as in the case of Cantonese SFPs where normal numbers represent pitch heights.

functions. He explains that *ne*$_1$ marks aspect, *ne*$_2$ expresses interrogation, *ba*$_1$ expresses interrogation, and *ba*$_2$ expresses directiveness.

Hu (1995) believes that the function of SFPs is to 帮助语气的表达，同时它能在语调的基础上增加色彩 'facilitate the expression of emotional moods and in the meantime add colors to utterances on the basis of intonation' (p. 376). He thinks there are only four kinds of *yuqi* that SFPs can express: declarative, interrogative, directive, and exclamative (p. 376). He discusses six case studies in terms of the function of each SFP: *de* (表示确实如此 'to confirm'), *le* (表示已经如此或出现新情况 'to indicate a state of affair has happened or a new state of affair arises'), *me* (表示可疑 'to cast doubt'), *ne* (表示不容置疑 'to affirm the undoubtedness of a state of affair'), *ba* (表示半信半疑 'to cast doubt'), and *a* (增加感情色彩 'to add colors to emotions') (p. 376).

Liu, Pan, and Gu (2001) point out that sentence-final particle (SFP) in Chinese is a complicated phenomenon for CSL learners (p. 411). Major SFPs include *a*, *ma*, *ba*, and *ne*. Their common function is to soften the sentence, i.e."缓和句子语气" (p. 411). The section then lists nine SFPs and illustrates their function, syntactic distribution, and prosodic features.

Shao (2016, pp. 16-17) divides *yuqi* expressed by SFPs into four categories:

- Declarative

le: 表示一种变化的新情况的出现 'To indicate a newly arising situation'

ne: 表示提醒 'To express a warning'

laizhe: 表示刚刚发生过 'To tell that an event just happened'

zhene: 表示对事实的确认 'To confirm a fact'

- Interrogative

ne: 用在特指问、选择问和正反问句末，表示深究的语气 'To emphasize the tone of interrogatives for wh-questions, A-not-A questions[22], and yes-no questions'

[22] In Mandarin, an A-not-A question is formed by juxtaposing both the affirmative and the negative form of a word. For example,
Nimen laodongli gou-bu-gou ne
2PL work.force enough-not-enough NE
'Do you have enough work force (or not)?' (Example in the DMC dictionary)

a: 表示惊疑 'To express a speaker's surprise'

ma and *ba*: 都用在是非疑问句末，只是"吗"表示怀疑的程度比较大，疑大于信；"吧"表示怀疑的程度比较小，信大于疑 'The two SFPs are both used at the end of yes-no questions. *Ma* expresses a higher degree of lack of information while *ba* expresses more speaker's confidence in some states of affairs'

- Directive

ba: 口气比较缓和，有商量的意味 'To soften the tone and express an intention of negotiation'

a: 在肯定祈使句中有催促的意味，在否定祈使句中有强调劝阻的意味 'To express an import of urgency in affirmative directive sentences and a strong tone of discouragement in negative ones'

- Exclamative

主要是"啊"，"呀、哇、哪"都是"啊"的语音变体 'Mandarin speakers mainly use *a* to express exclamation. *Ya, wa, na* are all phonological variants of *a*'

Doctoral Dissertations

Han (1988) is the first application of pragmatics theory to the study of SFPs that produces a comprehensive and in-depth understanding. This study has both an overview of distributional patterns of major SFPs and two detailed case studies: *ba* and *le*.

The overview discussion argues against previously held claim that SFPs are sentence mood markers. Han's finding is that SFPs "do not constitute a sufficient condition for classifying sentences as" (p. 20) declaratives, imperatives, interrogatives, or exclamatives. For instance, he points out that SFP *le*, *de*, and *ne* can occur in declaratives, imperatives, and exclamatives. Since a sentence cannot be simultaneously a declarative, an imperative, and an exclamative, the three SFPs are not sentence mood marker.

For *ba*, Han maintains that it changes the neustic "I say so" in declarative and imperative sentences to "I think so", indicating the speaker is being polite. It changes the neustic "I wonder so" in interrogative sentences to "I insist so", and thus conform

the interrogatives into directives, indicating the speaker is being angry.

For *le*, the author distinguishes post-verbal *le* from SFP *le*, where the former means cessation of an event and the latter inception of an event. Importantly, as Han advocates, SFP *le* itself does not have a pragmatic function, differing from SFP *ba*. The author puts it this way:

> "What needs to be clarified is simply that the inceptive meaning of the sentence-final *le* may be utilised by the speaker in a particular situation to contradict the real world affairs to meet the speaker's need, rather than the *le* particle itself having an ironic interpretation. The above does not exclude the possibility of a sentence containing *le* being used ironically, just as an interrogative form may be used rhetorically to express a speaker's annoyance etc". (pp. 185-186)

This observation of *le* is compatible with my theoretical framework, which will be explained in Chapter 2. In a nutshell as a preview, SFPs by themselves do not fulfill functions in human communication. SFPs are part of some "chunks" that do the job. You will know what the "chunks" are later.

Han's conclusion is that a pragmatic account of SFPs is essential because SFPs have a rich pragmatic load, which means they signal nuanced meanings in contexts. A pure syntactic account can only characterize the distributional patterns, but not the motivation underlying speakers' deployment of them. This of Han's observation coincides with my mini-corpus of pragmatic particle studies in that "pragmatics" is one of keywords therein. Han's highlighting the role of pragmatics in SFPs study is in line with the literature in general.

Lu (2005) advocates that SFPs are attitude markers. The author studies the pragmatics of six SFPs, namely *ma, ba, a, ne, de* and *me*, by providing both positive data, in the form of a general use condition for each of them, and negative data, in the form of listing situations where a target SFP will not occur following the prediction of those general use conditions for SFPs.

The data was taken "from some daily conversations in which I (Lu as the author) was engaged or that I accidentally overheard, from literary writings in the China Central

Daily, and from some pieces of Xiangsheng 'Cross Talk'" (p. 2). The analysis of data is conducted in two ways: (1) comparison between minimal pairs consisting of utterances ending with SFPs and corresponding utterances ending without; (2) comparison and contrast among the six target SFPs when each of them is used in a certain scenario.

Her conclusion is that "Taking the speaker's attitude and the speaker's belief about the addressee's attitude into consideration in generalizing the characterizations of six sentence-final particles in Mandarin Chinese has proved to be an effective way to distinguish the particles. It should be instructive to extend this approach to other sentence-final particles in Mandarin Chinese, as well as those in the other Chinese dialects, and to attitude markers in languages other than Chinese" (p. 166). This of Lu's observation echoes the fact that hear-oriented is one core property of PPs in general, as revealed by my mini-corpus findings.

Lu also points out that a constructional study of SFPs will be fruitful. She argues that,

> "The correlation between a construction and a particle has been investigated in this study. The correlation between a particle and a sub-construction, such as an adverb or an adverbial phrase, needs to be investigated, for example, the frequent co-occurrence of *cai* and the particle *ne* as in *Wo cai bu yao qu ne* 'But I just don't want to go'. It may shed some more light on the general characterizations for the uses of the particles" (p. 166).

This constructional view endorses my theoretical framework, which will be laid out in Chapter 2.

Bei, *Ne*, and *A* as Case Studies in this Dissertation

This current study will focus on three SFPs, namely *bei*, *ne*, and *a*, in terms of their distributions and functions. The review of each of the three SFPs begins with the dictionary definition offered in The DMC Dictionary. The dictionary entries can help you to get a basic sense of each particle.

Table 4 is a frequency table of the three target SFPs in the Conversation subcorpus of BCC.

Table 4: Frequencies of the three target SFPs in the Conversation subcorpus of BCC (http://bcc.blcu.edu.cn/zh/cid/3)

SFP	Frequency
a	1049004
ne	470482
bei	33314

A is the most frequently used SFP among the three. *A* and *ne* are in fact the two most frequently used SFPs in Mandarin (Li, 2013, p. 146). *Bei* clearly has the smallest population in the corpus. On the other hand, although *bei* is overall much less frequently used, it serves important interpersonal functions. It receives the least scholarly attention and is poorly understood. *Ne* and *a* have long been controversial cases for linguists (Simpson, 2014), as is revealed in the following literature review for the three SFPs considered.

1.4.3. 呗 *bei*

The DMC Dictionary defines *bei* in this way:

- Entry 1: 表示事实或道理明显，很容易了解 'To suggest that a fact or truth is clear and straightforward, which is therefore easy to understand'

For example,

(18) bu dong, jiu haohaoer xue bei
NEG understand adverb[23] diligently study BEI
'If you don't understand, then just study hard.'

- Entry 2: 表示勉强同意或勉强让步的语气 'To suggest an overtone of reluctance in one's consent or concession'

For example,

(19) qu jiu qu bei

[23] I gloss *jiu* as "adverb" because it is polysemous and it lacks an equivalent in English in any sense. Its interpretation depends on the specific context.

go then go BEI

'If going, then just go.'

In prior Chinese Linguistics literature, this particle is rarely touched upon. One exception is Liu, Pan and Gu (2001), in which the authors propose that *bei* indicates the propositional content of a host utterance is obvious, and *bei* speaker is characterized by an impatience tone when using *bei* (p. 429). The following is one example extracted from their discussion:

(20) - ni zenme lai de

you how come DE

'How did you arrive here?'

- zou lai de bei zher you mei che

walk come DE BEI this.area an emphasizing particle NEG vehicle

'(Of course) on foot! This area does not have available transportation.'

When asked how to be here, the *bei* speaker stated that he or she arrived on foot because there was no other available transportation method. The speaker used *bei* to convey the sense of obviousness by the assumption that this fact should be in the common ground between the speaker and the listener. Also, the utterance implies impatience.

Lyu (1999) also agrees that *bei* indicates obviousness, and he adds that *bei* can also be tagged with the construction [verb + 就 *jiu* 'adverb' + verb] to express a sense of triviality (p. 69). Following is one of his examples:

(21) xia jiu xia bei zanmen dai zhe yuyi ne

rain:fall adverb rain:fall BEI 1PL bring IMP raincoat NE

'You should not worry about the fact that it is raining because we have raincoats!'

Zhao and Shi (2015) believe that the core meaning of *bei* is 应而不愿 *ying er bu yuan* 'should but not be willing to' (p. 75), i.e. a speaker uses *bei* when he or she doesn't feel like replying but has to do so, thus expressing negative sentiment. Such negative sentiment is also captured by Shi and Huang (2016), who argue that *bei* "typically appears at the end of a main clause in a conditional sentence, to represent the speaker's not-a-big-deal attitude toward the consequence indicated by the main clause." (p. 39)

Shi and Huang's examples are:

(22) ta yao qu jiu qu bei

he will go adverb go BEI

'If he wants to go, just let him go'

(23) mei qian le, zanmen jiu yaofan bei

NEG money LE 1PL adverb beg BEI

'If we run out of money, we can just beg'

Satisfactory as it may seem, our understanding of *bei* can be enriched from at least two angles. First, a pilot study of the corpus data of *bei* by the current author proves that *bei*'s function is far from just asserting obviousness epistemically. At the very least, some *bei* utterances were also found to fulfill the speech act of requesting in my data. Second, previous studies on *bei* speak little about the discourse management functions of *bei*.

1.4.4. 呢 *ne*

Let's first see how The DMC Dictionary defines this particle:

- Entry 1: 用在句中表示停顿（多为对举） 'Used in sentence-internal position to mark a pause (such sentences usually contrast two things)'

For example,

(24) rujin ne, ke bi wangnian qiang duo le

nowadays NE, but compare previous.years better much LE

'Nowadays, it is much better than before'

- Entry 2: 用在陈述句的末尾，表示动作或情况正在继续 'Used in the sentence-final position of declaratives to mark the progressiveness of an activity or a situation'

For example,

(25) ta zai jing bian dashui ne

she at well side fetch.water NE

'She is fetching some water at the well'

- Entry 3: 用在陈述句的末尾，表示确认事实，使对方信服（多含夸张的语气） ‘Used in the sentence-final position of declaratives to confirm a fact so as to convince other interlocutors (such sentences often convey an overtone of exaggeration)’

For example,

(26) zhe ge yao ling de hen ne, fu shang jiu bu teng le
this CL medicine effective DE much NE, to.apply up adverb NEG pain LE
‘This medicine is very effective. Once applied and you will not feel pain anymore.’

- Entry 4: 用在疑问句（特指问、选择问、正反问）的末尾，表示提醒和深究的语气 ‘Used in the sentence-final position of interrogatives (wh-questions, alternative questions, A-not-A questions) to convey an overtone of reminding and strong enquiry’

Examples for Entry 4:

(27) zhe ge daoli zai naer ne
this CL reason at where NE
‘Where does this reason lie? / How to understand this issue?’

Ne is a controversial topic in Chinese Linguistics (Simpson, 2014, pp. 160-161). As Simpson summarizes, such controversy comes from the tension “whether it is necessary to assume multiple, homophonous SFPs in certain instances, each particle having a different function, or whether a unique particle can be posited to exist with an underlying, broad meaning that may be applied in a range of different contexts” (and this tension appears in the analysis of other SFPs as well) (p. 161). Simpson’s observation is actually a paraphrase of the contrast between maximalist and the minimalist approach in the literature (Wu, 2005), where linguists adopting the maximalist approach list different senses of a single particle, on the basis of various types of utterances that a particle accompanies, while linguists adopting the minimalist approach come up with a core, central function of a single particle and advocate that other functions are derived from it.

In the literature of *ne* studies, maximalists aim to categorize *ne* utterances into several distinct groups. The DMC Dictionary entries you just saw are an example of this approach. Other works adopting this approach yield similar results, such as Chao (1968), who thinks *ne* has four functions (pp. 801-802):

1. Questions in a Context

e.g.

ta lai-bu-lai ne

he come-not-come NE

'Will he come here or not?'

2. Questions with a Specific Point

e.g.

ta hui la tiqin; ni ne

he can play violin; 2SG NE

'He can play violin. What about you?'

3. Deliberate Pause

e.g.

jianglai de wenti ne, na jiu deng dao jianglai zai shuo

future DE question NE, then adverb wait arrive future again say

'As for the question about the future, let's just wait and solve it in the future'

4. Mild Warning

e.g.

zhe dao hen weixian ne

this indeed very dangerous NE

'This is indeed dangerous.'

By contrast, minimalists try to combine *ne*'s various functions into one grand function and thus arrives at a unified account for its functions.

Some authors think the core function of *ne* is to mark some contrast between messages conveyed in a discourse. For example, Li and Thompson (1981) argues that *ne* in both declaratives and interrogatives conveys the message that the speaker's words

is based on the experience, claim, expectation, or belief on the part of the hearer. Lin (1984) argues that "the general meaning or relational invariant meaning of *ne*" is the semantic feature [+contrastiveness]; "when the particle *ne* is used, the flavor of *contrast* is clearly and fully expressed" (p. 238). Lin maintains that *ne* is used to point out the distinction between what the speaker thinks is the way with "what has been guessed, claimed, expected or believed to be a certain way" (p. 237). Wu (2005) argues that the discourse function of *ne* is to mark hearer engagement for Common Ground (CG) negotiation, when it is used in statements. The author finds that "the dominant use of NE utterances is either to support the speaker's own claim or to contradict the hearer's" (p. 76), which are all about signaling the hearer to adjust the CG of the current interaction. In a similar vein, Li (2013) maintains that host utterances for *ne* reveal "a gap between what the speaker expects and what s/he actually sees" and such utterances "may be followed by the speaker's own remarks of reflection on their own subject knowledge" (p. 165).

Others think *ne* marks significance of some message. For example, Alleton (1981) advocates that the basic function of *ne* is "appealing to his listener's (or listeners') active participation" (p. 111). King (1986) maintains that the core function of *ne* is "making a metalinguistic comment and, by extension, expressing his attitude towards the content of the utterance" (p. 27). King also elaborates that *ne* cooccurs with the four types of evaluators in discourse: a. Repetition of previously mentioned or presupposed information; b. Rhetorical questions by which the speaker pauses and steps back to survey the situation; c. Explanations or additional details concerned with previously mentioned information, often in the form of parenthetical comments; d. Direct address by speaker to hearer. (p. 29) *Ne* marks the evaluative information more significant (p. 43). Li (2006) thinks that "When it (*ne*) occurs in declaratives, it indicates that the speaker considers the content that is being claimed to be extraordinary; when it occurs in wh- and A-not-A questions, it indicates that the speaker considers the matter that is being questioned to be of particular importance." (p. 21) Li is arguing that the unitary function of *ne* in both declaratives and interrogatives is to tell the hearer about the

unusualness in the speaker's words, in spite of different wordings in his description.

There are also scholars who have noticed a softening function of *ne* in regard to politeness. Lee-Wong (1998) endorses that *ne* signals speaker's uncertainty which "weakens the illocutionary force of a bald on record request in a context of high R" (p. 399). A high R means a request is of high imposition on the hearer. Liu, Pan, and Gu (2001) think that *ne*'s primary function is to soften the tone (主要功能是缓和语气) (p. 419) while it is used in utterances that attempt to solve some puzzles for the hearer (常表示说话人有些困惑，或说话人认为听话人对某事不清楚，试图为之解惑) (p. 423). Shi and Huang (2016) maintain that *ne* adds a gentle overtone to interrogative sentences or other sentence types (pp. 36-37).

In my mind's eye, some "now-generally-agreed-upon positions" (hereafter, NGAUPs)[24] have emerged from previous studies on *ne*:

NGAUP 1: *Ne* is used in conversations.

NGAUP 2: *Ne* can be used in both declaratives and interrogatives.

NGAUP 3: *Ne* is used when the speaker wants to say something vis-à-vis what the speaker/hearer has said or has in mind.

NGAUP 4: *Ne* marks the progressive aspect. (e.g. mentioned in Alleton (1981) and Wu (2005))

NGAUP 5: *Ne* is used due to a consideration of politeness by softening the tone.

I will return to these NGAUPs in the Discussion chapter.

1.4.5. 啊 *a*

- Entry 1：用在感叹句末，表示增强语气 'Used at the end of an exclamative sentence to strengthen the overtone'

For example,

(28) duo hao de tianer a
much good DE weather A
'What a nice day!'

[24] This phrase is borrowed from Enfield (2005, p. 194).

- Entry 2：用在陈述句末，使句子带上一层感情色彩 ‘Used at the end of a declarative sentence to add a flavor of emotion to the sentence’

For example,

(29) zhe hua shuo de shi a

this speech say DE make.sense A

‘This thought makes sense.’

- Entry 3：用在祈使句末，使句子带有敦促或提醒意味 ‘Used at the end of an directive sentence to add a hortatory tone or a sense of reminder’

For example,

(30) ni ke bie gaosu Xiaodeng a

you bear.in.mind don’t tell person.name A

‘Don’t tell Xiaodeng for God’s sake!’

- Entry 4: 用在疑问句末，使疑问语气舒缓些 ‘Used at the end of an interrogative sentence to soften the tone’

For example,

(31) ta shenme shihou lai a

he what time come A

‘When will he come?’

- Entry 5：用在句中稍做停顿，让人注意下面的话 ‘Used to mark a pause in sentence so as to remind the listener to pay attention to following utterances’

For example,

(32) zhe xie nian a, zanmen de rizi yue guo yue hao la

this few year A, we DE day more live more good PRT

‘These years our life is better and better.’

- Entry 6: 用在列举的事项之后 ‘Used after listed items’

For example,

(33) shu a, bao a, zazhi a, bai man le yi shujia

book A, newspaper A, magazine A, place(v.) full LE one bookcase

‘Books, newspapers, and magazines fill up all the bookcase.’

- Entry 7: 用在重复的动词后面，表示过程长 ‘Used after repeated verbs to indicate the lengthiness of action’

For example,

(34) xiangqinmen pan a, pan a, zhongyu pan dao le zhe yi tian
fellows hope A, hope A, finally hope arrive LE this one day
‘Fellows have been hoping all the time and finally this day comes.’

This SFP is a thornier issue (Simpson, 2014, p. 161). It has more dictionary entries than *ne*. The entries suggest that *a* can be used in declarative, directive, exclamative, and interrogative sentences. These are actually four of the five major Mandarin sentence types proposed by Zhan and Bai (2016). The fifth is vocative sentences, such as 老张 *lao Zhang* ‘Old Zhang (Zhang is a surname)’ (p. 403). My native speaker language feeling tells me that actually *a* can also be tagged with vocatives and therefore 老张啊 *lao Zhang a* ‘Old Zhang a’ is a perfectly acceptable token. Summarizing, *a* can appear in all five major Mandarin sentence types. Zhan and Bai in the same book chapter maintain that a is “the most commonly used sentence-final particle in exclamative sentences” (p. 412) and is also one of the typical SFPs that go with directive sentences (p. 422). These facts corroborate its great “flexibility and multifunctionality”.

Its functional multifunctionality motivates most scholars to adopt the minimalist approach in the study of a, because the minimalist approach seems to be an easier task by avoiding too many nuts and bolts. For example, Lee-Wong (1998) advocates that for *a*-tagged utterances of requests, its function is to signal informality and politeness in requests by contracting the distance between speaker and hearer and therefore rendering the requests casual. There are nevertheless scholars adopting the maximalist approach and come up with laundry-list style findings, such as Chao (1968).

As minimalists, Li and Thompson (1981) argue that *a* “performs the function of reducing the forcefulness of the message conveyed by the sentence” (p. 313). Wu (2004) thinks that a can be divided into interrogative *a* and non-interrogative *a* (p. 179). Wu believes that this particle can be further prosodically categorized as two different variants: "*a* with a notably low pitch and *a* with a flat or a slightly higher pitch " (p.

128). Low pitch *a* is mainly used in interrogatives. It is mainly used in *a*-formulated questions and *a*-attached questions. In such cases, "the speaker (also) conveys that the matter at issue is, from his or her perspective, not merely unexpected, i.e. previously not expected, but also counter to expectation " (p. 146). The flat or higher pitch *a* is mainly used in non-interrogatives (p. 179). In such cases, "final *a* may be used to indicate that the question is being launched to deal with some problematic and/or unexpected situation, or it may be used to mark the very launching of this question as in some way problematic" (p. 153). It is mainly used in three contexts, viz. informing, disagreeing, and the ones with "with some problem related to sequential contingency " (p. 214). A common function for all these final *a* tokens is to index "that the matter being addressed departs from how the matter should be, or normally is, from the a speaker's perspective " (p. 222). Wu's theory is echoed by Li (2013), who advocates that host utterances for *a* suggest "surprise arising from lack of background information" and such utterances "may trigger responses from the listener who may provide more information" (p. 165).

Those previous minimalist accounts shed lights upon *a*'s polarized pragmatic functions. It can interestingly both reduce the forcefulness while register a surprise or unexpectedness. Making the waters muddier, Li (2006) considers "*a* to be a discourse marker, which functions to highlight the relevance of the utterance in which it occurs to the discourse context" (p. 50). The brief review suggests that no NGAUP has been reached about *a*'s core function.

1.5. Research Gaps

My literature review suggests that scholars have produced extensive insights on the three SFPs. While they have laid the groundwork, many questions remain to be answered. In particular, the following research gaps have been identified in the review of previous studies.

First, there are gaps in the data. Previous studies tend to draw conclusions that are at variance with claims based on analysts' intuitions or on the particularities of small

datasets. Some researchers tapped into their native speaker intuition and thought of examples for their writing or resorted to sentences taken out from literary works (e.g. Lyu, 1999; Wu, 2005). When it comes to intuition, as Bybee (2013) points out, although it is "interesting and important", "language users are often unaware of the nature and frequency of certain structures that they use" (p. 50). In other words, intuition may not be reliable for linguistic generalization.

Others based their analysis on naturalistically collected data, but the limited sample size of those datasets leads to a narrow focus on just a limited number of uses of a certain SFP (e.g. Wu, 2004), typically within a Conversation Analysis framework. As pointed out by Jing-Schmidt (2019), this line of inquiry has yielded fruitful results particularly in the field of Conversation Analysis (CA) and sociolinguistics. As sociolinguistics is out of the scope of this study, breakthroughs can be made vis-à-vis previous CA studies. CA draws on a restricted database and selective quoted segments as data (Ten Have, 1990). I will use Wu (2005), one CA-based SFP study, to briefly illustrate both the CA approach and CA-based SFP studies. That study focuses on *ou.* One example provided by that study is reproduced as follows,

(5) (Tea Time s079b)

1W: yedan- ershisi:hao chi wan fan zhihou

Christmas the:twenty:fourth eat finish meal after

'Christmas- On the twenty-fourth, after finishing dinner,'

2W: ta jiu gen- yao gen ta nanpengyou fei dao- (.) xiaweiyi qu

she then with will with she boyfriend fly to Hawaii go

'she then with- was going to fly to- (.) Hawaii with her boyfriend.'

3W: chi wan- mashang jiu yao zou le **ou**"

eat finish immediately then will leave LE **OU**

'Right after she finished the meal- then she was planning to leave OU"'

4 (.)

5W: wo jiu shuo- ou zhende ou. na nide- fu- fumu bu hui zenmeyang

1SG then say PRT really PRT then your parents NEG will react

‘I then said- ‘‘Oh, really? Then won’t your- pa- parents react?’’’

6W: ta jingyade kan zhe wo

she surprisingly look IMP 1SG

‘She was staring at me, surprised,’

7W: fu(mu) zenme hui shuo zenmeyang ne

parents how:come will say react NE

‘(and was like), ‘‘Why should (my) parents react?’’’

8 (.)

9W: hen jingyade kan zhe wo, wen wo- (.) %why%

very surprisingly look IMP 1SG ask 1SG why

‘She stared at me, looking very surprised, and asked me, ‘‘%Why%?’’’

10W: (wo shuo) ou

1SG say oh

‘(I then said,) ‘‘Oh.’’’

11 (.)

12W:) wo jiu shuo, <na shihou yishi dao shuo ou tch!

1SG then say that time aware arrive say PRT

‘I then said, <(Only) then did I start to become aware that tch!’

13H:) bu yi[yang

NEG same

‘(it’s) differe[nt.’

14L:) [jiazhi guan bu tong la.

value view NEG same PRT

[‘the (cultural) values are different.’

This one excerpt illustrates the merits you can find in other CA studies as well. They are based on naturally occurring conversations. They have a well-developed transcription system that can almost reflect all particularities of conversations. Specific

to SFP studies, such fine-grained approach can really detail the nuts and bolts of SFP usage in concrete speech events. However, this seems a dilemma. The CA fine-grained analysis confines researchers' horizon since such an analysis is based on a very limited number of selected conversation excerpts. In a nutshell, the data adopted in CA studies is of a limited size.(Bybee, 2013)(Bybee, 2013)(Bybee, 2013)(Bybee, 2013)

Furthermore, as far as I know, there has not been a quantitative study of Mandarin SFPs yet. In other words, no researcher has conducted a quantitative study based on a relatively large sample of SFP utterances to obtain "the patterns or trends" of the target structure (Johnson, 2013, p. 288). This shortcoming would have been excusable if large Chinese language corpora were unavailable or inaccessible. However, corpus data availability has become a non-issue thanks to the rapid development of digital language technology in the last two decades.

Second, there are gaps in the theoretical framework. Previous studies mostly treated the functions of SPFs independent of the larger constructional context in which they are used. They also tend to insert an artificial boundary between SFPs and the preceding linguistic unit of an utterance while in real-world conversation a SFP is an organic part of an utterance. For example, when Zhan and Bai (2016) define the canonical structure of a Chinese sentence, they maintain that "(I)t consists of a main clause, which can have a complicated internal structure, and some peripheral elements, mainly sentence-final particles." (p. 401) This definition clearly separates SFPs from the rest linguistic elements in a sentence, especially forgetting the fact that the SPF serves a different function when used in a different constructional pattern. Jing-Schmidt terms this approach as "the morpheme-based approach" (personal communication), pointing out the limited focus of such an approach. Instead, my data show that a more reasonable view is to see the meaning of an SFP-tagged utterance as something depending on the larger constructional context. For example, which will be elaborated later in the dissertation, my data show that the SFP *bei* cooccurs with structures such as reduplicated verbs in utterances that have a function of mitigated requests while cooccurs with negatively evaluative expressions in utterances that have a function of

bald assertion. This insight is not captured by previous morpheme-based studies on this SFP.

Third, the gap in the theoretical framework has methodological consequences. Systematic classifications of the constructional subtypes of the SPFs and quantitative analyses of their usage patterns have not been provided in the literature.

In this study, I adopt a usage-based constructionist approach that pay attention to patterns of uses, as a remedy of this dilemma, which I will introduce in Section 2.2. In addition, keywords and terms related to my theoretical framework of this current study did not appear on the keyword and the term list that I presented in the Overview section. These keywords and terms include “construction”, “usage”, “semantic prosody”, and “collocation”. This fact suggests that scholars have not investigated SFPs through the lens of these theories because these keywords and terms do not have a high frequency of occurrence in literature on pragmatic particles. All in all, this means my dissertation will be a fruitful contribution to the current scholarly discussion. My theoretical framework then will be illustrated in Chapter 2.

1.6. Research Questions

Specifically, I will address the following research questions (RQs):

1) What are the distribution patterns of the three SFPs?

In other words, as SFPs are polyfunctional, it is challenging to come up with a description of their functions by a morpheme-based approach due to a lack of other contextual clues. Then, if I zoom in on the co-texts and spot what expressions they tend to co-occur, the result can offer new insights.

2) Are the functions of SFPs specific to the constructional contexts in which they are used? If so, what are those functions?

For now, you can just see “construction” as something like a “chunk”, which will be defined in Chapter 2. This concept is the core of the usage-based constructionist approach (defined in Chapter 2) that I adopt in this current study. I will see SFPs as part of constructions and associate their functions to the constructions, instead of associating

the functions to SFPs as morphemes.

1.7. Organization of this Dissertation

To answer these research questions, this dissertation is organized into eight chapters. Chapter 1 is the introductory chapter. Chapter 2 presents my theoretical framework. Chapter 3 introduces my data and methodology. Chapter 4 to 6 are case studies of SFP *bei*, *ne*, and *a* respectively, each focusing on the distribution patterns and construction-specific functions. Chapter 7 brings to you a discussion of my findings. Chapter 8 concludes the dissertation.

CHAPTER II

THEORETICAL FRAMEWORK

2.1. The Semantic Prosody Theory in Neo-Firthian Corpus Linguistics

Previous studies have shown that it is not effective to use either introspection-based data or small size empirical data in the study of SFPs. On the other hand, neo-Firthian corpus linguists have actually offered us some useful insights to address the two research gaps, i.e. the data and the approach.

First, the data. Let's start with some definitions. A corpus is defined by the Sketch Engine website (https://www.sketchengine.eu/corpora-and-languages/corpus-types/) as follows[25]:

> "A text corpus is a very large collection of text (often many billion words) produced by real users of the language and used to analyse how words, phrases and language in general are used. It is used by linguists, lexicographers, social scientists, humanities, experts in natural language processing and in many other fields. A corpus is also be used for generating various language databases used in software development such as predictive keyboards, spell check, grammar correction, text/speech understanding systems, text-to-speech modules and many others."

The terms "corpus" and "text corpus" will be used interchangeably in this dissertation. Nowadays a typical corpus is a "machine-readable collection of language used in authentic settings/contexts" (Gries & Newman, 2013, p. 258). "Machine-readable" means the corpus data can be analyzed by linguists via "a Web interface", a specifically designed tool, "a ready-made general corpus program", or a general purpose programming language (Gries & Newman, 2013, pp. 279-280). Prototypically, a corpus consists of "a body of naturally occurring language" (Gries & Newman, 2013,

[25] Lexical Computing Limited. (n.d.) *What is a corpus?* Sketch Engine. Retrieved November 10th, 2020 from https://www.sketchengine.eu/corpora-and-languages/corpus-types/.

p. 258).

Widely used large general corpora of English include The British National Corpus (BNC, 100 million words), The Corpus of Contemporary American English (COCA, 560 million words), etc. (A comprehensive list of such corporate can be found in the Appendix section of Gries and Newman (2013)) Widely used large general corpora of Mandarin Chinese include the CCL corpus (http://ccl.pku.edu.cn:8080/ccl_corpus/ , a 700-million-word corpus), the BCC corpus (http://bcc.blcu.edu.cn/ , a 15-billion-word corpus), etc. Part of the data for this present study comes from the BCC corpus. More information of this corpus will be provided in the Methodology chapter. In addition, small, specialized corpora have also yielded fruitful results in linguistic research. For example, Berkenfield (2001, cited from Gries & Newman, 2013) used a corpus of 10640 words for the author's research on the phonetic reduction of "that" in spoken English.

Speaking of the notion of corpus, this current study is a study in corpus semantics. This approach of semantics studies is the one "using corpus evidence to study meaning" (Stubbs, 2001, p. 4). The corpus I use in this study is the BCC corpus, which provides naturally occurring tokens of SFPs as data, instead of intuitive data. The corpus also provides a large amount of data, instead of a sample of a limited size. The ready-made corpus program Antconc and the software Excel assist me to find patterns of SFP usage, instead of misled by individual tokens. All those methodological details will be presented in the Methodology chapter.

Let me further clarify the merits of corpus for this current study by citing the two principles of corpus linguistics, summarized by Stubbs (2001):

> "Principle 1: The observer must not influence what is observed. Data and analysis must be independent. What is selected for observation admittedly depends on such factors as convenience, personal interests and prior hypotheses. Nevertheless, corpus data were part of natural language use and not produced for purposes of linguistic analysis.
>
> Principle 2: Repeated events are significant. The first task of corpus linguistics is

to describe what is usual and typical. Unique events certainly occur, but can be described only against the background of what is normal and expected. The frequent occurrence of lexical and grammatical patterns is good evidence of what is typical and routine in language use". (p. 221)

Therefore, according to Principle 1, corpus data as evidence is a better option than intuitive data, because I as the researcher is independent of the corpus data. According to Principle 2, a quantitative study based on corpus data is a better option than previous studies based on a limited number of selected conversation excerpts, because with the help of a corpus we will be able to see what is "typical and routine" of SFP use, instead of misled by particularities.

Second, the approach. As Albert Einstein have told us, "Problems cannot be solved at the same level of awareness that created them."[26] A central tenet of neo-Firthian corpus semantics is that the locus of meaning is not in isolated words but in larger units. As a truism goes, "no word is an island" (Hoffman et al., 2015, p. 1) This thought is inspired by the linguist J. R. Firth and can be best illuminated by his slogan: "You shall know a word by the company it keeps" (Firth, 1957, cited from Desagulier, 2017b, p. 10). In line with the central tenet, a key relevant notion adopted in this dissertation is "extended units of meaning", proposed by John Sinclair. Sinclair is "the most prominent proponent of the neo-Firthian approach", who carries the Firth's thought into practice in corpus linguistics (McEnery & Hardie, 2012, p. 122).

Sinclair argues against the traditional equation "word = unit of meaning" (Sinclair, 2004, p. 25). Instead, he proposes that the unit of meaning is what he coined as "lexical item", which consists of five categories: core (also called a node word) and semantic prosody as the obligatory ones and collocation (explained later in this section), colligation (the tendency for a core to co-occur with grammatical choices such as some certain word class), and semantic preference (the tendency for a core to co-occur with linguistic items that have some certain meaning) as three optional categories to arrive at the semantic prosody.

[26] *Master 60'*. Retrieved November 11th, 2020 from https://www.master60.com.tw/master-quote.php?id=126 .

The semantic prosody theory can be dated back to J. R. Firth' dictum "a word is characterized by the company it keeps" (Firth, 1968. cited in Jurafsky & Martin, 2019). The word "prosody" suggests that the locus of word meaning is not just in a word, but beyond that. A word creates meanings together with other units in its immediate contexts. In other words, the study of meaning needs to start from a lexical item, instead of a single word.

To use an oft-cited example in the semantic prosody literature as an example, let's start with the English expression "naked eye," as is detailed in Sinclair (2004). He argues that this phrase means far more than "unclothed organ of sight" (p. 31). Instead, he uses corpus evidence to demonstrate that this expression is actually part of the lexical item "visibility + preposition + the + naked + eye" (p. 33). On the N-3 position, i.e. the third position on the left of "naked eye", verbs or adjectives regarding "visibility" frequently occur. The most frequent verb is "see", others including "detect, spot, spotted, appear, perceived, viewed, recognized, read, studied, judged." The most frequent adjective is "visible", others including "apparent, evident, obvious and undetectable." On N-2, i.e. the section position on the left, prepositions occur, where "with" goes with verbs, as in ". . . you can see with the naked eye . . .", and "to" goes with adjectives, as in ". . . just visible to the naked eye . . ." (p. 32). On N-1, the position immediately on the left, "the" is the predominantly frequent word in 95 % instances (p. 31).

Furthermore, Sinclair reveals that this lexical item tends to co-occur with linguistic items which are laden with a semantic prosody of "difficulty" and these linguistic items are on the left of the N-3 position. For example, this semantic prosody is instantiated by adjectives such as "small, faint, weak and difficult" as in ". . . too faint to be seen with the naked eye . . ." (p. 33), by the adjective "invisible" (p. 34), or implied by the modal "can" or "could", as in ". . . these could be seen with the naked eye from a helicopter . . ." (p. 34). The semantic prosody of "difficulty" is the **function** of the lexical item "visibility + preposition + the + naked + eye" in communication. In other words, speakers use this lexical item to express "some kind of difficulty" regarding the process of seeing (p. 34).

Sinclair is right that "The beginning of the item is very difficult to detect normally, because it is so variable; on the other hand the end is fixed and obvious" (p. 34). The right end of this lexical item is invariably "naked eye", but on the left, especially the left of its N-3 position, what frequently co-occur include adjectives (small, faint, etc.) or modals (can or could), which cannot be easily categorized.

For identification of semantic prosody, I used collocation as the tool to arrive at SFP semantic prosody in my analysis of the corpus data. As mentioned earlier in this section, collocation is one category making up a lexical item (the other two being colligation and semantic preference). It has been shown to be an effective tool. For instance, apart from the "naked eye" example just laid out, Sinclair (2004) also reveals that the phrase "true feelings" not only means "genuine emotions", but also belongs to a lexical item that has a semantic prosody of reluctance by collocating with verbal structures such as "will never reveal, prevents me from expressing, careful about expressing, less open about showing, guilty about expressing" and a semantic prosody of inability by collocating with "try to communicate, incapable of experiencing, unable to share" (p. 36). In another study, Sinclair (2004a) demonstrates to us that the word "budge" belongs to a lexical item which has a semantic prosody of refusal by collocating with "won't, wouldn't . . . etc., and refuse . . ." and a semantic prosody of inability by collocating with "can't, couldn't" (p. 144).

A brief introduction to the concept of collocation should be in place. Collocation depicts what are the frequent co-occurring linguistic units in the immediate co-texts of a word. Such frequent "neighbors" of a word are referred to as collocates (Gries, 2009, p. 14). The Corpus of Contemporary American English (COCA) also defines "collocate" this way: "Collocates are words that occur near a given word (the node word), and they can provide very useful insight into the meaning and usage of the words near which they occur"[27]. The COCA definition especially emphasizes collocate's function in revealing the "meaning and usage" of the words, which again adumbrates the adoption of collocational analysis in this project as it can shed light on the meaning and usage of

[27] Mark Davies. (n. d.) *Collocates data*. The Corpus of Contemporary American English. Retrieved August 28th, 2020 from https://www.collocates.info/.

SFPs.

As an example, Gries shows us the collocate display of two English adjectives, *alphabetic* (Table 5) and *alphabetical* (Table 6), based on corpus data of the British National Corpus World Edition (BNC).

In the two tables, “L1” means the first position on the left of the node word, i.e. “alphabetic” or “alphabetical”. “R1” means the first position on the right. “Freq” stands for “frequency.” Note that “alphabetic” tends to co-occur with words related to writing, e.g. “literacy” and “character”, while “alphabetical” with words related to order, e.g. “order” and “list”. Such collocational disparity sheds light on the two words’ different meanings and usages.

To sum up at this point: I adopt the neo-Firthian approach to arrive at the semantic prosody of SFPs by zeroing in on their collocates as functional indicators, based on corpus data. In fact, little research has been done on semantic prosody of languages other than English (Andersen, 2017, p. 130; McEnery & Hardie, 2012). My work therefore can be a fruitful application of this theory in Mandarin Chinese, and in SFP studies particularly.

Table 5: Gries' collocate display of "alphabetic"

Word at L1	**Freq L1**	**Node word**	**Freq Node**	**Word at R1**	**Freq R1**
of	8	alphabetic	42	literacy	7
the	6			writing	5
an	5			order	3
in	2			character	3
such as	2			and	2
our	2			system	2
when	2			characters	2
widespread	1			culture	2
systems	1			in	1
varying	1			.	1

Table 6: Gries' collocate display of "alphabetical"

Word at L1	Freq L1	Node word	Freq Node	Word at R1	Freq R1
in	77	alphabetical	234	order	89
an	36			index	15
the	23			list	13
of	6			indexing	12
and	6			subject	12
.	6			sequence	11
,	6			listing	9
ascending	5			guest	6
or	5			and	5
strict	4			description	2

As a matter of fact, several previous studies on pragmatic particles have explicitly or implicitly adopted a methodology that has used collocation as an analytical tool. Aijmer (2002) points out that "the collocations and cooccurrences of discourse particles with other elements serve as linguistic clues to the interpretation of their functions" (p. 30). He lists collocation as one of the "functional indicators" that help interlocutors to arrive at the "intended interpretation" of discourse particles (p. 27). Aijmer dubs pragmatic particles as "discourse particles" and he is telling us that frequent linguistic units in immediate contexts of pragmatic particles can give us clue to what are the functions of the particles. In the case of particles that have multivariate functions, their co-occurring "neighbors" can help co-participants of conversations to unveil this multifunctionality. One relevant example is the Kishner and Gibbs (1996) study on the English pragmatic particle *just*, mentioned by Aijmer, which shows that intensifying *just* tends to co-occur with "other particles or hedges (most frequently *just sort of* ; 36 examples), with scalar adjectives or expletives (*just awful*, *just bloody well*) and with exclamatory so and such" (Aijmer, 2002, p. 161). Note that the "other particles or hedges", "scalar adjectives or expletives", and "exclamatory" are consonant with the intensifying meaning of *just*.

As another example, Zhao and Sun (2015)'s corpus-based study found that SFP *ba* in declarative sentences tends to cooccur with adverbs denoting uncertainty or speculation, such as *dagai*, *dayue*, *dadi*, *huoxu*, *yexu*, *kongpa*, *sihu*, *haoxiang*, *keneng*, *xiangbi*, *buhui*, *buneng*, *bujiande*, *suanshi*, and *gaibushi*, while repulses adverbs denoting certainty or strong assertion, such as *mingming*, *fenming*, *mingbaizhe*, *mingxian*, *xianran*, *gongran*, *dangran*, *nanguai*, *juran*, *jingran*, *guoran*, *guozhen*, *guaibude*, *yuanlai*, *pianpian*, and *pianqiao* (p. 125). The authors then point out that this of *ba*'s collocational profile is in line with its core meaning of "speculation."

Other studies have implicitly employed collocation in their analysis. For example,

- Lin 1984

The author argues that the central meaning of Mandarin *ne* is the semantic feature [+contrastiveness]. When he counters one previous claim that *ne* expresses the semantic feature [+exaggeration], the author maintains that for those *ne*-tagged utterances that have the semantic feature [+exaggeration], this meaning does not solely come from *ne*. The author believes the meaning is in fact "mainly expressed by other words in the context." (p. 218) The "other words" are clearly collocates and this of Lin's updated understanding of *ne* semantics shows that collocation is an effective indicator of the meaning this SFP.

- Luke 1990

The author compares LA and LO in Cantonese by investigating the "distributional and co-occurrence behavior of the two particles" (p. 188). He shows that the two semantically different particles systematically co-occur with different words or phrases in speech. LA is dubbed by the author as a "looking forward" particle, meaning LA tells other interlocutors that they need to pay attention to the speaker's following words, while LO is a "looking backward" particle which tells interlocutors to think about what has been said by the speaker because his or her previous words are the basis against which the speaker's current utterance can make sense. One example given by Luke is that LA is habitually followed by continuers such as "mhm" and "mm" whereas this happens less for LO. The continuers are semantically compatible with the "looking

forward" nature of LA as both of them indicate that more is to come. The continuers are not compatible with LO as this particle points to the opposite direction in discourse.

- Chan 1996

The author thinks that the Cantonese sentence-final particle *je* mainly serves a delimitative function and this function "is especially obvious in sentences containing some quantification" (p. 29). Another reading of this statement would be *je* tends to co-occur with quantification terms. Indeed, the corpus sentences given by the author all have such terms in them, e.g. 五百 (five hundred), 百分之幾 (a percent of), when she explicated this quantification function.

- Xiang 2011

In her explication on the *la¯h* particle in Shishan dialect in Southern China, Xiang explains that her attention "also focuses on co-occurrences of *la¯h* with other linguistic forms of restrictivity, such as rhetorical questions, tautological construction, and other restrictivity/exclusivity-marking particles." (p. 1382) She provides with examples of such co-occurrence. For instance, she finds that "Overall, in the 139 tokens of *la¯h*, as many as 44 tokens (31.7%) co-occur with other emphatic markers/constructions of restrictive functions" (p. 1384), and the author thinks this finding lends credence to her argument of this particle as "an emphatic marker of restrictivity" (same page).

- Lyu (1999) proposes three functions of *bei*:

a. 表示道理简单，无须多说 (To indicate that a state of affair is obvious and therefore no need to elaborate on it.)
b. 用在"动+就+动"的句子末尾，这种句子表示"没关系"、"不要紧"。(*bei* is used at the end of sentences that have "verb + *jiu* + verb" as the core structure and such sentences mean "it not a big deal" or "it does not matter")
c. 用在"就得了"、"就行了"等之后 (*Bei* follows short clauses such as "jiu de le", "jiu xing le", etc.)

Clearly, we see that in the latter two functions, *bei* tends to co-occur with *jiu*. The DMC Dictionary explains the function of *jiu* as 表示在某种条件或情况下自然怎么样 (to indicate what the state of affair will naturally be under some certain circumstances).

Both *jiu* and the particle *bei* convey an import that something should naturally happen and other co-participants in the discourse should not worry about this. Likewise, Zhao and Shi (2015) based their analysis on 29 example sentences of *bei*, and 15 of them (half of them) have *jiu* in them. Although the authors did not mention this, my careful inspection of their choice of example sentences reveals that *bei* tends to co-occur with *jiu*. In addition, Liu et al. (2001) provide three representative example sentences for their explanation for *bei* and one of them contains *jiu*: 你要去就去呗，跟我有什么关系！(If you want to go, just go. What does this have to do with me?).

- Sato (2017)

The author points out that English *I think* as a final particle can be deployed to express either a high certainty or a low certainty, as its subjective meaning components. I noticed from the examples given by the author that which subjective meaning is expressed is closely related its collocates. In the author's examples of *I think* that expresses high certainty, it cooccurs with words "absolutely", "certainly", and "quite". In the authors' examples of *I think* that expresses low certainty, the author, based on those examples, maintains that "the association of final *I think* with mitigation is attested in contexts where the speaker presents specific dates, numbers, or proper nouns as an instantaneous upshot based on limited memory" (p. 89).

Before I wrap up this section, I want to say that several other neo-Firthians have also proposed theories regarding the notion of "extended units of meaning." For example, based on corpus evidence, Stubbs (2001) argues that "combinations of words in phrases are therefore a good candidate for the basic semantic unit of language in use" (p. 14). His examples include an analysis of the word "surgery", where he shows that this word's four specific senses are actually determined by what phrase this word occurs in, as in 1) plastic surgery; 2) progress in surgery (has made heart transplants possible); 3) (she had) her surgery in Cemetery Road; 4) (she was taking) evening surgery. He thus concludes that "…cases of apparent multiple ambiguity at word level are usually illusory: they dissolve in context" (p. 14). He maintains that "it is not individual words which are the basic units of meaning, but longer phrases and collocations" (p. 57), thus

supporting the notion of extended units of meaning.

Another neo-Firthian Michael Hoey (2005) proposes the lexical priming theory[28]. He observes that:

> "As a word is acquired through encounters with it in speech and writing, it becomes cumulatively loaded with the contexts and co-texts in which it is encountered, and our knowledge of it includes the fact that it co-occurs with certain other words in certain kinds of context". (p. 8)

Therefore, he argues that collocation is pervasive in language.

In Section 2.3, I will argue that the semantic prosody approach (reproduced here: the neo-Firthian approach to arrive at the semantic prosody of SFPs by zeroing in on their collocates as functional indicators, based on corpus data) enlightened by the notion of "extended units of meaning" can be incorporated into the usage-based constructionist approach which will be elaborated in Section 2.2, thus providing a feasible, unified approach for this current study.

2.2. The Usage-based Constructionist Approach

In this section, I will argue that the usage-based constructionist approach also offers useful thoughts for this current study of SFPs.

"Construction" is a concept in Construction Grammar, a field which "is still very young, highly diverse, and undergoing rapid development." (Hilpert, 2014, p. xi) Therefore, different brands of Construction Grammar exist (Jing-Schmidt, 2015). This diversity thus necessitates a clear delineation of construction in the current study. Let's first see representative examples of construction, provided in Goldberg (2003, p. 220) and reproduced here in Table 7. Goldberg's examples indicate that constructions cover a plethora of types of linguistic units, ranging from morphemes to argument structures such as Ditransitive. This indicates the difficulty to define what a construction is. The definitions provided by Adele Goldberg, who has produced influential work in this field

[28] I will not present the details of the lexical priming theory because this is fundamentally a psychological theory and therefore falls outside the scope of this current study. Detailed information can be accessed in Hoey (2005).

(Hilpert, 2014, p. xii), are presented following Table 7:

Table 7: Examples of construction given in Goldberg (2003)

Construction	**Form/Example**	**Function**
Morpheme	e.g. anti-, pre-, -ing	
Word	e.g. Avocado, anaconda, and	
Complex word	e.g. Daredevil, shoo-in	
Idiom (filled)	e.g. Going great guns	
Idiom (partially filled)	e.g. Jog (someone's) memory	
Covariational-Conditional construction	Form: The X-er the Y-er (e.g. The more you think about it, the less you understand)	Meaning: linked independent and dependent variables
Ditransitive (double-object) construction	Form: Subj [V Obj1 Obj2] (e.g. He gave her a Coke; He baked her a muffin)	Meaning: transfer (intended or actual)
Passive	Form: Subj aux VP_{pp} (PP_{by}) (e.g. The armadillo was hit by a car)	Discourse function: to make undergoer topical and/or actor non-topical

- Goldberg (1995)

C is a CONSTRUCTION if and only if C is a form-meaning pair <*F*i, *S*i> such that some aspect of *F*i or some aspect of *S*i is not strictly predictable from C's component parts or from other previously established constructions. (p.4)

- Goldberg (2003)

Constructions are stored **pairings of form and function**, including morphemes, words, idioms, partially lexically filled and fully general linguistic patterns. (p. 219)

- Goldberg (2006):

Any linguistic pattern is recognized as a construction as long as some aspects of its **form** or **function** is not strictly predictable from its component parts or from other constructions recognized to exist. In addition, patterns are stored as constructions even if they are fully predictable as long as they occur with sufficient frequency. (p. 5)

- Goldberg (2019)

> constructions are understood to be emergent clusters of lossy memory traces that are aligned within our high- (hyper!) dimensional conceptual space on the basis of shared **form**, **function**, and contextual dimensions. (p. 7)

Two central criteria of construction identification emerge from Goldberg's definitions, whereby her most recent emphasis on memory traces in the accumulation of frequency effect adds nuance to the definitions:

- Non-predicability (i.e. A linguistic structure is a construction if the function of the whole construction is not equal to the function of any component or the sum of components' functions)
- Accumulating frequency (i.e. A linguistic structure is a construction if it is frequently used)

Let me explain the two criteria. First, non-predicability. This means a direct association between function and surface form (Goldberg, 1995, 2003; Jing-Schmidt, 2015) in a construction otherwise described as a **form-function pairing**. "Form" and "function", as readers probably have noticed, are the two keywords in Goldberg's definitions. Non-predicability tells us that the function of a surface form is not predicable from its components.

The direct association can be illustrated by two famous examples in the literature. Goldberg (1995) shows us that the sentence "John sneezed the napkin off the table" instantiates the English caused-motion construction, i.e. [Subject + Verb + Object + Directional]. This construction as a whole functions to express the meaning something like "Some agent causes some theme to move along some path." The construction, according to Goldberg, coerces the verb "sneeze" to be followed by an object, "the napkin" in this case. ("Coercion" is a widely discussed notion in construction grammar. Readers can refer to Chapter 5 of Traugott and Trousdale (2013) to get a sense of it.) The merit of such a constructionist theory is therefore that it can elegantly explain why a typical intransitive verb "sneeze" ("intransitive" means it is not supposed to be followed by an object, such as "the napkin" in the example sentence), is used as a

transitive verb (followed by object(s)) in the example sentence. Put it another way, the function of expressing the caused-motion is directly associated with the surface form of this construction rather than with the verb.

Similar examples in English include: the grammatical construction [Be + Going to + Verb] expresses future tense (Traugott & Trousdale, 2013); the construction [Noun Phrase + Had + (negative) + 've + Past Participle] expresses counterfactual, as in "If you had've eaten it, you would have died" (Fillmore, 1985); the [Clause A + Let Alone + Clause B] construction enables the understanding that the message conveyed by Clause A is more informative while the message by Clause B more confidently emphasized (Fillmore et al., 1988); etc. Here is how the direct association between function and surface form works in the three examples: the future tense interpretation is not predicable from "going to"; the counterfactual interpretation not from what Fillmore dubbed as the redundant have, i.e. 've; and the informativeness or the emphasis of message not from "let alone" in the construction; rather, the functions are associated with the constructions.

Let's see two relevant examples in studies of Mandarin. Jing-Schmidt (2015) demonstrates that for *ba*-sentences in Mandarin, a thorny issue in this field, their idiosyncratic meaning does not come from what transformation theorists believe that such sentences are transformed from their counterpart Subject-Verb-Object (SVO) sentences. Instead, *ba*-sentences instantiate *ba*-construction. The idiosyncrasy is associated with the construction, not with *ba* itself or any other component in the construction. Likewise, Jing-Schmidt (2017) argues that the linguistic representation of counterfactual thinking in Mandarin is done by counterfactual constructions rather than any lexical or grammatical category that appears in counterfactual utterances.

A central idea upheld by Construction Grammarians is that "the totality of our knowledge of language is captured by a network of constructions, i.e. a '**construct-i-con**'" (Goldberg, 2003, p. 219). In other words, the basic units of language are constructions (Goldberg, 1995, 2006). To know a language is just to know its constructions, or its construct-i-con (Hilpert, 2014b). Construction, Croft (2001)

maintains, is the only primitive grammatical unit (p. 362). By this he means that the traditional syntactic notions such as part of speech, heads, arguments, voice, etc. can all be accounted for by specific constructions in specific language. In other words, each language has its own construct-i-con.

How is construct-i-con developed? By schematization. Construction grammar maintains that language learning is item-based learning, which means children make generalizations over similarly structured utterances in language input to acquire abstract schemas (Hilpert, 2014, pp. 156-158). Simply put, children learn language by finding patterns from utterances. This process is coined as **schematization**, a process that is "fundamental to human cognition." (Hilpert, 2014, p. 161). Here is an example provided by Hilpert. After encountering expressions such as "my shoe", "my doll", "my key", etc., children acquire a pivot schema, [my + X] (pp. 164-165). A pivot schema has a fixed item ("my" in this case) and an open slot (the X). Then this schema, a mini-construction (as Hilpert calls it), will be stored in children's construct-i-con, and will help them to produce similar utterances on their own.

Studies have shown that this item-based learning process is not only true of the acquisition of pivot schemas, i.e. mini-constructions, but also true of the acquisition of "more abstract constructions, such as argument structure constructions and even complex clausal constructions" (Hilpert, 2014a, p. 177). In plain words, language learning is basically item-based learning of schemas/constructions of all sizes, which therefore includes constructions for sentence-final particles in this current study.

The significance of this central tenet of Construction Grammar for this current study is that: **TO KNOW ABOUT SFPs IS TO KNOW ABOUT THE RELEVANT CONSTRUCTIONS**. I use the word "relevant" because, according to the constructionist view, SFPs are nothing special but some building blocks of some constructions in the Mandarin construct-i-con.

To the best of my knowledge, only one previous study of SFPs has systematically adopted the constructionist approach. Lin (1984), as I have discussed in my literature review, argues that SFP *ne* is to express a flavor of contrast, which can be captured by

the semantic feature [+contrastiveness]. In his analysis, he shows us that the SFP *ne* "quite often is accompanied by *cai* 'really; simply; then and only then' and the semantic property of [+contrastiveness] is greatly enhanced" (p. 224). *Cai* is thus a collocate of *ne* because it "quite often" accompanies *ne*. Put it in the wording of collocation literature, *cai* frequently cooccurs with *ne*. He then observes that *cai* and *ne* constitute a *cai*…*ne* construction (p. 224). This construction can be presented as [*cai* + *ne*], which means *cai* and *ne* fill the two slots in this construction and the two XP slots can be filled by specific structures in specific speech events. His examples include

(35) nei yangzi, **cai** zhen shuai ***ne***
that look **indeed** really handsome **NE**
'That (toss) is really handsome. (Don't you thinkso?)' (p. 224)

(36) Wo **cai** bu yao ***ne***
1SG **indeed** NEG want.to **NE**
'(You think I want to ?) No way.' (p. 225)

His examples of *ne*-related constructions also include *hai (bu/mei)*…*ne*, *(hai) yiwei*…*ne*, Verb *zhe*…*ne*, (zheng)zai…*ne*, … *de ne*, (hai)…Verb *guo*…*ne*, … *le ne* (p. 223).

Lin's study shows us that SFPs can be studied by adopting a constructionist view. SFPs and their collocates constitute constructions. On the other hand, Lin comes to this conclusion based on data collected from "(1) native instructors in Mandarin (2) visiting scholars from China; (3) novels, plays and other texts published in Mainland China and Taiwan after the May Fourth Movement, particularly in the most recent years" (p. 223), not corpus data. Therefore, to use corpus data to conduct constructionist study of SFPs is one important novelty of my dissertation.

Then a working definition of "form" and "function" as the two core components of any construction should be in place, given that I adopt the definition of a construction as a "form-function pairing." [29] "Form" is straightforward. It refers to "morphosyntactic patterns" (Fillmore et al., 1988, p. 534) in which SFPs occur.

[29] Some authors use form-meaning pairing as the definition of construction (e.g. Bybee, 2013). I consistently use form-function pairing as the definition in the current study because SFPs are function words (Lin, 2007, p. 228).

"Function" is difficult to define because this is a rather general and broad concept in linguistics. Consider the fact that a school of thoughts in linguistics is called "functionalism."

In order to narrow down what I mean by "function", let's start again with what previous studies have discussed about SFP functions. Scholars have explained to us that SFPs are notoriously multifunctional. For instance, Xiang (2011) points out that SFPs, and more generally the whole category of pragmatic particles, "accomplish a complex range of interactional functions and convey speakers' nuanced stances vis-a`-vis the propositional content of the utterance, toward the addressee, and other elements of the interactional context" and they "also play a significant role in creating and maintaining textual cohesion, highlighting discourse relationships, facilitating conversational tasks, and, on the more macro-levels, indexing sociocultural identities" (pp. 1377-1378). It suggests that the very much of SFPs' function is about social actions.

"Social action" is a key topic in the field of Conversation Analysis (CA). (Sidnell & Stivers, 2013) Social actions are "what are conversationalists doing when they talk in interaction" (Stivers & Sidnell, 2013, p. 6). Typical examples include "requesting, inviting, granting, complaining, agreeing, telling, noticing, rejecting, and so on" (Schegloff, 2007, p. xiv, cited from Levinson, 2013, p. 104). I borrow this definition of social actions from CA as my working definition of "function" of construction.

Also recall one property of PPs, the big family to which SFPs belong, is "they accompany utterances that have illocutionary forces/acts in particular contexts. Stated in another way, such utterances are produced by speakers to do something for a certain interpersonal effect in specific conversations", as I have summarized based on findings from my mini-corpus of PP studies. This property justifies the adoption of "social action" as the working definition of "function" because **social actions, as "what are conversationalists doing when they talk in interaction", are exactly what linguists in the field of pragmatics call illocutionary forces/acts**.

The second criterion is high frequency. This diagnostic motivates the "usage-based" part of the whole usage-based constructionist approach.

The basic premise of the usage-based theory is that "experience with language creates and impacts the cognitive representations for language" (Bybee, 2013, p. 49). It thus endorses that knowledge of language can be obtained by investigating "experience with language", which can be paraphrased as "all the perception and production processes that are brought to the task of using language", including corpus data which record natural language uses (p.50). This theory and the construction grammar converge because both theories uphold the idea of "what-you-see-is-what-you-get" (Bybee, 2013, p. 51). As discussed previously, construction grammar maintains that a language form on the surface level is directly associated with its function so the seemingly ungrammatical form of the sentence "John sneezed the napkin off the table" can be explained by the function of the English caused-motion construction.

What is central to the usage-based theory is the exemplar representation of human beings' knowledge of language. In other words, human's knowledge of language is cognitively represented by exemplars. An exemplar is a category "formed from tokens of experience that are judged to be the same" (Bybee, 2013, p. 53). One of Bybee's examples of exemplars is that the vowels of *hit*, *swim*, and *sip* can be grouped together as one exemplar, in this case the phoneme [i], because they are similar to each other. Bybee also explains that exemplars vary in size, ranging from a single segment, such as the phoneme [i], to discourse-level linguistic units, such as the Pledge of Allegiance.

There are "strong" exemplars and not-so-strong exemplars (Bybee, 2013, p. 53). Strong exemplars are formed by a high frequency of patterns while not-so-strong ones by relatively low frequency. I borrowed the term "high-frequency exemplars" (namely, strong exemplars in Bybee's words) from Jing-Schmidt (2015) in this dissertation because I think this term offers more clarity. Jing-Schmidt explains that high-frequency exemplars are patterns of utterances in which each pattern has many instances in actual production of language structures. In plain language, such patterns occur a lot in people's words. According to Bybee, high-frequency exemplars are in the central position in peoples' cognitive representation of language and other low-frequency ones bear similarity with the high-frequency ones to some extent.

Jing-Schmidt, following Goldberg (1995), also shows that those high-frequency exemplars are "pathbreakers" (Jing-Schmidt, 2015, p. 8) in terms of language teaching, as language acquisition studies have shown that high-frequency exemplars are acquired first by language users. She therefore proposes that high-frequency exemplars need to be prioritized in language instruction. Her example is that prioritized attention needs to be given to *ba*-construction tokens meaning "change of absolute location", "change of orientation in space", "terminal change of state", and "change of identity or appearance" (pp. 12-13).

In the similar vein, I will prioritize the analysis of "pathbreakers" of SFP-related constructions in this dissertation. After all, they are the center of people's cognitive representation of the knowledge of SFPs. If the pathbreakers are dealt with as the first step, then the major task is completed.

The significance of the usage-based theory for this current study is thus that: **TO KNOW ABOUT SFP-RELATED CONSTRUCTIONS IS TO KNOW ABOUT THE RELEVANT SCHEMAS WHICH ARE ABSTRACTED FROM HIGH-FREQUENCY PATTERNS OF UTTERANCES**.

Up to this point, I have operationalized this current study from the perspective of the usage-based constructionist approach. To solve the "elusiveness" puzzle of SFPs is to find high-frequency schemas/patterns of them, by adopting the usage-based constructionist approach. Details regarding data and methods will be presented in the Methodology chapter.

For the sake of clarity, I adopt Traugott and Trousdale (2013)'s framework when presenting my findings for each SFP. Their framework captures the hierarchical relationships among constructions from the perspective of schematicity. Let's see an example of such a hierarchy of constructions so as to understand Traugott and Trousdale (2013, p. 17), shown in Figure 1:

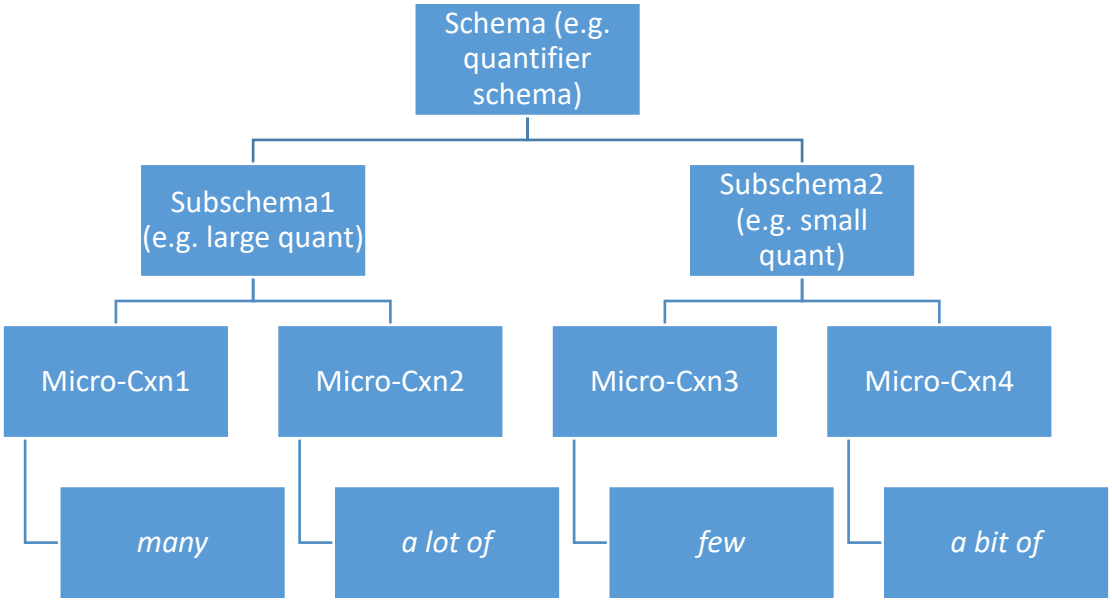


Figure 1: The example of the hierarchy of quantifier constructions, adapted from Traugott and Trousdale (2013)

Ranking highest as the first level is the Schema, which abstracts over specific constructions. The schema in the example is the "quantifier schema". Importantly, Traugott and Trousdale's use of "schema" is different from the "schema" used in the usage-based theory of language acquisition. The two scholars define schemas as "abstract, semantically general groups of constructions"; they are "abstractions across sets of constructions which are (unconsciously) perceived by language-users to be closely related to each other in the constructional network" (p. 14) Their definition shows that a "schema" is meant to be a functionally related group of constructions. However, the "schema" in the usage-based theory is an equivalent of a single construction. This discrepancy again shows that the field of construction grammar is still "highly diverse." Therefore, since I adopt Traugott and Trousdale's framework, I hereby remind readers that the "schema" and "subschema" occurring in later chapters refer to the two scholars' use of "schema", namely a group of constructions that share similar functions.

The second level is Subschemas. The subschemas in this example refer to the "large quant" subschema and the "small quant" one. Then the third level is construction, which is dubbed as Micro-construction (Micro-Cxn) in Traugott and Trousdale's work. I will

use “construction” in this dissertation, as what construction grammar studies do in general. The fourth level is Construct. A construct is a specific instance of constructions. In the example, constructs for the quantifier schema in the example are “many”, “a lot of”, “few”, and “a bit of”, as specific expressions of quantification.

2.3. Summary of the Theoretical Framework

The previous two sections demonstrate that the two theoretical traditions converge. First, both theoretical traditions notice that when speakers make meaning in communication, they usually do not rely on isolated words as units of meaning but what Sinclair coined as extended units of meaning (p. 122), semantic prosody as one example. For construction grammarians, since they believe that the basic unit of any language is construction, what neo-Firthians dub as extended units of meaning are also constructions as form-function pairings. Recall that constructions range from single words such as “Avocado”, “anaconda”, “and” to what Fillmore, Kay & O’Connor (1988) describe as “morphosyntactic patterns” (p. 534) such as the ditransitive or passive structure in English, as is shown in Goldberg’s examples, reproduced below. The more complex constructions such as ditransitive are clearly extended units of meaning because those constructions serve their functions by means of more than one component (at least in English). Ditransitive consists of at least two objects. Passive consists of at least a verb and agent and/or patient. Bybee (2013) calls constructions as “processing units or chunks – sequences of words (or morphemes) that have been used often enough to be accessed together” (p. 51). In other words, neo-Firthians and constructionists both uphold that the basic units of language are “chunks.”

Second, both theoretical traditions are usage-based. Semantic prosody studies identify patterns that are usually inaccessible to introspection (Louw, 1993; Stewart, 2010, cited in Partington, 2017, p. 193) and can only be revealed by corpus data. Semantic prosody is therefore usage-based in that linguists rely on corpus data, which consists of recorded usages of language speakers, to arrive at semantic prosody. The usage-based constructionist approach is usage-based, as I have explained in Section 2.2.

Table 8: Examples of constructions, adapted from Goldberg (2003)

Construction	**Form/Example**	**Function**
Morpheme	e.g. anti-, pre-, -ing	
Word	e.g. Avocado, anaconda, and	
Complex word	e.g. Daredevil, shoo-in	
Idiom (filled)	e.g. Going great guns	
Idiom (partially filled)	e.g. Jog (someone's) memory	
Covariational-Conditional construction	Form: The X-er the Y-er (e.g. The more you think about it, the less you understand)	Meaning: linked independent and dependent variables
Ditransitive (double-object) construction	Form: Subj [V Obj1 Obj2] (e.g. He gave her a Coke; He baked her a muffin)	Meaning: transfer (intended or actual)
Passive	Form: Subj aux VP_{pp} (PP_{by}) (e.g. The armadillo was hit by a car)	Discourse function: to make undergoer topical and/or actor non-topical

Third, both theoretical traditions have noticed the function side of linguistic expressions. As discussed in Section 2.1., Sinclair maintains that semantic prosody of a lexical item is its function in communication, i.e. the reason why speakers use the lexical item. Construction grammarians focus on construction as form-function pairing.

Last but not least, the two theoretical traditions provide mutual support methodologically. On the one hand, semantic prosody theory offers methodological insight into construction identification. As discussed earlier, Sinclair demonstrates that semantic prosody of a lexical item can be arrived at by investigating collocates of a node word. This is actually a specific scenario of constructional study: the semantic prosody (as a **function** expressed by a lexical item in communication) can be identified by studying the collocates of the node word in this lexical item (as a **form**). For example, recall that Sinclair demonstrates that the lexical item visibility + preposition + the + naked + eye expresses a semantic prosody of "difficulty." This [visibility + preposition

+ the + naked + eye] is as a matter of fact a construction because it satisfies the two criteria of construction identification: 1) Non-predicability, because the "difficulty" semantic prosody cannot be arrived at by just focusing on any of the component of this lexical item/construction; 2) High frequency, because Sinclair identified this lexical item/construction based on frequent co-occurring linguistic units with the node word "naked eye."

In other words, what is coined as "lexical item" by Sinclair is actually a construction. Sinclairan analysis of semantic prosody shows that collocation identification is a steppingstone to construction identification. Since construction is the basic unit of language and it is all-encompassing, ranging from morphemes to complex syntactic structures, it is able to incorporate what Sinclair coined as "lexical item" and also a lexical item's semantic prosody.

On the other hand, construction's schematicity presents an efficient way to organize semantic prosody findings of SFPs, which will be elucidated in the Methodology chapter.

CHAPTER III

METHODOLOGY

3.1. Overview of Methodology

This study investigates two new sources of data of SFP-tagged utterances, which had not been used in previous studies: corpus data from the BCC corpus and social media conversation data from WeChat. Systematically collected corpus data can reveal patterns of SFP use. WeChat data enable a more in-depth analysis of SFP use in specific contexts. The two sources of data can thus provide converging evidence. Notice that my WeChat data actually also constitute a mini-corpus, so it is essentially corpus data, too. I call this dataset "WeChat data" to distinguish it from the other dataset that consists of data from the large, general corpus of BCC.

The presentation of findings of both sources of data will follow a "3F Table" quantitative analysis and then a detailed qualitative analysis of examples. "3F" stands for "Function-Form-Frequency". Such a table will provide information of the constructions identified from both sources in the format of form-function pairings in the first two columns, followed in the third column by the token frequency of each construction. I present the Function column first because it can give a clear first-cut description of what the constructions are deployed for in conversations. In the tables, what is in bold font is information of the total number of a form/function or the grand total number of a table.

3.2. Corpus data

Comprehensive and sizable, corpus data facilitate the abstraction of constructions over a plethora of actual language usages. The corpus used in this study is BCC corpus (http://bcc.blcu.edu.cn/zh/cid/2) (Xun et al., 2016). It is a publicly available, free of

charge, 15-billion-word corpus of Mandarin Chinese. It has the following subsections:

- Newspaper (2 billion words)
- Literature (3 billion words)
- Weblog (3 billion words)
- Science and Technology (3 billion words)
- Classical Chinese (2 billion words)
- Multi-field (1 billion words)

The data used in this dissertation was downloaded from the "Multi-field" subsection through February 2019 to August 2020. Ten thousand tokens as a random sample were downloaded respectively for SFP *bei*, non-interrogative *ne* (NI-*ne*), interrogative *ne* (I-*ne*), non-interrogative *a* (NI-*a*), and interrogative *a* (I-*a*). It is "ten thousand" because the corpus only allows registered users to download ten thousand tokens for each search.

Then the raw corpus data from BCC was loaded into and segmented by NLPIR Chinese Lexical Analysis System (Zhang, 2012), which is an open source software package that can automatically segment and tag Mandarin data. To "segment" means to cut a line of Chinese text into individual words. Unlike English, which has spaces between words, no spaces exist between Chinese characters in the written format. Therefore, in order for computer software to process written Chinese, linguists have to cut them into word-by-word lines like English before further analysis. This segmentation process was completed by the NLPIR software in my study. To "tag" means to attach a part-of-speech tag to each segmented word of the Mandarin data. For example, the Mandarin word 照片 *zhaopian* 'photo' will be tagged as 照片/n, "n" standing for "noun", and 睡觉 *shuijiao* as 睡觉/vi, "vi" standing for "intransitive verb". The tagging makes the Mandarin data machine-readable and this paves the way for a KWIC search (explained later), via a free computer corpus software named Antconc (Anthony, 2020), which is perhaps the most useful, free, "ready-made general corpus program" (Gries & Newman, 2013, p. 279).

Note that the "tag" in the preceding paragraph is a corpus linguistics term, which

means to let computers add part-of-speech tags to words in a text. This sense of "tag" is different from the sense I refer to when I say a certain SFP is "tagged to an utterance." This second sense simply means an SFP occurs in the final position of an utterance.

Back to the KWIC search. "KWIC" stands for "key words in context", which is "probably the most widespread corpus-linguistic tool until now" (Gries, 2009, p. 16). KWIC search yields a display of every occurrence of the node word, i.e. the word you do a corpus search for, with a "user-specified context" (this is the "context", shortened for "C" in KWIC). Such context consists of what appears before and after the node word. An example of KWIC search result is presented in Figure 2,

Figure 2: An example of a KWIC search result

Such display of results is also called a concordance. The node word in this concordance is 呗 *bei*, in blue. "User-specified" means linguists can decide how many tagged words occur before and how many after the node word. The sum of the co-occurring texts is the window size. My KWIC searches in this current study are all with a search window size of 100, which means there are 50 tagged words before an SFP and 50 others after it. The academic term for this window size is "span", which is the sum of the number of words on a node word's left side and the number on its right. Therefore, a search window size of 100 is equal to a span of 100. Notice that tagged words include non-character symbols such as numbers and punctuations. This is because the NLPIR software package tags all symbols in a text by default and AntConc

counts every tagged symbol. The choice of a span of 100 is just a matter of preference because I found that with this span, concordances can display enough pretexts and post-texts of an SFP for me as the researcher to understand the meanings of such SFP-tagged utterances.

Also note that in Figure 2 above, preceding the node word *bei* are words displayed in three colors. They are the words appearing in the 1L (the first position on the left of the node word, in red), 2L (the second position on the left of the node word, in green), and 3L (the third position on the left of the node word, in purple). The sentences in this concordance display presented in Figure 2 are sorted alphabetically by the word in the 1L, 2L, and then 3L position. All concordance displays presented later in this dissertation follow this visualization rule.

Each presented concordance display is a screenshot of the original AntConc result page. The displays were captured by the built-in snipping tool on my laptop with Windows 10 operating system, on which AntConc was run. Each display has up to 27 lines in it because the snipping tool can only capture 27 lines at most in one computer screen. Each line contains one occurrence of an SFP and collocates in its 1L, 2L, and 3L position sorted alphabetically starting from 1L to 3L. A display with a number of lines lower than 27 means that a specific collocate has lower than 27 occurrences in the corpus, and that display captures all its occurrences.

I first used AntConc to generate a collocate display table for the position of 1L (viz. the first position to the left of an SFP), 2L, and 3L for each of the three SFPs considered. Such a table contains all collocates in that position, the frequency of each collocate, and the MI value of that collocate. Then each such display table was inserted into Excel for ranking and filtering so that significant collocates with a target SFP can be identified. A "significant" collocate means it has a strong association strength, which means such collocates tend to cooccur with SFPs more frequently than other collocates on the same position. To calculate the association strength, I followed the common practice in corpus linguistics research and adopted mutual information (MI) as the statistic to calculate the association strength of a collocate. As Desagulier (2017, p. 205)

explains, MI is a concept first proposed in information theory. It is calculated by the following formula, shown in Figure 3:

$$I(W_1, W_2) = log_2 \frac{P(W_1, W_2)}{P(W_1)P(W_2)},$$

Figure 3: Desagulier's presentation of Mutual Information (MI) formula

Here, I(W1, W2) is the MI value. P(W1, W2) is the probability of the co-occurrence of two words in a corpus, namely W1 and W2. P(W1)P(W2) is the product of the probabilities of occurrence of each word in the same corpus. A consensus is that significant collocates in Chinese have an MI ⩾3 (Guo & Li, 2016; Li & Guo, 2016). This amounts to say that a collocate with an MI lower than 3 is treated as noise and thus excluded from my further corpus data analysis.

To further screen significant collocates, I used another standard: a minimum frequency of five. This is because a frequency lower than five is considered too low for a collocate to be statistically significant, with MI as the measure, as explained in Kilgrarriff (2005). Accordingly, collocate displays in this dissertation only present collocates that satisfy both standards, namely with an MI value no less than 3 and a frequency no less than 5.

Table 9 shows all significant collocates in the 1L position of *bei*, which serves as an example of collocate display tables that satisfy both aforementioned standards for significant collocates. The total collocate types in 1L position is 306, which means there are 306 different kinds of collocates. Among them, 52 collocate types have a significantly strong association strength (MI ⩾ 3) and a frequency no less than 5. Those significant collocates are shown in the right column in Table 9 and ranked according to their token frequency shown in the left column. The middle column presents the values of the statistic MI. A token means an occurrence of a collocate in the corpus. Put it another way, a token is an instance of (a type of) a collocate.

Table 9: Collocate display of 1L of *bei*

Freq	Stat (MI)	Collocate
675	4.58165	背 *bei* ‘to recite’
192	3.84145	一下 *yixia* ‘one time’
100	3.36083	看看 *kankan* ‘to look a bit’
69	4.06336	试试 *shishi* ‘to try a bit’
34	3.48482	贱 *jian* ‘cheap, ignoble’
28	4.68276	屄 *bi* ‘cunt’
25	4.15669	玩玩 *wanwan* ‘to play a bit’
23	4.56246	抠 *kou* ‘stingy (vernacular)’
22	3.0761	名 *ming* ‘name’
20	4.68276	贫 *pin* ‘garrulous’
20	4.1467	钩 *gou* ‘hook’
20	3.96029	建议 *jianyi* ‘advice’
18	3.42642	玩儿 *wan’er* ‘to play a bit’
17	3.41267	骗 *pian* ‘to cheat’
16	4.0389	意见 *yijian* ‘opinion’
13	4.29573	得了 *dele* ‘enough’
13	3.85963	说说 *shuoshuo* ‘to talk a bit’
13	3.29573	测测 *cece* ‘to test a bit’
12	3.80829	聊聊 *liaoliao* ‘to chat a bit’
12	3.18026	得瑟 *dese* ‘to feel good about oneself’
12	3.13844	追 *zhui* ‘to chase’
12	3.05827	修 *xiu* ‘to fix’
11	4.05473	听听 *tingting* ‘to listen a bit’
11	3.33483	造 *zao* ‘to make, to create’
11	3.18799	脱 *tuo* ‘to take off clothes’
11	3.01291	混 *hun* ‘to muddle through’
10	3.48112	火锅 *huoguo* ‘hotpot’

(continued)

Freq	Stat (MI)	Collocate
9	3.94579	转转 *zhuanzhuan* ‘to walk around a bit’
9	3.04533	熊 *xiong* ‘to be coward’
9	3.04533	再说 *zaishuo* ‘to postpone something’
8	4.36083	膀 *bang* ‘arm’
8	4.22333	拉倒 *ladao* ‘to forget about it’
8	3.22333	想想 *xiangxiang* ‘to think a bit’
7	4.49011	飚 *biao* ‘to show off’
6	4.46036	直说 *zhishuo* ‘to honestly say something’
6	4.26772	慢慢来 *manmanlai* ‘to do something gradually’
6	3.46036	唠嗑 *laoke* ‘to chat (casually)’
6	3.09779	抵制 *dizhi* ‘to boycott’
6	3.01979	懒 *lan* ‘to be lazy’
5	4.68276	掉价 *diaojiao* ‘to have one’s status lowered’
5	4.41972	走着瞧 *zouzheqiao* ‘to wait and see’
5	4.41972	走人 *zouren* ‘to leave some place’
5	4.41972	试试看 *shishi kan* ‘to try a bit’
5	4.41972	梵 *fan* ‘Sancrit’
5	4.41972	唬 *hu* ‘to scare’
5	4.19733	走走 *zouzou* ‘to walk a bit’
5	4.19733	上街 *shangjie* ‘to shop’
5	4.00469	降 *jiang* ‘to drop’
5	3.68276	赚钱 *zhuanqian* ‘to make money’
5	3.68276	瞧瞧 *qiaoqiao* ‘to glimpse a bit’
5	3.54525	离婚 *lihun* ‘to divorce’
5	3.41972	奋斗 *fendou* ‘to endeavor’

All such collocate displays are available in the Appendix section.

The next step is to identify meaningful high-frequency constructions out of the

collocates. The method is to manually check the concordance of every collocate to identify what Stubbs (2001) calls "lexico-grammatical frame" (pp. 91-95). This method conforms to the norm of concordance-based corpus linguistic studies that a concordance display "usually needs a human analyst for interpretation" (Gries & Newman, 2013, p. 277). Overall, I manually checked all concordances for all significant collocates on 1L, 2L, and 3L position for all three target SFPs, in order to identify lexico-grammatical frames.

Stubbs (2001) provides an example of the lexico-grammatical frame of the English word "undergo", which is represented as:

"passive or modal + *undergo* + adjective + abstract noun" (pp. 91-95)

Typical collocates for "passive or modal" include "forced to, required to, must, etc.", which encode a semantic prosody of "involuntary". Typical collocates for "adjective" include "further, extensive, major, severe, etc." Those for "abstract noun" come from the lexical fields of "medical procedure, testing, training, change, trauma, etc." Those adjectives and abstract nouns encode a semantic prosody of "unpleasant." Clearly, [passive or modal + *undergo* + adjective + abstract noun] is a construction because it is a form-function pairing in which the function of the whole construction is not equal to the function of any component or the sum of components' functions.

Simply put and as discussed in Section 2.3., SFP collocations help me to identify SFP constructions. This construction identification method is inspired by the notion of lexical item proposed by Sinclair, as discussed in Chapter 2. In other words, this construction identification process can be skematically presented as follows:

Collocates → Lexico-grammatical frames → Constructions

Let me explain this construction identification process a little bit more by one of the significant collocates of *bei*, 一下 *yixia* 'one time'. It occurs 192 times in the corpus. My glossing for *yixia* and other collocates in previously presented Figure 2 are in Table 10:

Table 10: One example of collocates on the 1L, 2L, and 3L position of *bei*

3L	**2L**	**1L**	**SFP (node word)**
我 *wo* 'first person singular pronoun'	change (a verb; an instance of code-switching)	一下 *yixia*	呗 *bei*
今天 *jintian* 'today'	rock (a verb; an instance of code-switching)	一下 *yixia*	呗 *bei*
就 *jiu* 'immediately'	tomorrow (used as a verb; an instance of code-switching)	一下 *yixia*	呗 *bei*
去 *qu* 'to go to'	争取 *zhengqu* 'to strive'	一下 *yixia*	呗 *bei*
时间 *shijian* 'time'	争取 *zhengqu* 'to strive'	一下 *yixia*	呗 *bei*
啊 *a* 'an SFP'	二 *er* 'being stupid or weird'	一下 *yixia*	呗 *bei*
给 *gei* 'to give'	交涉 *jiaoshe* 'to negotiate'	一下 *yixia*	呗 *bei*
亲 *qin* 'an address term used by sellers to address buyers in e-commerce setting'	亲 *qin* 'to kiss'	一下 *yixia*	呗 *bei*
亲 *qin* 'an address term used by sellers to address buyers in e-commerce setting'	亲 *qin* 'to kiss'	一下 *yixia*	呗 *bei*
亲 *qin* 'an address term used by sellers to address buyers in e-commerce setting'	亲 *qin* 'to kiss'	一下 *yixia*	呗 *bei*
亲 *qin* 'an address term used by sellers to address buyers in e-commerce setting'	亲 *qin* 'to kiss'	一下 *yixia*	呗 *bei*
亲 *qin* 'an address term used by sellers to address buyers in e-commerce setting'	亲 *qin* 'to kiss'	一下 *yixia*	呗 *bei*
亲 *qin* 'an address term used by sellers to address buyers in e-commerce setting'	亲 *qin* 'to kiss'	一下 *yixia*	呗 *bei*

(continued)

3L	**2L**	**1L**	**SFP (node word)**
王子 *wangzi* ‘prince’	亲 *qin* ‘to kiss’	一 下 *yixia*	呗 *bei*
送 *song* ‘to give; to deliver’	人 *ren* ‘human being’	一 下 *yixia*	呗 *bei*
的 *de* ‘a particle, used in associative clause or as an SFP’	介绍 *jieshao* ‘to introduce’	一 下 *yixia*	呗 *bei*
给 *gei* ‘to give’	介绍 *jieshao* ‘to introduce’	一 下 *yixia*	呗 *bei*
水 *shui* ‘water’	休息 *xiuxi* ‘to have a rest’	一 下 *yixia*	呗 *bei*
培训 *peixun* ‘to train’	体验 *tiyan* ‘to experience’	一 下 *yixia*	呗 *bei*
窗 *chuang* ‘window’	关 *guan* ‘to close’	一 下 *yixia*	呗 *bei*
们 *men* ‘a suffix for the plural form of personal pronouns’	关注 *guanzhu* ‘to pay attention’	一 下 *yixia*	呗 *bei*
微薄 *weibo* ‘This is actually a typo. The correct form should be 微博 *weibo* ‘weblog”’	关注 *guanzhu* ‘to pay attention’	一 下 *yixia*	呗 *bei*
们 *men* ‘a suffix for the plural form of personal pronouns’	分享 *fenxiang* ‘to share’	一 下 *yixia*	呗 *bei*
呀 *ya* ‘an allophone of SFP a’	分享 *fenxiang* ‘to share’	一 下 *yixia*	呗 *bei*
呢 *ne* ‘an SFP’	分享 *fenxiang* ‘to share’	一 下 *yixia*	呗 *bei*
啊 *a* ‘an SFP’	分享 *fenxiang* ‘to share’	一 下	呗 *bei*

		yixia	

(continued)

3L	**2L**	**1L**	**SFP (node word)**
图 *tu* ‘picture’	分享 *fenxiang* ‘to share’	一 下 *yixia*	呗 *bei*

Table 10 shows that in the 1L position, all collocates are *yixia*; in the 2L position, 26 out of 27 collocates are verbs. Based on this observation, I endorse that I have found a construction for *bei*: [V + *yixia* + *bei*]. This is an instance of all constructions I found from the corpus data. Put it in another way, I rely on collocates as functional indicators to identify and categorize constructions.

The next step is to organize constructions for a clear presentation. I resorted to the notion of schematicity, as introduced in Section 2.2. Traugott and Trousdale endorse that constructions are on a scale of schematicity (2013). Some constructions are “completely schematic constructions,” for instance, the English ditransitive construction [SUBJ V OBJ1 OBJ2] (Traugott & Trousdale, 2013, p. 14), meaning that such constructions do not have any slots that have been lexically filled. In this case, the slot SUBJ, V, OBJ1, and OBJ2 can theoretically be filled by any English words.

Some constructions are less schematic because they include both specific lexical items and open slots in them. This can be illustrated by examples of construction given by Goldberg (2003) (previously shown in Table 7). For instance, in the construction [jog + someone’s + memory], “someone’s” can be “his”, “her”, “John’s”, etc. and thus this slot is not lexically filled while “jog” and “memory” are two specific words and therefore they fill two slots. Then we can say this construction is less schematic than [SUBJ V OBJ1 OBJ2] because the former construction has more lexically filled slots. By contrast, other constructions are completely not schematic, which means they are fully lexically filled. For example, as Goldberg exemplifies, “avocado”, “Going great guns”, etc.

The same is true of my example, [V + *yixia* + *bei*]. V is a semi-open (or we may

also say semi-filled) slot because it indicates that linguistic units on this position are predominantly verbs, but it does not specify which verbs occur. *Yixia* and *bei* are lexical items that fill the two slots. Therefore, this construction is partially schematic because it has one open slot. All other constructions that I have found from the corpus data are partially schematic because a certain SFP and its collocates as fixed lexical items fill some slots in those constructions. Such constructions can be boiled down to the template "collocate + SFP."

One caveat of this corpus-based approach is that the constructions identified may overlook contextual information associated with constructs. As Traugott and Trousdale (2013) point out:

> "Constructs are empirically attested tokens (e.g. attested *I gave Sarah a book, She needed a lot of energy*), instances of use on a particular occasion, uttered by a particular speaker (or written by a particular writer) with a particular communicative purpose. Constructs are very rich, imbued with a great deal of pragmatic meaning, much of which may be unrecoverable outside of the particular speech event". (p. 16).

A general purpose, large-size, corpus cannot have it for both times. It provides me with significant collocates and thus helps me to find SFP-related constructions. In other words, such constructions reveal what Stubbs' Principle 2 calls "typical and routine in language use." The Principle is reproduced here:

> "Principle 2: Repeated events are significant. The first task of corpus linguistics is to describe what is usual and typical. Unique events certainly occur, but can be described only against the background of what is normal and expected. The frequent occurrence of lexical and grammatical patterns is good evidence of what is typical and routine in language use".

However, such corpus findings cannot at the same time supply specific contextual information for each token of a target structure. That is to say, the corpus data are able to provide me with patterns, but it cannot simultaneously provide me with contextualized details for a qualitative analysis. This is where WeChat data come in.

3.3. WeChat Data

Lacking information of turn-by-turn evolvement of conversations, corpus decontextualizes each use of a node word (Gries & Newman, 2013, p. 275). "Decontextualize" means it is impossible to know the specific context of each occurrence of a word. For example, we cannot pin down the exact speaker-hearer identity of a conversation recorded in a corpus. The BCC corpus indeed has a "context" function by which users can trace the larger stretch of text where an SFP token occurs, but the exact speaker-hearer identity is often not clear. The reason is that the BCC corpus only compiles textual data, without providing demographic data of interactants. We cannot either know the exact activities the speakers are engaged in.

That is to say, in terms of Construction Grammar, my corpus data can reveal constructions, but cannot be used to investigate constructs. A "construct" is an instance of real usage of a construction, as you have seen in Traugott and Trousdale's illustration of the quantifier schema in Figure 1. Whereas constructions are abstract and overall but at the same times lack particular information, constructs are what speakers produce in real communications and thus are imbued with nuts and bolts of speech events. This suggests that we should also take into account "the particular speech event" of SFPs reflected in constructs so that we can study the pragmatics of them.

Recall that pragmatic particles (PPs), to which SFPs belong, have the following four characteristics:

- They predominantly occur in dialogues in spoken language.
- They do not change or contribute to the propositional content of an utterance they accompany.
- They accompany utterances that have illocutionary forces/acts in particular contexts. Stated in another way, such utterances are produced by speakers to do something for a certain interpersonal effect in specific conversations.
- They are speaker-initiated (as Deng (2015) coins it) and they accompany utterances that are hearer-oriented.

WeChat data can make significant contributions to my current study as another source

on PPs. First, the WeChat data consist of casual dialogues among users, where PPs likely occur. Second, the WeChat data let me know the pretext and post-text of a dialogue and therefore shed important lights on the illocutionary force of an SFP-tagged utterance. Third, the WeChat data, consisting of screenshots of dialogues, give clear clues on the speaker-hearer identity. It is clear which turn is produced by which speaker in a WeChat dialogue, as is the same with other instant communication tools.

WeChat data has other advantages in terms of data collection. First, WeChat is the most widely used social media application in China (Sandel et al., 2019, p. 228). WeChat "is a multi-modal online application that allows users to send short text messages, pictures, emoji, graphical interchange formats (GIFS), audio files, and/or some combination of each." (Sandel et al., 2019, p. 228) Thus, there is reason to believe that WeChat data can generally represent the everyday usage of Mandarin Chinese in casual situations. Second, it is digitally mediated so that it is convenient to collect the data because participants can just text to me their screenshots of their WeChat interaction. Third, similar to other social media platforms, in daily life Chinese people use spoken language to compose WeChat messages, although they are actually messages in the written format. Stated in another way, WeChat messages are "spoken language written down." Therefore, SFPs are likely to occur in such messages, as SFPs mainly occur in spoken language, according to prior studies.

I collected 532 WeChat screenshots of episodes that contain target SFPs. An "episode" is defined as a WeChat conversation revolving around a particular topic, following the definition adopted in Lee (2018). Also following Lee's research, a "turn" is defined as a WeChat message and a "sequence" is defined as "two turns or actions in which the first action performed by one speaker invites a particular type of second action to be performed by another speaker, for example, a question in the first turn invites an answer in the next" (p. 155).

During the data collection, the screenshots were sent to me via WeChat. In the first place, I posted two recruitment messages in my WeChat "moments" to broadcast the messages to all of my WeChat friends to voluntarily send me screenshots, for a

financial reward of 0.5 RMB for each screenshot, which was as a matter of fact declined by most contributors. My "WeChat friends" include friends, relatives, teachers, and other acquaintances. After the screenshots were collected, all identifying information on the individuals was redacted to anonymize the data which included person names, address terms that may disclose private information, location names, etc. This part of the research was approved by the University of Oregon Research Compliance Office.

Each screenshot was then renamed by the SFP-tagged utterance in the episode captured by the screenshot. All the new file names were compiled into a txt file via a .bat file. Used on my laptop with a Windows operating system, the content of the .bat file for each target SFP is "dir *.* /b > Bei_WeChat.txt", "dir *.* /b > I_Ne_WeChat.txt" (for screenshots having interrogative *ne* tokens), "dir *.* /b > NI_Ne_WeChat.txt" (for non-interrogative *ne* tokens), "dir *.* /b > I_A_WeChat.txt" (for interrogative *a* tokens), and "dir *.* /b > NI_A_WeChat.txt" (for non-interrogative *a* tokens). Each .bat file was placed in the folder where I stored the screenshots. Upon double clicking the .bat file, a .txt file was generated which contained all the file names of the screenshots in that folder. Since the file names are basically SFP-tagged utterances found in the WeChat data, each .txt file then contained all the utterances for each target SFP.

A new .txt file is in this sense a mini-corpus in that it has all tokens of a certain SFP. The data in the .txt file was then loaded into an Excel spreadsheet, where all SFP episodes were then coded in a case-by-variable format, as is exemplified in Table 11.

The first column consists of the raw "WeChat utterances" of a target SFP. The second column is the variable "Pre-SFP sentence type", namely the sentence type of the structure preceding a certain SFP, such as declarative or directive. The third column is the variable "Form" in a constructionist sense. The fourth column is the variable "function." All the coding was done and double-checked by me as the researcher. Once the coding was done, I used the Pivot Table function in Excel to generate statistical results, which are available in the next three chapters.

Table 11: Examples of case-by-variable coding of WeChat data

WeChat utterances	Pre-SFP sentence type	Form	Function
也不能入境呗 *ye buneng rujing bei* 'Just still cannot enter this country'	Declarative	[VP + *bei*]	Response
傻逼呗 *shabi bei* 'Just a stupid cunt'	Declarative	[NP + *bei*]	Response
国内经济形势不好呗 *guonei jingji xingshi buhao bei* 'The domestic economic situation is just not good'	Declarative	[VP + *bei*]	Response

I focus on sentence type in construction identification of my WeChat data because this (mini-)corpus made up of WeChat data cannot provide me with collocational profile for each target SFP, due to its small size. Instead, I use sentence types to obtain "sentence type + SFP" style constructions because I can manually find the sentence type of each SFP-tagged utterance in this dataset. By categorizing SFP-tagged tokens according to the sentence types of pre-SFP linguistic units, I can more easily identify patterns under each sentence type. Without such categorization, my scrutiny would be overwhelmed by individual tokens. At the same time, upholding a usage-based constructionist approach as is laid out in Chapter 2, I do not assume there is a boundary between the sentence type and the SFPs. Pre-SFP linguistic units and SFPs as a whole constitute constructions.

I follow Zhan and Bai (2016) in recognizing five major Mandarin sentence types in Mandarin Chinese:

- Declarative
- Exclamative

- Interrogative
- Directive
- Vocative

I did both a quantitative and a qualitative analysis based on the WeChat data. The Excel Pivot table results provided the distribution patterns of each target SFP-tagged utterances, as a quantitative analysis. Such results are organized according to the sentence types of the structure preceding an SFP, namely "Pre-SFP sentence type", which are available to me in the manual coding process. I did not organize the WeChat data analysis findings according to schemas because it is impossible to obtain collocates from the WeChat data that are adequately representative as I can do from the corpus data, due to the much smaller size of WeChat data.

Then for the high frequency exemplars in the WeChat data, i.e. "pathbreakers", of each target SFP, I did a qualitative analysis by providing a turn-by-turn explanation of one exemplar WeChat conversation captured in screenshot(s). In a nutshell, the significance of WeChat data lies in its contribution to the qualitative analysis.

3.4. Romanization

Pinyin[30] is used to Romanize the Mandarin data. The pinyin system is made up of a Roman letter part and a tone marking part (because Mandarin is a tone language). I just maintain the Roman letter part in my romanization because tones are not focused in my semantic and pragmatic analysis. This practice has been carried out in representative studies of SFPs such as Wu (2004) and Yang and Wiltschko (2016). All pinyin symbols in example sentences remain vertical and those in the running texts are italicized.

I also present Chinese characters for the following categories of data in order to maintain the originality of my findings:

- Mandarin Chinese dictionary entries.
- Quotations from works originally published in Chinese.

[30] Wikipedia. (April 21th, 2020). *Pinyin*. https://en.wikipedia.org/wiki/Pinyin.

- Expressions extracted from raw corpus concordances.
- Expressions extracted from raw WeChat screenshots.

3.5. Glossing and Translation

I use both broad and narrow rules for glossing and translation. The broad rules mainly serve to get the main message of an example to readers, leaving out unessential details. The narrow rules will be used if a fine-grained analysis is needed.

- Broad rules

Only free English translations are provided for non-English examples, raw corpus findings, and sentences extracted from previous publications, when details of those expressions do not matter.

- Narrow rules:

Non-English words and expressions cited in the running text for the first time will be displayed in three consecutive layers: Mandarin Chinese characters, italicized pinyin, and free English translation within single quotation marks. If they are cited again in the running text, only italicized pinyin will be used.

Examples that are separately presented from the running text will be transcribed according to the Leipzig Glossing Rules[31]. The glossing will be in three lines. The first line is the Romanization of an example clause or sentence in the object language (in this case, Mandarin Chinese). The second line is a morpheme-by-morpheme gloss of the first line. The third line is a free English translation of the example.

[31] Department of Linguistics at Max Planck Institute for Evolutionary Anthropology. (April 21th, 2020). *Leipzig Glossing Rules*. https://www.eva.mpg.de/lingua/resources/glossing-rules.php.

CHAPTER IV

RESULTS FOR 呗 *BEI*

4.1. Constructions of *bei* in Corpus Data

Several constructions emerged from the corpus data for *bei*. They are presented in Table 12.

The HORTATORY REQUESTING Schema

The first is what I call HORTATORY REQUESTING schema, paired with an directive sentence structure that ends in the SFP *bei*. Under this are two subschemas: Subschema 1 "Requesting an Action to Be Taken Which does not Require Much Effort" and Subschema 2 "Requesting Something in Small Quantity." The first category of constructions under Subschema 1 contains the collocate 一下 *yixia* 'one time.' One is [V + (O) + 一下 *yixia* 'one time' + *bei*], "V" standing for "verb" and "O" standing for "object". The verb can be an intransitive verb, such as in the structure 休息一下呗 *xiuxi yixia bei* 'Just have a rest' or a transitive verb with its object omitted, as in 分享一下呗 *fenxiang yixia bei* 'Just share it with me'. The other construction under this subschema is [VO + *yixia* + *bei*] in which the verb co-occurs with its object, as in 展望一下 2012 呗 *zhanwang yixia 2012 bei* 'Just predict what will happen in 2012', in which *zhanwang* 'to predict' is the verb and *2012* is the object.

The morpheme 一下 *yixia* consists of the numeral *yi* 'one' and the verbal classifier *xia* that that quantifies an action. The explanation offered by The DMC Dictionary for *yixia* is "用在动词后面，表示做一次或试着做" (used after a verb to mean doing something for one time or trying tentatively to do something). One can argue that the numeral is not understood literally, but that together *yixia* conveys tentativeness. On the other hand, verbs occurring in the constructions include 亲 'to

kiss', 分享 'to share something', 转(发) 'to forward (a message, an email, etc.)', etc., which describe some actions that the speaker hopes the hearer can take so that the speaker can be the beneficiary. Those verbs mean a request and *yixia* mitigates the efforts needed to meet a request, as is indicated by the dictionary explanation. By using *yixia*, the speaker is connoting "just do something for one time as a trial; it is not a big deal"; the speaker is conveying a tone of being tentative and non-imposing by emphasizing the low workload of possible imminent action required: the action is just a one-time thing. *yixia* in this mitigation sense is a "diminutive" (Lee-Wong, 1998, p. 394).

The second category is the verbal reduplication construction. Reduplicated verbs found from the corpus data include the following items as is shown in Table 13. In Mandarin Chinese, reduplication of volitional verb signals "a delimitative aspect, *doing something a little bit* (italicization in the original text)" (Sun, 2006, p. 92). Li and Thompson (Li & Thompson, 1981) also use the term delimitative aspect to characterize its function (pp. 232-236). Li and Thompson further specify that this aspect is "particularly likely to occur in requests" (p. 235) and explain that:

> "When one wishes to soften a request so that it will not appear harsh, the delimitative aspect is a perfect device to use, since it reduces the 'weight' of the request on the hearer by saying that the action can be done 'just a little'". (p. 236)

In my case, as is shown in Table 13, all the verbs are reduplicated to mean "to do something a bit", for example, 看看 *kankan* means "to look a bit." The last example in the column, 试试看 *shishi kan*, is another form of verbal reduplication, i.e., a verb is reduplicated, that is *shi* is reduplicated as *shishi*, and is then tagged with a clitic 看 *kan* (here *kan* is a clitic and thus different from the content word *kan* 'to look', which has the 4th lexical tone, although they share the same Chinese character). Sun (2006) points out that the clitic 看 *kan* "functions to emphasize the testing nature associated with the delimitative aspect" (p. 92).

Table 12: Constructions in corpus data of *bei*

Function	Form	Frequency
HORTATORY REQUESTING	**[Directive + *bei*]**	**518**
Subschema 1 "Requesting an Action to Be Taken Which does not Require Much Effort"		
	• [V + (O) + 一下 *yixia* 'one time' + *bei*]	192
	• [V-V + *bei*]	275
Subschema 2 "Requesting Something in Small Quantity"		
	[V + CL + N + *bei*]	51
	• [V + 个 *ge* 'one thing, classifier' + 名 *ming* 'name' + bei]	15
BALD ASSERTION	• [给 *gei* 'to give' + 个/点(儿) *ge/dian(er)* 'one thing, classifier/a bit, classifier' + 建议 *jianyi* 'advice/suggestion'/意见 *yijian* 'opinion' + *bei*] **[Declarative + *bei*]**	36 **840**
	• [Single Word+ *bei*]	840
Grand Total		**1523**

Table 13: Frequency of reduplicated verbs occurring before *bei* in the construction [V-V + *bei*]

Reduplicated Verb	Frequency
看看 *kankan* 'to look a bit'	100
试试 *shishi* 'to try a bit'	69
玩玩 *wanwan* 'to play a bit'	25
说说 *shuoshuo* 'to talk a bit'	13
Reduplicated Verb	Frequency
测测 *cece* 'to test a bit'	13
聊聊 *liaoliao* 'to chat a bit'	12
听听 *tingting* 'to listen a bit'	11
转转 *zhuanzhuan* 'to walk around a bit'	9
想想 *xiangxiang* 'to think a bit'	8
走走 *zouzou* 'to walk a bit'	5
瞧瞧 *qiaoqiao* 'to look a bit'	5
试试看 *shishi kan* 'to try a bit'	5
Total	**275**

For all the verbal reduplication tokens, the speaker is requesting the hearer to try something and at the same time is implying that such a trial does not need too many efforts at all. These verbal reduplications functioning as mitigators are called "syntactic downgraders" by Lee-Wong (1998, p. 392). One such utterance is presented as Example (37),

(37) 你 看看 呗
ni kan-kan bei
you look-look BEI
'You can just take a look. (It's not a big deal.)'

The category of constructions under Subschema 2 is [V + CL + N + *bei*], CL standing for "classifier" and N standing for "noun". One construction within this subschema is [V + 个 *ge* 'one thing, classifier' + 名 *ming* 'name' + *bei*]. The speaker of all such utterances is requesting the hearer to do something about the hearer's name, such as 注个名 *zhugeming* 'to register' and 签个名 *qiangeming* 'to sign one's name'. As is shown Example (38),

(38) 给 我 签 个 名 呗
gei wo qian ge ming bei

give me sign CL name BEI

'Just give me a signature.'

Another construction within this subschema is [给 *gei* 'to give' + 个/点(儿) *ge/dian(er)* 'one thing, classifier/a bit, classifier' + 建议 *jianyi* 'advice/suggestion'/意见 *yijian* 'opinion' + *bei*]. 点 *dian* 'a bit' is a classifier and 点儿 *dianer* is the classifier followed by a diminutive marker 儿 *er*. 建议 *jianyi* 'advice/suggestion' and 意见 *yijian* 'opinion' are two synonymous collocates. *Jianyi* appears in the construction [*gei* + *ge/dain(er)* + *jianyi* + *bei*]. The speaker uses this construction to request some advice from the hearer. Likewise, *yijian* predominantly appears in the construction [*gei* + *ge/dian* + *yijian* + *bei*]. The speaker uses this construction to request the hearer's opinion. One such use is shown as Example (39),

(39) 大家 给 点 建议 呗

dajia gei dian jianyi bei

everyone give bit advice BEI

'Can anyone just give me a piece of advice?'

The BALD ASSERTION Schema

Schema 2 is what I call BALD ASSERTION schema, which is paired with a short declarative utterance ending in *bei* that is used as an evaluative remark or even insult. The utterance is typically concise and curt, as is shown in Example (40), which is extracted from the corpus data,

(40) 什么 是 温柔？ 贱 呗！

Shenme shi wenrou? Jian bei!

what be tender cheap BEI

'—What is tenderness? – That is just being cheap!'

Let's take a look at the collocates that fill in the slot of the construction under this schema, which are shown in Table 14,

Table 14: Collocates filling in the slot in [Single Word + *bei*]

Collocate	Frequency
贱 *jian* 'cheap, ignoble'	34
屄 *bi* 'cunt (vulgar)	28
抠 *kou* 'stingy (vernacular)'	23
贫 *pin* 'garrulous'	20
骗 *pian* 'to cheat'	17
得瑟 *dese* 'to be cocky'	12
混 *hun* 'to muddle along'	11
熊 *xiong* 'to be coward'	9
懒 *lan* 'lazy'	6
掉价 *diaojia* 'to have one's status lowered'	5
Total	**165**

The most frequent one is 贱 *jian* 'cheap, ignoble'. The second is 屄 *bi* 'cunt (vulgar)'[32]. It turns out that in this case *bei* is collocated with the phrase 晒傻屄 *shai shabi* 'to make a stupid cunt bask in sunshine'. *Shai*, here as a new coinage used in

[32] My explanation of this collocate is based on this website: https://chinese.yabla.com/chinese-english-pinyin-dictionary.php?define=%E5%B1%84 , September 3rd, 2019.

Cyber China community, means "to display or show off something on the Internet". *Shabi* means "Stupid cunt, or describes something as being very stupid. Often shortened to 'SB', 'sha bi', 傻 B *sha B* 'stupid B', 2B (2 = S), or 傻 *sha* 'stupid' + any Chinese character with the 'bi' sound"[33].

These collocates fall into a narrow semantic category of disparaging evaluations verging on insult. Each of them together with *bei* constitute a construction in the form of [Single Word + *bei*]. Its function is what I follow Chappell (1991) to call "bald assertion", as she uses this term when she describes one function of *me* as "marker of disagreement in face of a self-evident situation" (p. 52). One example of such use has been presented as Example (15), which is reproduced below.

B: Xianzai shou zhei zhong chuguochao yingxiang de ren tai duo le
now suffer this kind go:abroad:trend influence DE people too many LE
'There are far too many people being influenced by the trend to go abroad'

C: Zhe ye shi hao shi me!
this also be good matter ME
'That's something good too!' (adapted from Chappell, 1991, pp. 54-55)

B and C were talking about "Chinese Education System". B thought a lot of Chinese people were influenced by the trend of studying overseas. C replied with a me-tagged utterance and showed his or her attitude that this was actually a good thing. C's assumption was that many people studying abroad should be a self-evident good thing and he or she uses *me* to indicate disagreement and impatience with B in questioning the positiveness of this.

In this example, Speaker C offers "offers no reason or elaboration as to why he disagrees with B in viewing overseas study as something beneficial" (p. 54). That is why Chappell thinks this assertion of Speaker C is bald. This is also true of this schema of bei, where such utterances are used in response with no explanation on the part of speaker why he or she thinks so. Additionally, unlike *me* which does not convey an import of insulting, such bei utterances are bald also because they are bluntly insulting

[33] Quora. (September 3rd, 2019). *What does the Chinese word 傻逼 mean in English?* https://www.quora.com/What-does-the-Chinese-word-%E5%82%BB%E9%80%BC-mean-in-English.

with a tone of contempt.

Interestingly, the top 1 collocate in the 1L position is 背 *bei* 'to recite', as is pointed out by a member of this dissertation committee. The construction [*bei* 'to recite' + *bei*] collocate can still be categorized under the schema of Bald Assertion as it conforms to the form of schema, i.e. [Single Word + *bei*], although it is not a negatively evaluative term. I manually checked the raw data and confirm that the assertation is bald because no qualifications exist in specific contexts.

To sum up at this point, the corpus data suggest that bei occurs in constructions in the form of [Directive + *bei*], which function as hortatory requesting and in construction in the form of [Declarative + *bei*] functioning as bald assertion.

4.2. Constructions of *bei* in WeChat Data

Altogether, screenshots for 179 tokens of *bei* were collected. Those *bei*-tagged utterances perform three functions in discourse: Response, Request, and Suggestion. The three functions are lined up well with the two constructional schemas already found in the corpus data. The HORTATORY REQUESTING schema and the Request as well as the Suggestion function can be categorized as what Givón (2001) coined "manipulative speech acts" (more generally, non-declarative speech acts) while the Response function and the BALD ASSERTION schema as "declarative speech acts."

Declarative *Bei*-tagged Utterances

48 Response tokens were identified in total, as is shown in Table 15.

Table 15: Distribution of declarative *bei*-tagged utterances in WeChat data

Function	Form	Frequency
Response	**[Declarative + *bei*]**	**48**
	• [VP + *bei*]	38
	• [VP + *jiu* + VP + *bei*]	5
	• [NP + *bei*]	5

Table 15 shows that in the WeChat data of response tokens, *bei* is predominantly tagged with Mandarin declarative sentences. 38 out of 48 utterances are constructs of the construction [VP + bei], as is shown in Figure 4,

	A: Yo (an exclamation), your field is also impacted by the trade war?
B: The domestic economic situation is just not favorable.	
	A: [an emoji meaning "I see"]

Figure 4: *guonei jingji xingshi buhao bei* 'The domestic economic situation is just not favorable.'

A: 你们 也 受 贸易战 影响？

nimen ye shou maoyizhan yingxiang

2PL also receive trade.war impact

'Is your industry also impacted by the trade war?'

B: 国内 经济 形势 不 好 呗

guonei jingji xingshi bu hao bei

domestically economy situation not good BEI

'The domestic economic situation is just not favorable.'

In this conversation, Speaker A and Speaker B were talking about the influence of the trade war between USA and China on Speaker B's field. B responded to A's question with the bei-tagged utterance to imply that since the trade war rendered the domestic

economic situation unfavorable, all fields, including B's own field, was negatively impacted. In this utterance, *bu hao* 'not good' is the verb phrase.

Five out of forty-eight are tagged with a sentence in the idiosyncratic schematic construction [VP + *jiu* + VP + *bei*]. The five constructs of [VP + *jiu* + VP + *bei*] are as follows:

他 说 加 就 加 呗

ta shuo jia jiu jia bei

he say add adverb add BEI

'(If) he said to add, (then) add. (It's fine.)'

感染 就 感染 呗

ganran jiu ganran bei

infected adverb infected BEI

'(If) infected, (then) infected. (It's fine.)'

不 去 就 不 去 呗

bu qu jiu bu qu bei

NEG go adverb NEG go BEI

'(If) you don't want to go, (then) you don't go. (It's fine.)'

人 多 就 人 多 呗

ren duo jiu ren duo bei

human many adverb human many BEI

'(If) there are many people, (then) there are many people. (It's fine.)'

脏 就 脏 呗

zang jiu zang bei

dirty adverb dirty BEI

''(If) it is dirty, (then) it is dirty. (It's fine.)''

For the five constructs, structurally speaking, identical or near identical structures appear before and after the adverb *jiu*. This adverb is polysemous. The DMC Dictionary's explanation for *jiu* for its use in such a construction is 放在两个相同的成分之间，表示容忍 'To be put between two identical components to mean tolerance.'

As an example, let's see the conversation for the first construct, 他说加就加呗 'If he said to add him as a co-author, then I'll do that. It's fine.':

A: 你 要 给 X 加 这 个 author 吗?

ni yao gei X jia zhe ge author ma

you want give X add this classifier author MA

'Will you include X as a co-author?'

B: Since he said that, I just include him as a co-author.

B: 肯定 加 了 不然 还 能 怎样

kending jia le buran hai neng zenyang

certainly add LE otherwise other can how

'Of course. I have no other choices.'

B: **他 说 加 就 加 呗**

ta shuo jia jiu jia bei

he say add adverb add BEI

'If he said to add him as a co-author, then I'll do that. It's fine.'

Speaker A asked Speaker B whether B was going to add X as a co-author. Speaker B said that he or she certainly would do that as the only choice available. Then Speaker B produced the *ta shuo jia jiu jia bei* token to underscore the fact that he or she was willing to do this. In the pre-SFP structure, the verb *jia* 'to add' appears both in the preceding position and following position of the adverb *jiu*, so that this is an instance of the construction [VP + *jiu* + VP + *bei*].

In addition, five out of forty-eight constructs are of the construction [NP + bei], which means bei follows a noun phrase. This form is very similar to the form of the [Single Word + *bei*] construction identified from the corpus data. One instance of [NP + bei] construction is shown in Figure 5,

Figure 5: *wenqing bei* 'I think it just means I'm one of those young people who love arts and humanities.'

A: (I) feel (you) have the quality of being a scholar in the field of A&H	
	B: I like math, computer science, etc.
A: [emoji meaning "wow"]	
	B: Speaking of the quality of being a scholar in the field of arts and humanities [an emoji meaning "to cry and laugh at the same time"] **I think it just means I'm one of those young people who love arts and humanities.**
A: You can do interdisciplinary research in the future then.	
	B: [two emojis meaning "to laugh"]

In this conversation, Speaker A and Speaker B were talking about their choice of STEM or Arts and Humanities (A&H) as their major. A said that s/he felt that Speaker B had the quality of being a scholar in the field of A&H. This utterance served as a compliment. Speaker B then adopted a deflection strategy as the response to compliment by saying 文青呗 *wenqing bei* 'I think it just means I'm one of those young people who love arts and humanities.' Here *wenqing* is the noun phrase, which means "young people who love arts and humanities" and we can see that it is directly followed by the SFP *bei*.

This conversation provides another example of the bald assertion function of *bei*-tagged utterances. It shows that Speaker B's real interest lies in math and computer science, but perhaps because of some reasons s/he is majoring in A&H. Although A complimented that Speaker B possessed the quality of being a scholar in A&H, Speaker B is probably not entirely happy with this reality since, after all, B is not majoring in his or her preferred disciplines. That explains why B just used an NP (*wenqing*)

followed by *bei* in his or her response, without offering further explanations or elaborations.

Textually speaking, for the 48 *bei*-tagged Response tokens overall, five types of immediately subsequent tokens were found. 98% of such subsequent tokens perform the function of terminating an ongoing sequence. Their distribution is presented in Table 16:

Table 16: Types of immediately subsequent tokens for *bei*-tagged Response tokens

Types of Immediately Subsequent Tokens for *Bei*-Tagged Response Tokens	**Terminating Function**	**Frequency**	**Percentage**
Terminal Explanation by The Same Speaker	Yes	20	42%
Initiation of New Sequence by Another Speaker or The Same Speaker	Yes	15	31%
Terminal Confirmation by Another Speaker	Yes	8	17%
Terminal Silence by All	Yes	4	8%
Disagreement by Another Speaker	No	1	2%
Total		48	100%

I will define the five types of immediately subsequent tokens in the following part by providing examples.

By "Terminal Explanation by The Same Speaker," I mean in such cases a speaker first produces a *bei*-tagged utterance and then another utterance, explanation in nature, which further clarifies the speaker's overtone that he or she is not willing to continue the current sequence. That is why I term it as "terminal". Textually speaking, such terminal explanation wraps up the current speaker's turn. As is shown in Figure 6 and Figure 7:

Figure 6: Pretext of *buqu jiu buqu bei* 'If you don't want to go, it's fine.'

Figure 7: *buqu jiu buqu bei* 'If you don't want to go, it's fine.'

Figure 14

	A: I saw someone posted it. I don't know whether it is reliable or not. Do you want to have a try?
B: Have a try of what?	
	A: That interview.
B: Why do you mention this to me?	
	A: I just saw this opportunity.
B: Why do I need to go?	
	A: [An emoji meaning "byebye"]

Figure 15

	A: If you don't want to go, it's fine.
	A: It seems you are very mad.
B: Because I don't see any point in it.	
	A: I see but I don't know what activity you are engaged in right now.
B: I'm drawing pictures.	

In this example, Speaker A was mentioning an interview opportunity to Speaker B. After Speaker A produced the *bei*-tagged utterance 不去就不去呗 *buqu jiu buqu bei* 'If you don't want to go, it's fine', and then A offered a what I call "terminal explanation" by saying 火气好大的样子 *huoqi haoda de yangzi* 'It seems you are very mad'. I think this wrap-up explanation's function is twofold. The first function is that A clarifies that A thinks B should not have been mad at A. Clearly A is doing a favor to B by breaking this interview opportunity information to B, who is supposed to

show gratitude. The second function is that A's terminal explanation demonstrates her purpose to wrap up the sequence initiated by B's question 我为什么要去 *wo weishenme yao qu* 'Why do I need to go?'. A is demonstrating her disappointment with B's outrageous reply and such disappointment makes A want to end the sequence; or to a great extent, it makes A almost speechless.

By "Initiation of New Sequence by Another Speaker or The Same Speaker", I mean in such cases a *bei*-tagged utterance is at the end of the current sequence. What follows up is a new sequence initiated by the current speaker or another speaker. Figure 8 depicts such a scenario where a new sequence is initiated by the same speaker:

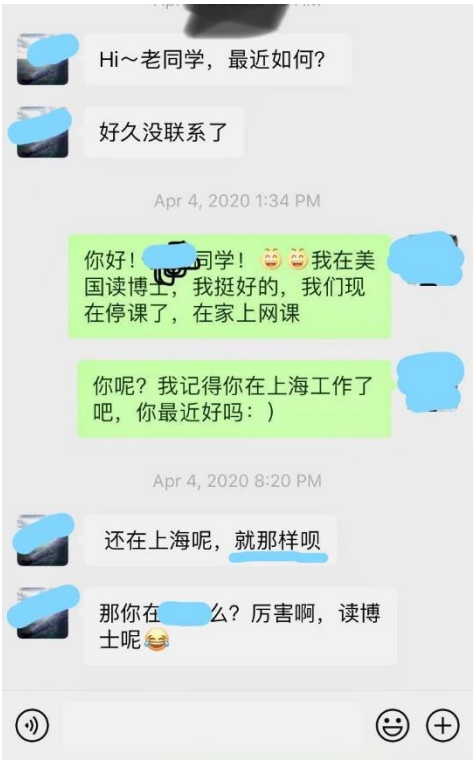


Figure 8: *jiu nayang bei* 'I'm just doing OK.'

B: Hi, my dear classmate. How are you recently?	
B: We haven't been in touch for quite a long time.	
	A: Hi! [emoji meaning "smile"] I'm currently studying for my PhD in the states. I'm doing OK. We are currently doing remote learning because of the COVID-19.
	A: How about you? I remember that you are working in Shanghai. How are you? :)
B: I'm still at Shanghai. **I'm just doing OK.**	
B: Are you living in X? Awesome, you are studying in a PhD program now.	

In this example, Speaker A and B were greeting each other when chatting after a long while. After the very first several ice-breaking turns, A initiated a new sequence by asking about the current status of B: 你呢？我记得你在上海工作了吧，你最近好吗:) *ni ne? wo jide ni zai shanghai gongzuo le ba, ni zuijin hao ma* 'How about you? I remember that you are working in Shanghai. How are you? :)' Then

B replied with the *bei*-tagged utterance 还在上海呢，就那样呗. *hai zai shanghai ne, jiu nayang bei* 'Still in Shanghai, nothing special' This utterance marks the end of the current sequence because B then initiated a new sequence by checking with A whether A was living in X or not. This initiating turn also includes B's compliment of A's admission to a PhD program 厉害啊，读博士呢 *lihai a, du boshi ne* 'Awesome, you are studying in a PhD program now'.

The above example is produced in a situation where a new sequence is initiated by the same speaker after a *bei*-tagged utterance. Such tokens actually take up only a small part of the "New Sequence" type (3 out of 15). The rest 12 tokens are produced when a new sequence is initiated by another speaker. Let's see an example of this kind in Figure 9, which has been discussed earlier for the [XP + *jiu* + XP + *bei*] construction.

A: Will you include X as a co-author?	
	B: Of course. I have no other choices.
	B: If he said to add him as a co-author, then I'll do that. It's fine.
A: Is this paper about to appear?	
	B: Far from it.
	B: major revision
A: [An emoji indicating receipt of the message]	
A: I see.	
A: Perhaps he thinks the paper is going to be accepted.	

Figure 9: *ta shuo jia jiu jia bei* 'If he said to add him as a co-author, then I'll do that. It's fine.'

A: 你　要　给 X　加 这　个　　author 吗？

ni　yao　gei X　jia zhe　ge　　author　ma

you　want give X　add this classifier author　MA

‘Will you include X as a co-author?’

B: Since he said that, I just include him as a co-author.

B: 肯定　　加 了 不然　　还　　能　怎样

kending　jia le　buran　　hai　　neng zenyang

certainly add LE otherwise　other　can　how

‘Of course. I have no other choices.’

B: **他　　说 加　就　　加 呗**

ta　shuo　jia　jiu　　jia　bei

he　say　add　adverb　add　BEI

‘If he said to add him as a co-author, then I’ll do that. It’s fine.’

A: 已经　　要　发　　了？

yijing　yao　fa　　le

already　want　publish　LE

‘Is it already going to be published?’

Speaker A initiated the Q & A sequence by asking whether Speaker B was going to list X as a co-author of a paper. B answered that she definitely would and then said 他说加就加呗 ‘If he said to add him as a co-author, then I’ll do that. It’s fine.’ After this *bei*-tagged token, A initiated a new sequence by asking B whether this paper was going to be published.

By “Terminal Confirmation by Another Speaker”, I mean each such *bei*-tagged utterance precedes a confirmation utterance which ends the current sequence. This confirmation is offered by a speaker different from the *bei* speaker. As is shown in Figure 10:

B: Why?
A: The situations in China and South Korea have been improved
B: And then?
A: (The restrictions) can be lifted then.
B: OK.
B: President Trump is fair then.
A: He made this announcement when answering questions.

Figure 10: *jiu keyi jiechu le bei* '(The restrictions) can be lifted then'

In this WeChat group chat, interactants were talking about COVID-19 status. Speaker A told group members that President Trump may lift the travel restrictions on China. A then provided the background information that the situation concerning COVID-19 in China and South Korea had been improved. B then invited A to say more about it by saying 嗯，然后呢. *en, ranhou ne* 'and then' It is at this time that A replied 就可以解除了呗 *jiu keyi jiechu le bei* 'The restrictions can be removed then', a *bei*-tagged utterance. This utterance was followed by B's confirmation utterance 好吧 *hao ba* 'OK', after which B added that 川普很公正嘛 *chuanpu hen gongzheng ma* 'President Trump is fair' A concluded the sequence by saying 也是回答问题的时候说的 *yeshi huida wenti de shihou shuo de* 'He made this announcement when answering questions.' Although B's confirmation token 好吧 (OK) does not lie at the very end of this sequence, it heralds the end. Indeed, it is just two more turns before the conversation comes to an end.

"Terminal Silence by All" is probably something novel regarding the WeChat data. It depicts a situation where a silence follows a *bei*-tagged utterance. The silence means all interlocutors do not send messages anymore and thus marks the end of the

current sequence. It does not mean that all interlocutors are rendered speechless. Instead, it is a way to end a conversation in instant communication tools such as WeChat. When all interactants feel that they do not need to give input to an on-going conversation anymore, they can just stop texting. As a convention, this "silence" suggests that all speakers think there can be the end. A new topic can be initiated later at any time. This way of closing a conversation stands in sharp contrast with face-to-face communication. In such situations, if all people present stop talking abruptly, this probably suggests someone has just said something very inappropriate and then other interactants are speechless. Let's see the following screenshots as an example:

A: But no one takes the flight. Have no idea.
A: I have to wait for another three hours.
A: Can any of you chat with me?
B: I can.
B: The travel industry in France lost at least 3% revenue.
B: LV and other luxury brands
B: all suffered.
A: You are also suffering from a shortage of face mask?
B: We?
B: No, we can still get face masks.

Figure 11: Pretext of *fanzheng ni wuliao wo jiu pei ni bei* 'Anyway, if you feel bored, I can talk with you.'

B: But we cannot mail face masks back.
A: Why?
B: Because the Customs only permit face masks as donations through the official channel.
Jan 29, 2020 08:06
B: Medical supplies mailed by private sector will probably be confiscated.
A: Didn't you mail some of them earlier?
A: Never mind.
Jan 29, 2020 08:07
A: Which carrier did you choose?
B: French Post.
B: And then EMS.

Figure 12: Extra pretext of *fanzheng ni wuliao wo jiu pei ni bei* 'Anyway, if you feel bored, I can talk with you.'

B: Whatever carrier you choose
B: Your package will be checked at the Customs.
A: [Two emojis meaning frustrated]
A: My friend's SF package from US to China was also lost.
B: So you know.
B: Under the current circumstances.
B: There are some people
B: Acquire profit in time of trouble.
A: Yes.
B: Anyway, if you feel bored, I can talk with you.
Jan 29, 2020 09:08

Figure 13: *fanzheng ni wuliao wo jiu pei ni bei* 'Anyway, if you feel bored, I can talk with you.'

We can see in Figure 11 Speaker A said 谁陪我唠唠嗑 *shui pei wo laolao ke* (as is shown in the first circled turn) 'can any of you chat with me' as a way to solicit company when he was waiting for his transfer flight at the airport. Then Speaker B jumped in to start chatting with Speaker A. This chatting happened around 08:00 AM on Jan. 20, 2020. Till the end of this chat, Speaker B uttered 反正你无聊我就陪你呗 *fanzheng ni wuliao wo jiu pei ni bei* 'If you feel bored, I can talk with you' (a bei-tagged utterance) to restate that he was willing to keep Speaker A in company during the three-hour long waiting time. It is at this point a "Terminal Silence" comes in where both Speaker A and B did not send messages anymore, neither did other group members in this WeChat group. A new sequence was initiated at 09:08 AM on the same day (as is shown by the circled time on Figure 13), almost one hour later after the previous

conversation.

The fifth type is "Disagreement by Another Speaker". For this type, a *bei*-tagged utterance is followed by an utterance showing another speaker's disagreement towards a previous speaker's statement. Since this type has only one token and thus it is not a typical use of *bei* in the dataset, no detailed analysis is offered.

To sum up at this point, the WeChat data show that the *bei-tagged* utterances behave like a "conversation terminator" when it is used as Response. 98% of their immediately subsequent turns serve as a termination of an ongoing sequence.

Manipulative *Bei*-tagged Utterances

One hundred and thirty-one manipulative *bei*-tagged utterances were found in the WeChat data, as is shown in Table 17.

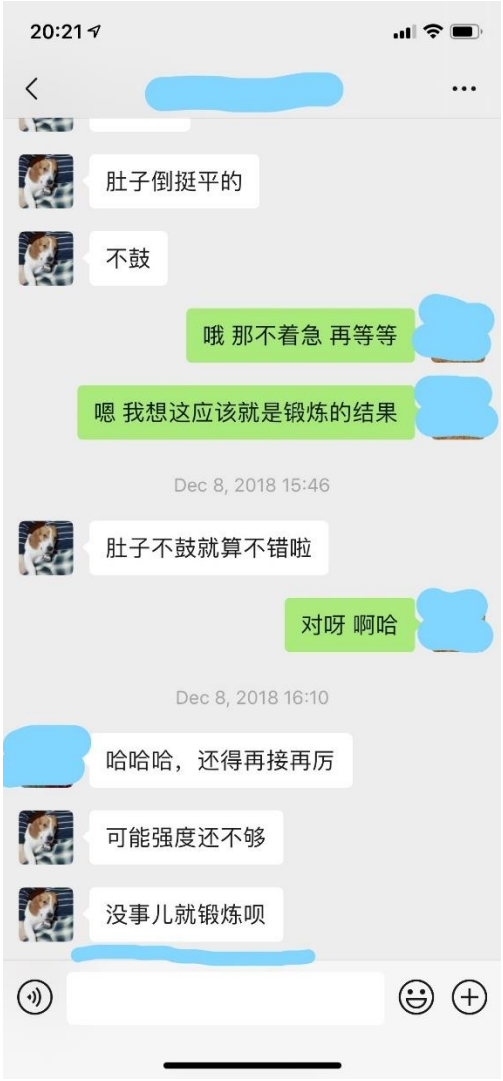
Table 17: Distribution of manipulative *bei*-tagged utterances in WeChat data

Function	Form	Frequency
Suggestion	**[Directive + *bei*]**	**70**
	• [VP + *bei*]	65
	• [VP + *jiu* + VP + *bei*]	2
	• [NP + *bei*]	2
	• [NP_{voc} + VP + *bei*]	1
Request	**[Directive + *bei*]**	**61**
	• [VP + *bei*]	40
	• [NP_{voc} + VP + *bei*]	21
Grand Total		**131**

All manipulative *bei* uses are directive sentences, which serve as either suggestion (70 out of 131, as the majority) or request (61 out of 131). For the function of suggestion, the construction [VP + bei] has the predominant token frequency (65 out of 70).

An example of suggestion in the form of [VP + bei] is presented in Figure 14, Speaker A, B, and C were talking about the effect of doing abs workout. Speaker A was

more of an expert in this aspect, so he offered the suggestion to take full advantage of one's part time to exercise by using the underlined token 没事儿就锻炼呗.



A: Your abdomen is indeed flat.	
A: Not fat.	
	B: Oh, no worries, let's wait and see.
	B: Yeah, I think this suggests the effect of doing exercises.
A: You know just a flat abdomen is already a good job.	
	B: Yes, indeed.
C: Ha ha ha, still need more practice.	
A: Probably your exercise is not intense enough.	
A: If you have free time, then just do exercises	

Figure 14: An example of *bei* used in suggestions with the form [VP + bei]

The commonality of all such suggestion tokens is that the *bei* speaker sounds like he or she has a high epistemic stance in the issue at hand (Heritage, 2013, pp. 376-378). Epistemic stance can be loosely explained as a speaker's perception of his or her knowledge about a domain. Here in the example captured by Figure 14, the domain is abs workout. Speaker A offers his suggestion by using the *bei*-tagged utterance 没事儿就锻炼呗 *meishier jiu duanlian bei* 'If you have free time, then just do exercises', which has no qualification or hedges. It sounds very confident. This results from his self-perceived high epistemic stance vis-à-vis other interlocutors. (For detailed information, readers can refer to that Heritage's chapter on epistemic stance and

epistemic status.)

For the function of request, about 70 percent of tokens are in the form of [VP + bei], which is exemplified by the conversation captured by Figure 15,

A: Is you IELTS score 7.5?	
	B: Yes.
A: Teach me a bit then.	
A: [An emoji meaning the speaker is looking forward to replies.]	
	B: Ha ha, no problem.
	B: You need to first tell me your strengths and weaknesses in your English proficiency.
A: No strengths.	

Figure 15: An example of *bei* used in requests

We can see in the conversation represented in Figure 15, Speaker A first asked Speaker B's IELTS score. Then Speaker A went on to request Speaker B to teach him English by the token 教教我呗 *jiaojiao wo bei* 'Teach me a bit then', with the reduplication of verb 教 *jiao* 'to teach' and without hedges or qualifications.

The rest tokens (21 out of 61) are in the form of [NP_{voc} + VP + bei], where NP_{voc} means a vocative is used with a verb phrase to achieve the social function of request. As is shown in Example (41),

(41) 姐姐， 把 你 拍 的 蛋糕 照片 发 一下 呗

jiejie ba ni pai de dangao zhaopian fa yixia bei

elder.sister ba you take DE cake photo send one.time BEI

'Elder sister, send me the cake photo you took.'

I did not use the original screenshot of the conversation of this example because I think only a glossing of the utterance with a vocative suffices to illustrate such usage of *bei*.

4.3. Interim Summary of *bei*

For RQ1, which is about distribution patterns, my data clearly show that *bei* can be used not only in declarative utterances, as has been claimed in all prior *bei* related studies to the best of my knowledge that *bei* only occurs in declarative utterances, but also in manipulative utterances, notably in suggestions and requests. The corpus data show that *bei* occurs in constructions in the form of [Directive + *bei*], which function as hortatory requesting and in construction in the form of [Declarative + *bei*] functioning as bald assertion.

In [Directive + *bei*], *bei* tends to cooccur with structures meaning "doing something a little bit", such as *yixia* or reduplicated verbs, and with structures denoting a small quantity, such as *ge* and *dian*. The WeChat data also shows that [VP + *jiu* + VP + *bei*] is an idiosyncratic construction, compared with other *bei*-tagged utterances. In [Declarative + *bei*], bei notably cooccurs with curt words that are bluntly evaluative or even insulting.

For RQ2, which is about SFP-related constructions and the associated functions, my corpus data show that the hortatoryness and the baldness are attributed to constructions. The hortatoryness is indicated by *bei* and its collocates such as *yixia* 'one time' and the baldness is indicated by this SFP and its single-word collocates such as *jian* 'ignoble.' More generally, both the corpus data and the WeChat data suggest that the functions performed by *bei* constructions are correlate with the sentence types that *bei* cooccurs with. The corpus data reveal that, under the Hortatory Requesting schema, *bei* is tagged with directive sentences, while under the Bald Assertion schema, it is with

declarative sentences. The WeChat data show that, the declarative *bei*-tagged utterances consist of *bei* with declarative sentences while the manipulative *bei*-tagged utterances consist of *bei* with directive sentences.

For more nuanced functions, the WeChat data show that *bei*-tagged utterances serve as "conversation terminator" that will bring an ongoing sequence to a close. Also, what emerges from the WeChat conversations I collected is the fact that *bei*-tagged utterances express suggestions with a notable sense of speaker self-perceived confidence and requests that are straightforward. Both findings are in line with the hortatoriness and the baldness of the schemas identified from the corpus data.

CHAPTER V

RESULTS FOR 呢 *NE*

5.1. Constructions of Non-interrogative *ne* (NI-*ne*) in Corpus Data

The corpus data only provides me with one schema of NI-*ne*, as is shown in Table 18,

Table 18: Constructions in NI-*ne* corpus data

Function	Form	Frequency
Dissuasion	**[Question Word + *ne*]**	**98**
	[何必 *hebi* 'why bother' + *ne*]	59
	[何苦 *heku* 'why bother' + *ne*]	16
	[那又/又能/又/能 *nayou/youneng/you/neng* 'concessive' + 怎样 *zenyang* 'how' + *ne*]	13
	[还能/又能/又/能 *haineng/youneng/you/neng* 'concessive' + 怎么样 zenmeyang 'how' + *ne*]	10
Mock Question	**[干 *gan* 'to do' + question word + *ne*]**	**89**
	[干 *gan* 'to do' + 嘛 *ma* 'what' + *ne*]	48
	[干 *gan* 'to do' + 啥 *sha* 'what' + *ne*]	19
	[干 *gan* 'to do' + 什么 *shenme* 'what' + *ne*]	14
	[干 *gan* 'to do' + 吗 *ma* 'what' + *ne*]	8
Grand Total		**187**

The DISSUASION Schema

I found that question words frequently appeared in this schema, even if the target

SFP is the NI-*ne*[34]. This suggests that the construction as a whole does not fulfill the function of interrogation in terms of seeking information, although it has question words. Some of the question words have conventional meanings by themselves. For example, 何必 *hebi* 'why bother' and 何苦 *heku* 'why bother' form rhetorical questions. For the question word 怎样 *zenyang* 'how' and 怎么样 *zenmeyang* 'how', the two question words collocate with concessives such as 那又 *nayou* / 又能 *youneng* / 还能 *haineng* / 又 *you* / 能 *neng* 'concessive'. Such utterances convey a dissuasion because speaker sees the state of affairs as being beyond one's control and nothing can be done about it and any action would be unnecessary or useless.

Consider Example (42) as an illustration,

(42) 这么 冲动， 何必 呢。
zheme chongdong, hebi ne
so impulsive why.bother NE
'(You are) so impulsive (in this situation), why bother?'

The MOCK QUESTION Schema

This schema is [干 *gan* 'do' + Question Word + *ne*]. Question words occurring in this slot are 什么 *shenme* 'what', 啥 *sha* 'what (colloquial)', 嘛 *ma* 'what', and 吗 *ma* 'what'. Functionally speaking, such utterances are mock questions. Take Example (43) as an illustration,

(43) 大家 都 在 干 什么 呢，准备 过年 吗？
dajia dou zai gan shenme ne, zhunbei guonian ma
everyone all at do what NE, prepare celebrate.the.spring.festival MA
'What is everyone doing? Preparing for the celebration of the Spring Festival?'

I say this is an NI-ne token because this utterance's interrogative tone is much weaker than the second utterance that ends with *ma*. While the *ma*-ending utterance really ask

[34] Readers might notice that some sentences in the concordances of NI-*ne* are followed by a question mark. Those tokens are not the key word in my KWIC search for NI-*ne*; they actually appear in the co-texts of a NI-*ne* token. However, the Antconc software cannot distinguish between key word NI-*ne* tokens and other *ne* tokens in their span. That is why some interrogative *ne* tokens, which are followed by a question mark, still appear. However, they, very small in number, do not disrupt the current analysis of NI-*ne*.

the question whether everyone is preparing for the Spring Festival, the *ne*-ending utterance is more of a topic initiator which brings up the issue of everyone's activity. That is why it ends with a comma, instead of a question mark.

Others constitute rhetorical questions which discourage the hearer to continue doing some activity, as is shown in Example (44),

(44) 就 不要 每天 装 纯 了 活 着 那么 累 干 什么 呢
jiu buyao meitian Zhuang chun le, huo zhe name lei gan shenme ne
just do.not everyday pretend naïve LE, live DUR so tired do what NE
'Just don't pretend to be naïve anymore. Why live such a tiresome lifestyle!'

The corpus data show that NI-ne interestingly cooccur with question words to constitute mock question utterances that seemingly exist in the form of question while perform other functions, such as topic initiating, dissuasion, etc.

5.2. Constructions of NI-*ne* in WeChat Data

The WeChat data show that all but one NI-*ne* tokens appear in declarative sentences, with the one token appearing in a directive sentence which serving as a request, as is shown in the last row above the Grand Total in Table 19. The table also reveals that the utterances perform a variety of social actions as their functions. The major functions revolve around epistemics as the most frequent function types belong to this kind. Let's see examples for the top three functions of declarative sentences tagged by NI-*ne*. Recall that the three function types are what usage-based constructionist approach terms as "pathbreakers", namely high-frequency exemplars.

"Eager Confirmation" is the most frequent one (32 out of the 89 declarative tokens). NI-ne in this case occurs in two types of utterances: 嗯呢 en ne 'Yes/OK' and 是呢 shi ne 'Yes/OK', as is underlined in Figure 16 and Figure 17 respectively. En and shi both have a confirmatory meaning. Both types of utterances serve as a confirmation of what other interlocutors have said, with an overtone of eagerness. In other words, they function as not just confirmation, but eager confirmation. Structurally speaking, both types of utterances are idiomatic formulae. In the example presented in Figure 16,

interlocutors were chatting about a smartphone accessory. After Speaker A pointed out the discrepancy between that accessory's size and Speaker B's phone, Speaker B said 嗯没事 我看看还有没有别的款式 en meishi wo Kankan hai youmeiyou bie de kuanshi 'No problem. I'll see if they have other sizes in their stock'. Speaker A then said 嗯呢 en ne 'Yes', as confirmation of B's solution. In the example presented in Figure 17, the speakers were chatting about a person who was famous for her tying flowers craftsmanship. In Speaker A's turn, he or she said 这个人扎花确实好看 zhe ge ren zhahua queshi haokan 'This person's tying flowers work looks really good'. Then in Speaker B's turn, he or she said 是呢 她干了那么多年了 shi ne ta gan le name duo nian le 'Yes. She has done this job for many years', as a confirmation.

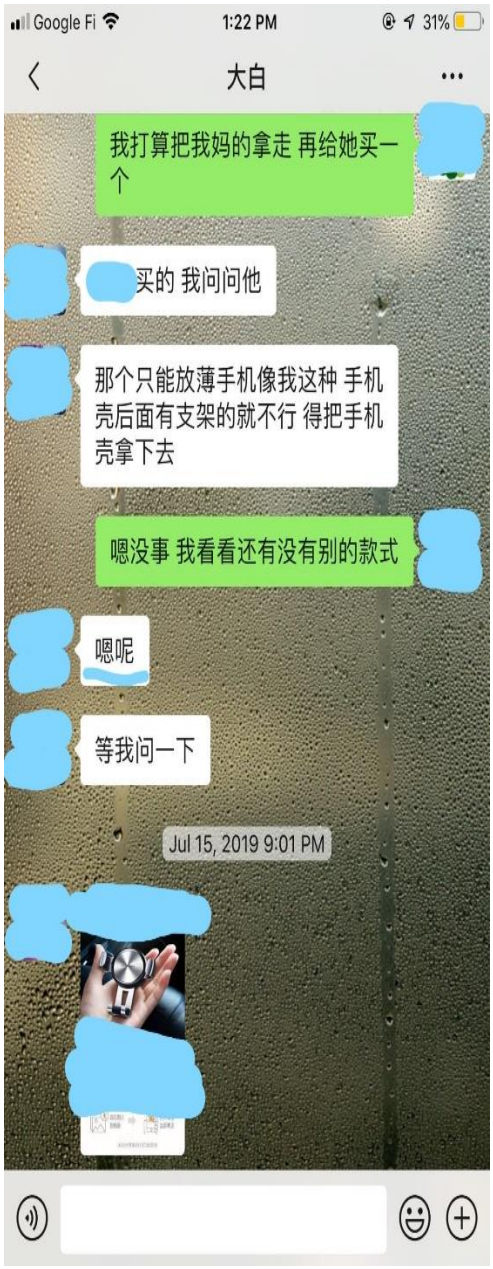
The second on the list is "Registering an Epistemic Stance". One example is presented in Figure 18, where speakers were talking about an app that could provide its users with real time boarding information. After Speaker A asked whether this app worked both in their home country and in other countries, Speaker B explicitly registered his or her own epistemic stance on this issue by saying 我不知道国外如何呢 *wo buzhidao guowai ruhe ne* 'I don't know whether it works in other countries.'

The third major function is "Drawing Attention to Speaker Current Situation". The example for this function is captured in Figure 19. Although the participant only provided me with the WeChat screenshot for only two turns, I can clearly reconstruct the situation. The *ne* speaker (Speaker A) had very likely been asked by another conversationalist about his or her whereabouts. Speaker A then gave an update on his or her current situation by saying 我在外面呢，手机静音了 *wo zai waimian ne* 'I'm outside, and my phone has been silenced'. Notice that the speaker current situation can be fully depicted by the utterance *wo zai waimian* 'I'm outside'. Accordingly, the *ne*-tagged utterance explicitly draws hearer's attention to this message. In other words, the whole utterance is markedly hearer-oriented, with the presence of this SFP.

Most of the rest functions in the declarative tokens, such as "Informing" and "Reminder", are also about epistemics in conversation.

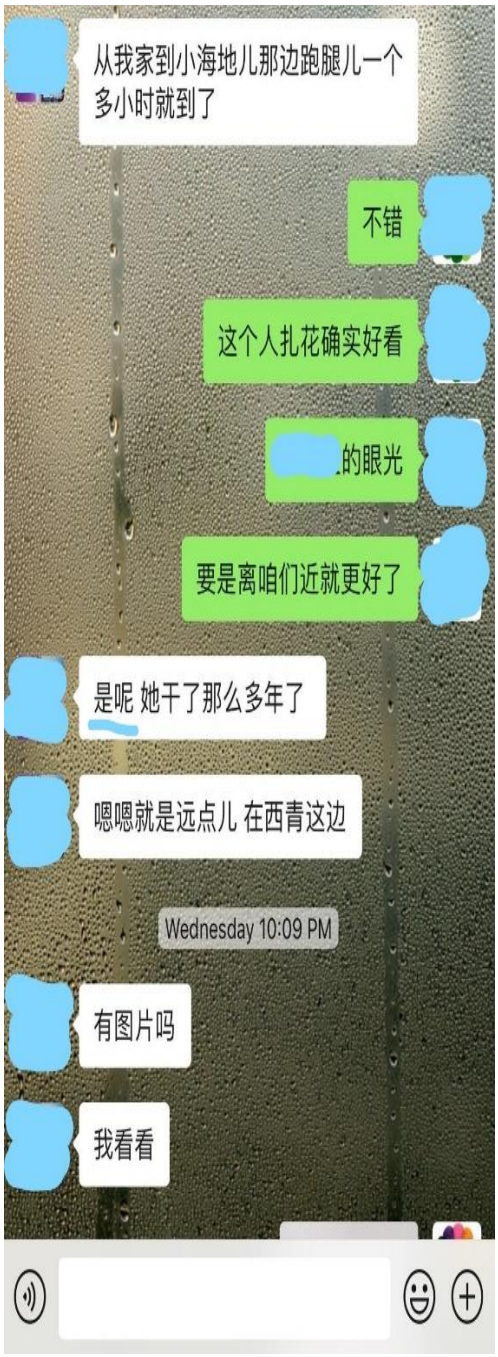
Table 19: Distribution of NI-ne-tagged utterances in WeChat data

Function	Form	Frequency
	[Declarative + ne]	**89**
Eager Confirmation		32
	[嗯 en 'Yes/OK' + ne]	16
	[是 shi 'Yes/OK' + ne]	16
Registering an Epistemic Stance	[VP + ne]	22
Drawing Attention to Speaker Current Situation	[VP + ne]	11
Negation	[VP + ne]	4
Informing		3
	[NP + ne]	2
	[VP + ne]	1
Reminder	[VP + ne]	3
Expression of Happiness	[VP + ne]	3
Complaint	[VP + ne]	2
Elaboration	[VP + ne]	1
Suggestion	[VP + ne]	1
Showing Agreement	[VP + ne]	1
Initiation of Elaboration	[VP + ne]	1
Expression of Friendliness	[VP + ne]	1
Compliment	[VP + ne]	1
Faultfinding	[VP + ne]	1
Rejection	[VP + ne]	1
Expression of Worry	[VP + ne]	1
	[Directive + ne]	**1**
Request	[VP + ne]	1
Grand Total		**90**



	B: I'll take my mom's and buy her a new one later.
A: X bought one. I'll ask him/her about it.	
A: That one can only be used on thin smartphones like mine. Those smartphones with cases that have a mount are not compatible. If one wants to use the accessory, one has to remove the case.	
	B: No problem. I'll see if they have other sizes in their stock.
A: OK.	
A: I'll ask about it.	

Figure 16: *en ne* 'OK'



B: It takes only one hour from my home to Xiaohaidier (a place name).	
	A: Nice.
	A: This person's tying flowers work looks really good.
	A: X's vision (X has a real vision on this.)
	A: I wish we could live closer to that place.
B: Yes. She has done this job for many years.	
B: Yeah. The only problem is the distance. It's nearby Xiqing (a place name).	
B: Do you have pics of their products?	
B: Let me take a look.	

Figure 17: *shi ne* 'Yes'

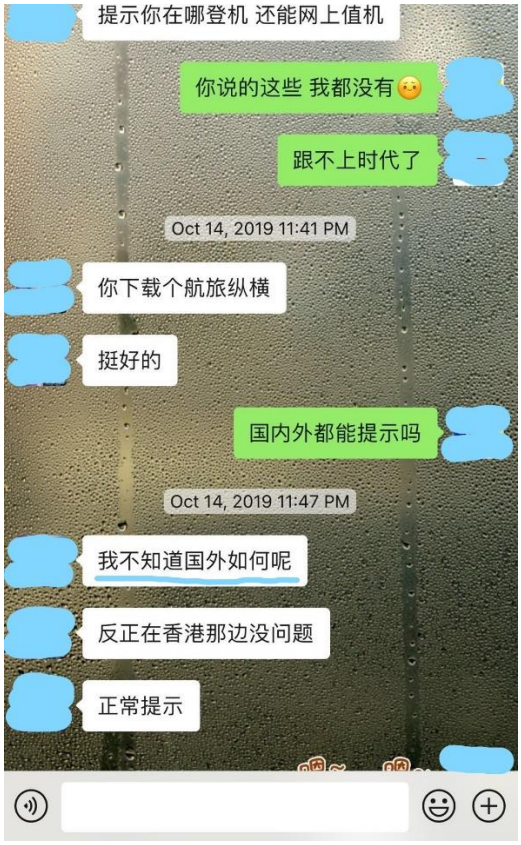
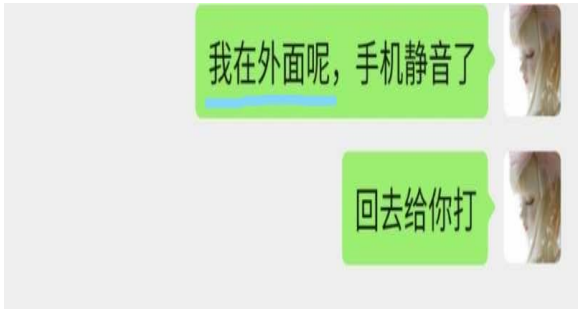


Figure 18: *wo buzhidao guowai ruhe ne* 'I don't know whether it works in other countries.'

B: (This app) can provide you with information on your gate number. It can also be used when check in online.	
	A: What you've said are new to me. [An emoji meaning embarrassment]
	A: I'm out.
B: You can just download an app called *Hanglyu Zongheng*.	
B: It's pretty nice.	
	A: Does it work both in their home country and in other countries?
B: I don't know whether it works in other countries.	
B: I know it works in Hong Kong.	
B: Works well (in Hong Kong).	



A: **I'm outside**, and my phone has been silenced
A: I'll call you back when I'm home.

Figure 19: *wo zai waimian ne* 'I'm outside'

5.3. Interim Summary of NI-*ne*

For RQ1, the NI-*ne* is mainly used in two situations, according to the "pathbreakers" that I've found. NI-ne interestingly cooccur with question words to constitute "non-question" (mock question) utterances that seemingly exist in the form of question while perform other functions, such as topic initiating or dissuasion. This finding is in line with the established fact that, functionally speaking, questions are not just used to solicit answers; they can implement many other social actions including requests, offers, challenges, criticisms, etc. (Hayano, 2013) Such NI-*ne*-tagged utterances are mock questions because they have question words in them, but such utterances are not interrogative in nature.

Another situation is when the speaker is saying something about epistemics in conversation, as is shown in the WeChat data analysis. In such cases, the speaker mainly produces confirmatory utterances, registers his or her epistemic stance, or provides updates on his or her current situations. The WeChat data findings have enriched the corpus data findings by revealing the highly frequently used [嗯 *en* 'Yes/OK' + *ne*] and [是 *shi* 'Yes/OK' + *ne*] confirmatory constructions.

For RQ2, this SFP' functions are clearly dependent on the constructions, because the NI-ne occurs in utterances that serve as contradictory functions. For example, the corpus data show that [Question Word + *ne*] can perform the dissuasion function while [嗯 *en* 'Yes/OK' + *ne*] or [是 *shi* 'Yes/OK' + *ne*] construction performs the confirmation function. Without this constructionist representation, the two opposite interpretations of its functions would no wonder point to its so-called "elusiveness."

5.4. Constructions of Interrogative *ne* (I-*ne*) in Corpus Data

Schemas and constructions for I-*ne* are presented in Table 20:

Table 20: Constructions in I-*ne* corpus data

Function	Form	Frequency
INSISTENT INTERROGATION		
Subschema 1	**[Nouns of Complexity + ne]**	**81**
	[有 you 'have' + 何 he 'what' + Disyllabic Nouns + ne]	11
	[作 zuo 'make' + 何 he 'what' + Disyllabic Nouns + ne]	8
	[意味着 *yiweizhe* 'to have a certain significance' + 什么 *shenme* 'what' + *ne*]	6
	[什么样 *shenmeyang* 'what outlook' + 的 *de* 'a particle' + Nouns of Complexity + *ne*]	43
	[答案 *daan* 'answer' + 是 *shi* 'copula' + 什么 *shenme* 'what'/多少 *duoshao* 'how many/how much' + *ne*]	13
Subschema 2	**[Verb + Question Word + ne]**	**105**
	[干 gan 'do' + Question Word + ne]	67
	[V + 些 *xie* 'a certain number' + 什么 *shenme* 'what' + *ne*]	38
Grand Total		**186**

The REQUESTING ELABORATION Subschema

Collocates in the 1L position play a significant role in this subschema, as is shown in Table 21. Multiple question words (highlighted in bold) are among the collocates,

including 什么 *shenme* ‘what’, 怎么办 *zenmeban* ‘what to do/how to deal with it’, etc.

For the construction [Nouns of Complexity + *ne*], those nouns (underlined in Table 21) are part of constructions for requesting explanations for a complex problem. Such I-*ne*-tagged utterances convey an aura of complexity; answers to questions formulated by such utterances are very likely to be of sophistication and subtlety.

Following is an explanation for what I call “nouns of complexity”:

- 关系 *guanxi* ‘relationship’ (A relationship involves at least two entities and the clarification of it consists of at least two steps: a clear knowledge of each of the entities involved and the interdependence among them. This is usually not an easy task.)
- 意义 *yiyi* ‘significance’ (A clarification on significance normally requires an answerer to present all meaningful aspects of something based on a deep analysis.)
- 原因 *yuanyin* ‘reason (noun)’ (No doubt that a reasoning process is usually sophisticated. A verbalization of a reasoning process should be fine-grained accordingly.)
- 后来 *houlai* ‘later’ (Though itself not a semantically rich word, this collocate appears in the utterance 后来呢 *houlai ne* ‘What happened later’, which is typically used to solicit more information about a situation of a later time in a narrative discourse. The response thus should be an expanded text.)
- 想法 *xiangfa* ‘thought (noun)’ (The DMC dictionary explains *xiangfa* as 思索所得的结果 *sisuo suode de jieguo* ‘the results of thinking and investigating.’ This indicates that *xiangfa* is not just a flicker of thought but a point reached by a deep analysis of a problem. As a rule, the clarification of a thought may only be achieved by extended discourses.)
- 看法 *kanfa* ‘opinion’ (A speaker needs elaboration, instead of a very short response, for other interlocutors to understand his or her opinion.)
- 作用 *zuoyong* ‘impact (noun)’ (The English definition of “impact” (an

equivalent of *zuoyong*) is "the effect or influence that an event, situation etc. has on someone or something" (Longman Dictionary of Contemporary English (4th edition), p. 812). It normally requires a speaker to provide detailed information about the event or situation and the relationship between all parties involved before he or she can crystallize an effect or influence therein.)

- 特征 *tezheng* 'feature (noun)' (This collocate appears in utterances such as 有什么特征 *you shenme tezheng* 'what feature(s) does it have', 有何特征 *you he tezheng* 'what feature(s) does it have' and 有哪些特征 *you naxie tezheng* 'what features does it have'. 哪些 *naxie* 'which ones' is a question word to ask questions about an entity as a whole which has more than one part. When this question word is used, it indicates that the speaker assumes there must be more than one element involved in the answer. Its equivalent in English is "which ones" (Sun, 2006, p. 172); a response to such questions is supposed to include several aspects or a list of several items. My point is that a satisfactory answer to a question about features of some entity probably constitutes a discourse-level response.)
- 现象 *xianxiang* 'phenomenon' (The explanation for this word offered by the DMC Dictionary is 事物在发展、变化中所表现的外部的形态和联系 'the external form and connection of things in their development and change'. Such I-*ne*-tagged questions are about a certain phenomenon, primarily about the underlying reasons of a certain phenomenon, and to fully answer those questions requires an elaborated and organized discourse.)
- 处置 *chuzhi* 'to handle' (All questions containing this collocate are about how to handle a situation. A satisfactory answer needs at least a list of procedures of a solution and rationales thereof.)
- 区别 *qubie* 'difference' (All the questions are about differences between two entities. Therefore, a qualified response should consist of a description of the two involved entities and the dissimilarity therein.)
- 角色 *juese* 'role' (This collocate is used both literally and metaphorically in the

exemplar sentences. The questions are primarily about nuts and bolts of one role or multiple roles, either in a real play setting when a role means "the character played by an actor in a play or film" or in a metaphorically similar contexts such as daily life when the meaning of a role is "the way in which someone or something is involved in an activity or situation, and how much influence they have on it" (Longman Contemporary English Dictionary (4th edition), p. 1426). To describe a role is very likely to be sophisticated.)

- 命运 *mingyun* 'destiny' (Obviously the question of destiny is intriguing for everyone. One will face considerable complication when answering questions concerning destiny.)

Some collocates appearing in other positions are also meaningful. One such collocate is 何 *he* 'what', in the 2L position. An examination of corpus shows this collocate is part of the construction [作 *zuo* 'make'/有 *you* 'have' + 何 *he* 'what' + Disyllabic Nouns + *ne*]. The noun in this construction in almost all cases is a disyllabic word, including 感想 *ganxiang* 'reflective thoughts', 评价 *pingjia* 'evaluation', 解释 *jieshi* 'explanation', 窘境 *jiongjing* 'dillema', 打算 *dasuan* 'plan', 意义 *yiyi* 'significance', etc. In Chinese, disyllabic words are typical of the written and formal register. *He* is a question word inherited from classical Chinese. The juxtaposition of *he* and the disyllabic nouns sounds very formal, which conveys an overtone that the speaker wants a serious answer from the respondent. There is no simple answer to those questions, as is illustrated in Example (45) and Example (46),

(45) 你 现在 作何 打算 呢？
ni xianzai zuo he dasuan ne
you now make what plan NE
'What plan do you have now?'

(46) 这 种 差别 有 何 意义 呢？
zhe zhong chabie you he yiyi ne
this kind difference have what significance NE
'What is the significance of such a difference?'

Table 21: Collocate display of 1L of I-*ne*

Freq	Stat (MI)	Collocate
519	3.06059	**什幺 *shenme* 'what'**
67	4.19732	**怎幺办 *zenme ban* 'what to do'**
63	3.21422	**哪里 *nali* 'where'**
58	3.92781	**怎幺样 *zenme yang* 'how'**
28	3.06983	**何必 *hebi* 'why bother'**
27	3.32222	办 *ban* 'do'
24	3.04993	然后 *ranhou* 'afterwards'
21	4.65479	玩玩 *wanwan* 'to play a little bit'
19	3.04474	礼物 *liwu* 'gift'
17	3.31599	关系 *guanxi* 'relationship'
16	3.7006	意义 *yiyi* 'significance'
13	3.0333	原因 *yuanyin* 'reason (noun)'
12	3.5959	后来 *houlai* 'later'
11	4.02676	方舟 *nuoya fangzhou* 'Noah's ark'
11	3.8663	**何苦 *heku* 'why bother'**
11	3.41378	想法 *xiangfa* 'thought (noun)'
10	4.65479	**何乐而不为 *hele er buwei* 'Why not'**
10	3.65479	多久 *duojiu* 'how long'
9	3.36529	变化 *bianhua* 'change'
8	3.84744	看法 *kanfa* 'opinion'
8	3.7479	作用 *zuoyong* 'impact (noun)'
8	3.40686	**干吗 *gan ma* 'Do what'**
8	3.40686	合适 *heshi* 'appropriate'
8	3.33286	状态 *zhuangtai* 'status'
7	3.76171	特征 *tezheng* 'feature (noun)'
7	3.21422	反应 *fanying* 'reaction'

(continued)

Freq	Stat (MI)	Collocate
7	3.06983	上网 *shangwang* 'to surf the Internet'
6	3.78032	做梦 *zuomeng* 'to dream / to daydream'
6	3.15229	现象 *xianxiang* 'phenomenon'
5	4.65479	益处 *yichu* 'advantage'
5	4.65479	心动 *xindong* 'to start to feel interested in'
5	4.16936	处置 *chuzhi* 'to handle'
5	3.97672	区别 *qubie* 'difference'
5	3.39176	角色 *juese* 'role'
5	3.27628	女朋友 *nyupengyou* 'girlfriend'
5	3.06983	命运 *mingyun* 'destiny'

Another significant collocate in the 2L position is 意味着 *yiweizhe*, which means 含有某种意义 'to have a certain significance', according to the DMC Dictionary. It is part of the construction [意味着 *yiweizhe* 'to have a certain significance' + 什么 *shenme* 'what' + *ne*]. I categorize this collocate, although a verbal expression, as nouns of complexity because it is related to the noun 意味 *yiwei* 'profound meaning'. This construction constitutes a question about the significance of something. To answer such a question requires a thorough explanation, as is illustrated in Example (47),

(47) 这-对-中国-经济-和-社会-发展-意味着-什么-呢？

zhe-dui-zhongguo-jingji-he-shehui-fazhan-yiweizhe-shenme-ne

this-to-China-economy-and-society-development-to.have.a.certain.significance-what-NE

'What does this mean to the economic and societal development of China?'

Also, the collocate 什幺样 *shenmeyang* 'what outlook', in the 3L position. It is part of the construction [什幺样 *shenmeyang* 'what outlook' + 的 *de* 'a particle' + Nouns of Complexity + *ne*]. Nouns filling the slot of this construction also tend to belong to the formal register, such as 状态 *zhaungtai* 'status', 作用 *zuoyong* 'impact', 世界观 *shijieguan* 'worldview', 反应 *fanying* 'response', 思考 *sikao*

'thinking', 影响 *yingxiang* 'impact', etc., and again a familiar situation is encountered: the answer to such questions should be sophisticated. One utterance is presented as Example (48),

(48) 这 将 对 我们的 生活 带来 什么样 的 影响 呢？
zhe jiang dui womende shenghuo dailai shenmeyang de yingxiang ne
this will to our life bring what outlook DE impact NE
'How will this impact our life?'

Likewise, the collocate 答案 *daan*, which is in the 3L position, is explained by the DMC Dictionary as 对问题所做的解答 'the answer to a question'. This collocate is part of the construction [答案 *daan* 'answer' + 是 *shi* 'copula' + 什么 *shenme* 'what'/多少 *duoshao* 'how many/how much' + *ne*], meaning "what is the answer to XP". It usually takes efforts to clearly explain the answer to other interlocutors. As is illustrated in Example (49),

(49) 你 这 个 谜语 的 答案 是 什么 呢？
ni zhe ge miyu de daan shi shenme ne
you this CL riddle DE answer be what NE
'What is on earth the answer to this of your riddle?'

The QUESTIONING ONE'S ACTIVITIES Subschema

There are two collocates that are particularly relevant with this subschema. 干 *gan* 'do', in the 2L position, is part of the construction [干 *gan* 'do' + Question Word + *ne*]. Recall [干 *gan* 'do' + Question Word + *ne*] is also a construction for NI-*ne*. The difference is that constructs of this construction under the current schema are followed by a question mark while those constructs of the construction with the almost identical form under the previous Quasi-Questioning One's Activity schema are not. Constructs under the previous schema are "quasi" questions (mock questions) because their functions are not interrogative while the current "real" questions are genuinely asking about one's activity, via three question words: *shenme* is a general one and *sha* and *ma* are ones used in colloquial speech. As is shown in Example (50) and Example (51),

(50) 你们 准备 干 什么 呢？

nimen zhunbei gan shenme ne

you prepare do what NE

‘What are you going to do?’

(51) 下午 都 去 干 啥 呢？

xiawu dou qu gan sha ne

afternoon all go do what NE

‘What (are you going to) do in the afternoon?’

The second collocate is 些 *xie*, also in the 2L position, which is defined by the DMC Dictionary as 表示不定的数量 ‘to indicate a certain number’；一些 *yixie* ‘a few’. This collocate mainly appears in the construction [V + 些 *xie* ‘a certain number’ + 什么 *shenme* ‘what’ + *ne*]. *Xie* here is a classifier while the question word refers to the object of the verb. The questions expressed by constructs of this construction therefore imply that there are several things, instead of only one, as the object of the verb. This again indicates the complexity of answering such questions. The answerer has to provide at least a list of things so as to provide a satisfactory answer. As is shown in Example (52),

(52) 周末 打算 做 些 什么 呢？

zhoumo dasuan zuo xie shenme ne

weekend plan do some what NE

‘What do you plan to do in the weekend?’

Summing up at this point, interrogative *ne* utterances can be categorized under the Schema Insistent Interrogation, which consists of two subschemas: [Nouns of Complexity + *ne*], which has the function of requesting elaboration, and [V + Question Word + *ne*], which is deployed to “really” ask one’s activity, as is different from the “quasi” questions conveyed by the non-interrogative *ne* constructs. I call the grand schema “Insistent Interrogation” because I find that questions under this schema are not just questions, as the message of interrogation can be expressed by the pre-SFP utterances, but questions with a tone of insistence. In other words, the speaker

particularly uses the ne-tagged utterance to let the hearer know that he or she is very much interested in getting the answer.

5.5. Constructions of I-*ne* in WeChat Data

Table 22: Distribution of interrogative *ne*-tagged utterances in WeChat data

Function	Form	Frequency
Genuine Question (Inquiry about Information)		**13**
	• [WHQ + ne]	12
	• [YNI + ne]	1
Faultfinding	**[WHQ + ne]**	**4**
Confirmation Request		**3**
	• [YNI + ne]	2
	• [WHQ + ne]	1
Genuine Question (Inquiry about Reason)	**[WHQ + ne]**	**3**
Suggestion		**2**
	• [WHQ + ne]	1
	• [YNI + ne]	1
Initiation of a Topic	**[WHQ + ne]**	**2**
Showing Worry	**[WHQ + ne]**	**1**
Registering Epistemic Stance	**[WHQ + ne]**	**1**
Grand Total		**29**

The WeChat findings for I-*ne* are rather straightforward, as is shown in Table 22. It occurs in both wh-questions (shortened as "WHQ") and yes-no questions (In the table, "YNI" stands for "yes-no interrogative". YNI is used in Conversation Analysis literature and it refers to yes-no questions. I adopted the term YNI simply because "interrogative" is a more explicit, linguistic label of the sentence type of such an

utterance).

25 out 29 tokens are WHQ tokens. Answers to WHQs are very likely not a simple answer but one with complexity to some degree. 15 out of 25 WHQ tokens are genuine questions, meaning they are used as questions to ask for either information or reasons for something.

	A: Never mind.
B: Yeah. My six days here… Usually I won't have a problem even if I take a whole day off.	
	A: [An emoji meaning happiness]
B: When will you arrive home?	
	A: [An emoji meaning feel blessed]
	A: Saturday afternoon.
B: So soon!	
	A: I didn't buy X in the states. There was no discount then. It's not a good deal I thought. Do you need me to buy X here?

Figure 20: *ni shashihou dao jia ne* 'when will you arrive home'

is an example of I-*ne*-tagged WHQ inquiring about information where Speaker B used the I-*ne*-tagged utterance 你啥时候到家呢 *ni shashihou dao jia ne* 'when will you arrive home' to ask Speaker A's travel itinerary. Figure 21 is an example of I-*ne*-tagged WHQ inquiring about reason where Speaker A asked Speaker B about the reason why B was still awake at midnight, as is indicated by the time "01/17/2020 12:29 AM" in the screenshot.

Figure 21 *ni za hai bu shui ne* 'Why don't you sleep'

A: Why in the whole world are you still awake?	
	B: I'm thinking too many things now.
	B: Can't fall asleep.
	B: Never mind.
A: Why the hell are you thinking so much?	

5.6. Interim Summary of I-*ne*

For RQ1, I-*ne* mainly appears in WHQs or YNIs, which parallel with findings based on corpus data, as questions addressed to interlocutors generally request explanation or elaboration. Notably, the corpus data show that I-*ne* is typically used in the subschema of [Nouns of Complexity + *ne*], which has the function of requesting elaboration, and the subschema [V + Question Word + *ne*], which is deployed to "really" ask one's activity, as is different from the "quasi" questions conveyed by NI-*ne* constructs. The frequent cooccurrence of I-*ne* with question words is understandable as I-*ne* is an SFP used in interrogative utterances. Its frequent cooccurrence with what I call "nouns of complexity" is an overlooked fact by previous studies. The grand schema is what I call "Insistent Interrogation."

For RQ2, the data show that the constructions in which I-*ne* occurs clearly invite elaboration or explanation on the part of other interlocutors. The overtone of inviting an elaborated response is conveyed by the constructions as a whole and marked by the presence of multiple question words and nouns of complexity.

CHAPTER VI

RESULTS FOR 啊 *A*

6.1. Constructions of Non-Interrogative *a* (NI-*a*) in Corpus Data

Constructions of NI-*a* are summarized as in Table 23,

Table 23: Constructions in NI-*a* corpus data

Function	Form	Frequency
EXPRESSION OF AFFECTIVE STANCE	**[VP/NP + *a*]**	**820**
	[VadjP + *a*]	**662**
	[伤不起 *shangbuqi* 'cannot be hurt anymore' + *a*]	150
	[给力 geili 'It really works' + a]	105
	[无聊 wuliao 'boring' + a]	54
	[不容易 burongyi 'not easy' + a]	40
	[Degree Adverb + 累 lei 'tired' + a]	40
	[了不起 liaobuqi 'great' + a]	21
	[激动 jidong 'excited' + a]	20
	[难受 nanshou 'unbearable' + a]	18
	[不易 buyi 'not easy' + a]	16
	[无奈 wunai 'helpless' + a]	15
	[郁闷 yumen 'depressed' + a]	14
	[麻烦 mafan 'troublesome' + a]	13
	[苍天 cangtian 'the sky' + a]	12
	[来之不易 laizhibuyi 'hard-won' + a]	11
	[诱人 youren 'attractive' + a]	10
	[受不了 shoubuliao 'cannot stand anymore' + a]	10
	[Degree Adverb + 丑 chou 'ugly' + a]	8
	[Degree Adverb + 热闹 renao 'hassle and bustle' + a]	8
	[好笑 haoxiao 'funny' + a]	8
	[吓人 xiaren 'scary' + a]	8
	[奢侈 shechi 'luxurious' + a]	7
	[威武 weiwu 'grand' + a]	7
	[悲哀 beiai 'sad' + a]	7

(continued)

Function	Form	Frequency
	[好强 haoqiang ‘so strong’ + a]	6
	[喜庆 xiqing ‘festive’ + a]	6
	[心疼 xinteng ‘frustrated’ + a]	6
	[惬意 qieyi ‘comfortable’ + a]	6
	[sorry + a] (an instance of codeswitching)	6
	[难看 nankan ‘ugly’ + a]	5
	[亲切 qinqie ‘accepting and friendly’ + a]	5
	[猥琐 weisuo ‘wretched’ + a]	5
	[淫荡 yindang ‘lewd’ + a]	5
	[得意 deyi ‘delighted’ + a]	5
	[惭愧 cankui ‘guilty’ + a]	5
	[VP + a]	**117**
	[坑爹 kengdie ‘to deceive father (generally means “to deceive”)’ + a]	67
	[闹哪样 naonayang ‘What the hell are you doing’ + a]	19
	[霸气 baqi ‘the quality of being domineering’ + 侧漏/外漏 celou/wailou ‘to leak’ + a]	8
	[没有天理 meiyou tianli ‘don’t have fairness (not fair)’ + a]	7
	[是 shi ‘copula’ + 王道 wangdao ‘kingly way’ + a]	6
	[作孽 zuonie ‘to be evil’ + a]	5
	[在民间 zai minjian ‘in the folk’ + a]	5
	[NP + a]	**41**
	[人才 rencai ‘talent’ + a]	17
	[奇葩 qipa ‘strange flower (to be weird)’ + a]	9
	[差距 chaju ‘gap’ + a]	9
	[美味 meiwei ‘delicious food’ + a]	6
INTENSIFIED INSULT	**[NP + a]**	**149**
	[你妹 nimei ‘your younger sister (an insulting word)’ + a]	137
	[畜生 chusheng ‘animal (in a derogative sense)’ + a]	7
	[你妹夫 nimeifu ‘your younger sister’s husband (an insulting word) + a]	5
EAGERLY CALLING FOR ACTION	**[VP + a]**	**86**
	[一起来抢 yiqilaiqiang ‘Let’s grab’ + a]	41

(continued)

Function	Form	Frequency
	[救命 jiuming ‘save life (help me)’ + a]	23
	[小心 xiaoxin ‘small heart (to watch out)’ + a]	22
Grand Total		**1055**

The EXPRESSION OF AFFECTIVE STANCE Schema

This schema straddles both [VP + *a*], such as [坑爹 *kengdie* ‘to deceive father (generally means “to deceive”)’ + a], and [NP + *a*], such as [奇葩 *qipa* ‘strange flower (to be weird)’ + a], as the form. Within [VP + *a*], a subtype is [$V_{adj}P$ + *a*], which basically means an adjective cooccurs with *a*. It is termed as “$V_{adj}P$” as I follow Li and Thompson (Li & Thompson, 1981) in upholding that Mandarin adjectives are categorized as adjectival verbs in terms of syntactic structure (pp. 141-147). In other words, an adjective phrase is seen as a subcategory of VP (verb phrase).

The common denominator of all constructions under this schema is that they are used by the speaker to express an affective stance (see Wang et al., 2021, a recent study on affective stances expressed by a Mandarin constructional idom [you *X* you *Y*]). It can be a positive affective stance, such as Example (53),

(53) 超赞 的 白切 羊肉， 诱人 啊！
chaozan de baiqie yangrou, youren a
fabulous DE white.cut lamb attractive A
‘This is really fabulous white-cut lamb, so attractive!’

or a negative affective stance, such as Example (54),

(54) 一 个 人 上班， 好 无聊 啊
yi ge ren shangban, hao wuliao a
one CL person go.to.work so boring A
‘One has to go to work alone. This is so boring.’

The INTENSIFIED INSULT Schema

Constructions under this schema are used by the speaker to insult someone, as is shown in Example (55),

(55) 一天-到-晚-就-是-钱-钱-钱，钱-你妹-啊

yitian-dao-wan-jiu-shi-qian-qian-qian, qian-nimei-a

through.out.the.day-reach-night-adverb-copula-money-money-money-money-your.sister-A

'You are talking about money throughout the day. Money your sister.'

I call it "Intensified Insult" because the insult can be expressed by the noun phrases preceding a. With a tagged, the utterances deliver intensified insults. This finding is in line with findings about the previous EXPRESSION OF AFFECTIVE STANCE schema in that constructions under both schemas are affectively loaded.

The EAGERLY CALLING FOR ACTION Schema

By "EAGERLY CALLING FOR ACTION" I mean such constructions are used to enact other interlocutors to take some actions, with an overtone of eagerness. As is shown in Example (56),

(56) 救 命 啊

jiu ming a

save life A

'Please save life (help me)!!!'

The message of calling for action can be expressed by the utterance *jiuming* 'save life (help me)'. With *a* tagged, the whole utterance sounds much more eager, which is typically used when the speaker is facing imminent danger.

To sum up at this point, non-interrogative *a* occurs in constructions that are notably used to express the speaker's stance, with the eagerly calling for action tokens as a small portion (86 out of 1055). EXPRESSION OF AFFECTIVE STANCE has the highest token frequency (820 out of 1055). Such Affective stance can be either positive,

negative, or neutral, which suggests its vast compatibility in contexts.

6.2. Constructions of NI-*a* in WeChat Data

Table 24: Distribution of non-interrogative *a*-tagged utterances in WeChat data

Function	Form	Frequency
Evaluation	**[Exclamative + a]**	**77**
	• [VP + a]	69
	• [NP + a]	5
	[Declarative + a]	**3**
	• [VP + a]	2
	• [NP + a]	1
Emphatic Confirmation	**[Declarative + a]**	**41**
	• [是 shi 'copula' + a]	38
	• [VP + a]	3
Expression of Surprise		**31**
	Zero Form	**11**
	[Declarative + a]	**13**
	• [VP + a]	10
	• [NP + a]	3
	[Exclamative + a]	**7**
	• [NP + a]	7
Showing Consent	**[Declarative + a]**	**8**
	• [VP + a]	8
Suggestion		**7**
	[Directive + a]	**5**
	• [VP + a]	5
	[Declarative + a]	**2**
	• [VP + a]	2

(continued)

Function	Form	Frequency
Registering an Epistemic Stance	**[Declarative + a]**	**5**
	• [VP + a]	5
Request		**5**
	[Directive + a]	**4**
	• [VP + a]	4
	[Declarative + a]	**1**
	• [VP + a]	1
Expression of Desire		**4**
	[Exclamative + a]	**2**
	• [VP + a]	
	[Declarative + a]	**2**
	• [VP + a]	
Expression of Worry	**[Declarative + a]**	**4**
	• [VP + a]	4
Description of Hardship	**[Declarative + a]**	**4**
	• [VP + a]	4
Friendly Reminder		**3**
	[Directive + a]	**2**
	• [VP + a]	2
	[Declarative + a]	**1**
	• [VP + a]	1
Expression of Good Wish	**[Declarative + a]**	**2**
	• [VP + a]	2
Showing Disagreement	**[Declarative + a]**	**2**
	• [VP + a]	2
Response	**[Declarative + a]**	**2**
	• [VP + a]	2

(continued)

Function	Form	Frequency
Pause	**[Declarative + a]**	**1**
	• [VP + a]	1
Promise	**[Declarative + a]**	**1**
	• [VP + a]	1
Informing	**[Declarative + a]**	**1**
	• [VP + a]	1
Expression of Gratitude	**[Declarative + a]**	**1**
	• [VP + a]	1
Reminder	**[Directive + a]**	**1**
	• [VP + a]	1
Encouragement	**[Directive + a]**	**1**
	• [VP + a]	1
Invitation	**[Directive + a]**	**1**
	• [VP + a]	1
Description of Problem	**[Declarative + a]**	**1**
	• [VP + a]	1
Expression of Inner Feeling	**[Exclamative + a]**	**1**
	• [VP + a]	1
Grand Total		**204**

Obviously, the WeChat data shown in Table 24 indicates that NI-*a* appears in a variety of sentence types. Two major sentence types are exclamative (84 out of 204) and declarative (95 out of 204). The third sentence type is directive, which is relatively infrequent (14 out of 204).

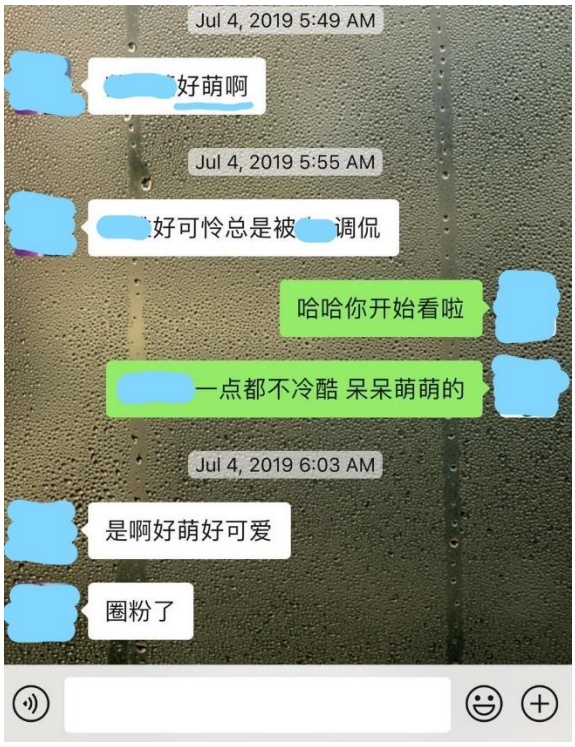
Notice that the dataset has revealed one special sentence type which I named "Zero Form." It means NI-*a* in this case stands alone in discourse to express surprises. I will not present an analysis of this type either, because for one thing, it is not frequent (11

out of 204); for another, this usage is not a "sentence-final particle" in its strict sense because it is not attached to anything, but an utterance in its own right.

NI-*a* in Exclamative Sentences

The majority of such sentences (74 out of 84) perform an evaluation function. The rest of sentences in this category are used to express speaker feelings. This pattern is well in line with the corpus data findings as the schema with the highest token frequency is the EXPRESSION OF AFFECTIVE STANCE schema.

Let's first see an example of NI-*a* used in the evaluation situation, which is the predominantly major function (the example conversation is captured in Figure 22). The two speakers were chatting about some celebrities' performance probably in some live shows. Speaker A initiated the evaluative language episode by saying X 好萌啊 X *hao meng a* 'X was really cute'.



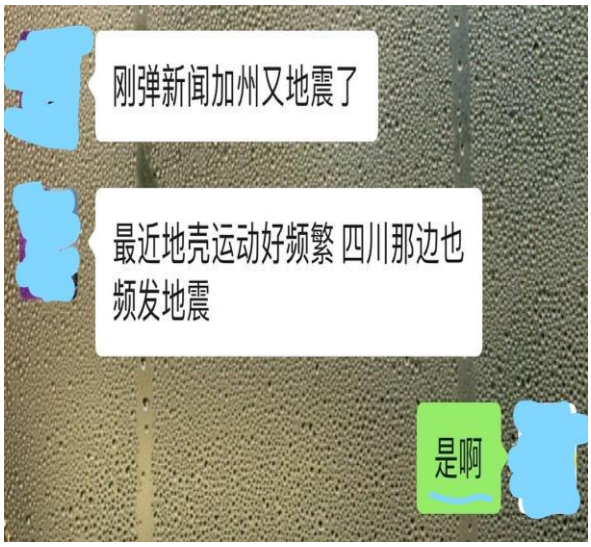
A: X was really cute.	
A: Poor X. He/she was always teased by YYY.	
	B: Ha ha, it seems you've started streaming the show then.
	B: X was not cool at all. He/she was just being cute.
A: Yeah. So cute, so adorable.	
A: Attracting so many fans now.	

Figure 22: X *hao meng a* 'X was really cute'

NI-*a* in Declarative Sentences

The top three functions for NI-*a*-tagged declarative sentences are "Emphatic

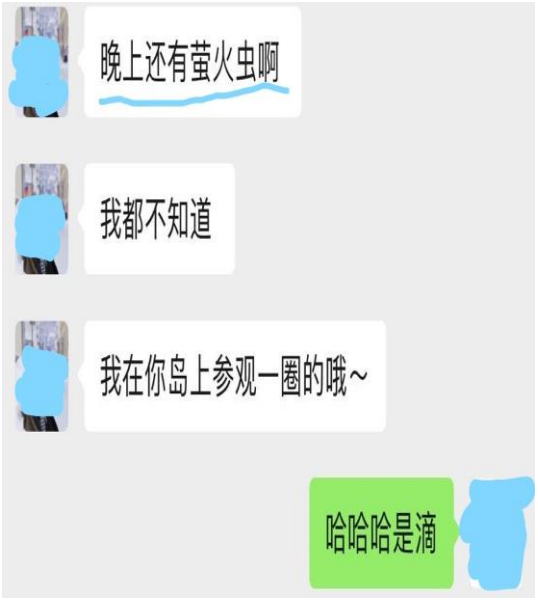
Confirmation", "Expression of Surprise", and "Showing Consent". Together they make up more than 70% of the data (59 out of 82). For the 41 confirmative tokens, three of them are 对啊 *dui a* 'Yes', and the rest of them are all 是啊 *shi a* 'Yes'. Both types of utterances are idiomatic formulas. An example of *shi a* used as confirmation is presented in Figure 23. In this conversation, Speaker A told Speaker B that an earthquake had just happened in California. Speaker A added that he or she thought earthquake had been frequent worldwide around that time. One earthquake also had happened in the Sichuan Province in China. Speaker B then confirmed this statement by saying *shi a*. Notice that the message of confirmation is clearly expressed by the copula *shi*. With *a* tagged, the utterances serve as emphatic confirmations, which show that the speaker shares similar feelings with the hearer and the speaker is comfortably make this confirmation.



A: The news just showed that there had been an earthquake in CA.	
A: It seems this is happening a lot recently. There was one in Sichuan, China, too.	
	B: Yes.

Figure 23: *shi a* 'Yes'

"Expression of Surprise" means NI-*a* appears in expressions showing the speaker's surprise, but at the same time the sentence type is declarative, not exclamative. One example is provided in Figure 24. Speaker A was surprised at the fact that Speaker B's island (probably some virtual one in some video game) had fireflies, by saying 晚上还有萤火虫啊 *wanshang haiyou yinghuochong a* 'There are even fireflies at night'.



A: There are even fireflies at night	
A: I don't know that at all.	
A: Let me take a tour on your island.	
	B: Ha ha ha, yes.

Figure 24: *wanshang haiyou yinghuochong a* 'There are even fireflies at night'

"Showing Consent" means the speakers agree on a proposal made by other interlocutors. One example is shown in Figure 25. Speaker A proposed that he or she with Speaker B went to swim and play badminton on Saturday. Speaker B then gave his or her consent by saying 好啊 *hao a* 'OK' and went on to show his or her preference over swimming.

A: Let's go to swim and play badminton on Saturday.	
	B: OK.
	B: I prefer swimming.

Figure 25: *hao a* 'OK'

NI-*a* in Directive Sentences

The two major functions performed by NI-*a*-tagged utterances are suggestion (5 out of 14) and request (4 out of 14), which is exemplified in Figure 26 and Figure 27,

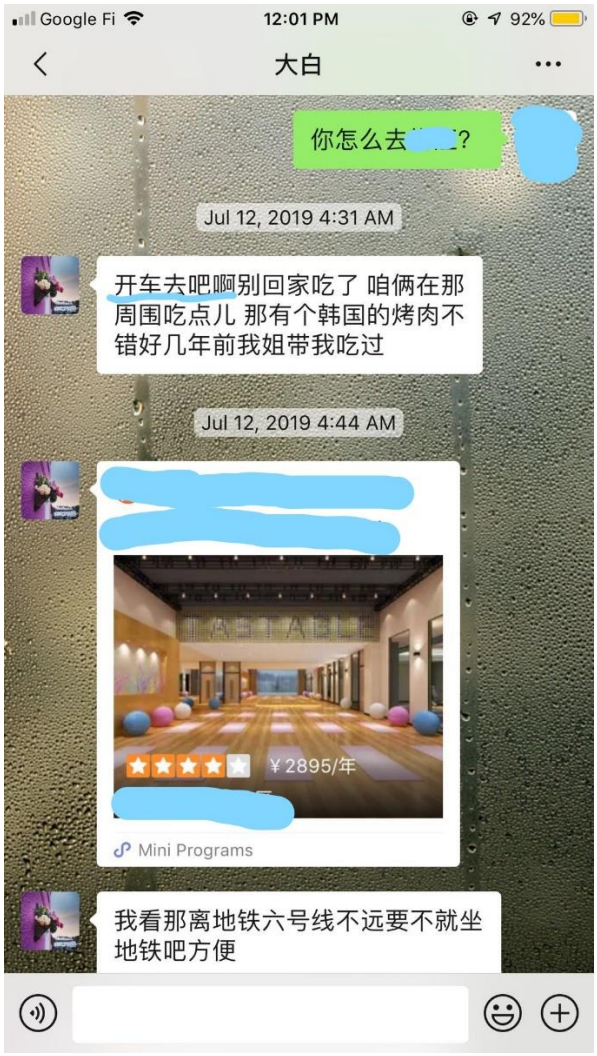


Figure 26: *kaiche qu ba a* '(I suggest) we drive to X.'

	A: How will go to X?
B: **(I suggest) we drive to X.** We don't eat at home. We can dine out at some restaurant near X. There is a great Korean BBQ restaurant. I've been to that place with my elder sister several years ago.	
B: [A link to the website of that Korean BBQ restaurant]	
B: (On a second thought) I see that that restaurant is not far from Subway Line 6. How about we take Line 6? It's convenient.	

In the conversation captured in Figure 26, Speaker A and Speaker B were chatting about their imminent trip to a place X. After A initiated the topic, B suggested by the *a*-tagged utterance 开车去吧啊 *kaiche qu ba a* '(I suggest) we drive to X.' And then s/he went on to suggest that they could have a mean at a Korean BBQ restaurant nearby X.

A: X's songs are also pretty good.	
	B: OK, I see. I'll check it out. **You can recommend some songs of his or her to me if you find any**.
A: OK. OK.	

Figure 27: *ni you haoting de jiu tuijian wo a* 'You can recommend some songs of his or her to me if you find any.'

In the conversation captured in Figure 27, Speaker A and Speaker B were chatting about a singer X. After A initiated the topic by praising X's songs, B requested by the *a*-tagged utterance 你有好听的就推荐我啊 *ni you haoting de jiu tuijian wo a* 'You can recommend some songs of his or her to me if you find any.'

6.3. Interim Summary of NI-*a*

NI-*a* constructions fulfill much more functions than any other SFPs discussed in the current study. Its extraordinary flexibility has already been noticed by Li (2006). A discussion of this finding will be presented in the Discussion chapter.

6.4. Constructions of Interrogative *a* (I-*a*) in Corpus Data

Schemas and constructions of I-*a* are summarized in Table 25:

Table 25: Constructions in I-*a* corpus data

Function	Form	Frequency
SHOWING CONCERN TO HEARER	**[Question Word / NP + a]**	**709**
	[哪里/哪儿 nali / naer 'where' + a]	159
	[怎么办/肿么办/咋办 zenme ban 'What to do' + a]	86
	[怎么样 *zenme yang* 'how' + *a*]	71
	[怎么回事 zenme hui shi 'What's the matter' + a]	60
	[什么意思/啥意思 shenme yisi/sha yisi 'What's the meaning' + a]	44
	[什么事 shenme shi 'what matter' + a]	41
	[怎样 zenyang 'how' + a]	29
	[真的假的/真的还是假的 zhende jiade/zhende haishi jiade 'true or false' + a]	29
	[什么情况/神马情况/啥情况 shenme qingkuang/shenma qingkuang/sha qingkuang 'What's the matter' + a]	27
	[多久 duojiu 'how long' + a]	17
	[行不行 xingbuxing 'OK or not' + a]	15
	[咋样 zayang 'OK or not' + a]	14
	[咩事 meishi 'what matter' + a]	13
	[几时 shenme shihou / jishi 'when' + V + a]	13
	[什么节目/咩节目 shenme jiemu / mei jiemu 'what program' + a]	12
	[什么时候/几时 shenme shihou / jishi 'when' + 结婚 jiehun 'marry' + a]	10
	[怎么一回事 zenme yi hui shi 'What's the matter' + a]	10
	[啥事 shashi 'what matter' + a]	8
	[有 you 'to have' + 什么/啥/咩/神马 shenme/sha/mei/shenma 'what' + 区别 qubie 'difference' + a]	7
	[什么玩意儿/啥玩意儿/嘛玩意儿 shenme wanyier/sha wanyier/ma wanyier 'What's the matter'+ a]	7
	[什么关系/啥关系 shenme guanxi/sha guanxi 'What is the relationship' + a]	6
	[咋回事 za hui shi 'What's the matter' + a]	5
	[这些事/这种事 zhe xie shi/zhe zhong shi	5

	'these kinds of things/this kind of thing' + a]	
	[什么样 shenme yang 'what appearance' + a]	5
	[升级 shengji 'upgrade (of some softwares)' + a]	6
	[什么歌 shenme ge 'what song(s)' + a]	5
	[什么书 shenme shu 'what book(s)' + a]	5
MOCK QUESTION AS GRUMBLE AND COMPLAINT	**[VP + a]**	**51**
	[坑爹 kengdie 'to deceive father (generally means "to deceive")' + a]	16
	[闹哪样 naonayang 'what the hell is going on?' + a]	8
	[神经病 shenjing bing 'neuropathy (an insulting word)' + a]	8
	[有完没完 youwanmeiwan 'endless' + a]	7
	[有毛病 you maobing 'have problem(s)' + a]	6
	[NP + 何在 hezai 'where' + a]	6
Grand Total		**760**

The SHOWING CONCERN TO HEARER Schema

Under this schema are constructions that are used to insistently request information as a way to show the speaker's concern to the hearer. One scenario is to request explanation. In such cases, 事 *shi* and 样 *yang* are two frequent morphemes in these collocates of I-*a*. The DMC Dictionary's entries of *shi* revolve around denotations of "matter" or "incidence". It occurs in collocates such as 什么事 *shenme shi* 'what matter' 怎么一回事 *zenme yi hui shi* 'What's the matter' in the corpus. In fact, 7 out of 18 constructions under this schema constitute a question of "What's the matter", as is reflected by my glossing in **Error! Reference source not found.**. Example (57) is an illustration,

(57) 究竟 是 什么 事 啊？
jiujing shi shenme shi a
on.earth copula what matter A

‘What on earth is the matter?’

The DMC entries of 样 *yang* are generally about “appearance” or “situation.” It occurs in collocates such as 什么样 *shenme yang* ‘what appearance’ and 怎么样 *zenme yang* ‘how’. As is shown in Example (58),

(58) 考试 成绩 怎么 样 啊？

kaoshi chengji zenme yang a

exam results how appearance A

‘How is the exam? / What are the results of your exam?’

Such constructions with *shi* and *yang* are similar with other collocates in that constructions under this schema are all used to invite other interlocutors to offer an explanation. Notably, the construction [真的假的/真的还是假的 *zhende jiade/zhende haishi jiade* ‘true or false’ + *a*] is also categorized under this schema. This particular construction’s function is to express doubt over the state of affairs. I think the expression of doubt naturally invites the other party to explain or defend his or her position. As is shown in Example (59),

(59) 你 说 你 没 玩 过 是 真 的 还是 假 的 啊？

ni shuo ni mei wan guo shi zhen de haishi jia de a

you say you not play EXP copula real DE or fake DE A

‘You said you did not play it. Is this real or not?’

Another scenario is to request details about a plethora of real-life situations. For example, the three collocates, 怎么办 *zenme ban* ‘what to do’ 肿么办 *zhongme ban* ‘what to do’ 咋办 *zaban* ‘what to do’ are all question words regarding possible solutions to a problem, such as in Example (60),

(60) 那 你 下周 该 怎么 办 啊？

na ni xiazhou gai zenme ban a

then you next.week should how do A

‘Then what should you do next week?’

The collocates 几时 *jishi* ‘when’ 多久 *duojiu* ‘how long’ are enquiries of time. As is in Example (61) and Example (62),

(61) 工资单　几时　发　啊？

gongzidan jishi　fa　a

payroll　when send A

‘When is your payroll sent? / When do you receive your payroll?’

(62) 你去　玩　多久　啊？

ni　qu wan duojiu　a

you go play how.long A

‘How long will you play?’

The collocate 哪里 *nali* ‘where’ is asking about location. As is in Example (63),

(63) 捐　到　哪里　啊？

juan　dao　nali　a

donate reach where A

‘Where (does my) donation (go)?’

There are also other “encyclopedic-style” collocates ranging from 歌 *ge* ‘song’ to 升级 *shengji* ‘upgrade (of a software)’. As is shown in Example (64),

(64) 唱　过　什么　歌　啊？

chang guo shenme ge　a

sing　EXP what　song A

‘(What song(s)) has X sung?’

I call it “Showing Concern to Hearer” to illustrate the point that the a-preceding utterances can already serve as questions. With *a* tagged, the whole utterances adds an extra layer of meaning, that the speaker concerns about the hearer so that he or she asks a question in this insistent manner.

The MOCK QUESTION AS GRUMBLE AND COMPLAINT Schema

Paired with semantically forthright and aggressive collocates such as 有毛病 *you maobing* ‘something is wrong (derogative)’ 闹哪样 *nao nayang* ‘What the hell are you doing’ 神经病 *shenjingbing* ‘neuropathy (derogatively, stupid)’, I-*a* occurs in such constructions that constitute grumbles and complaints, masquerading as questions.

As is shown in Example (65),

(65) 你 这 人 是 不 是 有 毛病 啊？

ni zhe ren shi bu shi you maobing a

you this person copula not copula have problem A

'Do you have any (mental) problems? (You should or shouldn't have done something.)'

In a nutshell, the majority of interrogative *a* tokens (709 out of 772) occur in constructions that have question words. Overall, I-*a* constructions perform two functions: showing concern to hearer and grumbling or complaining.

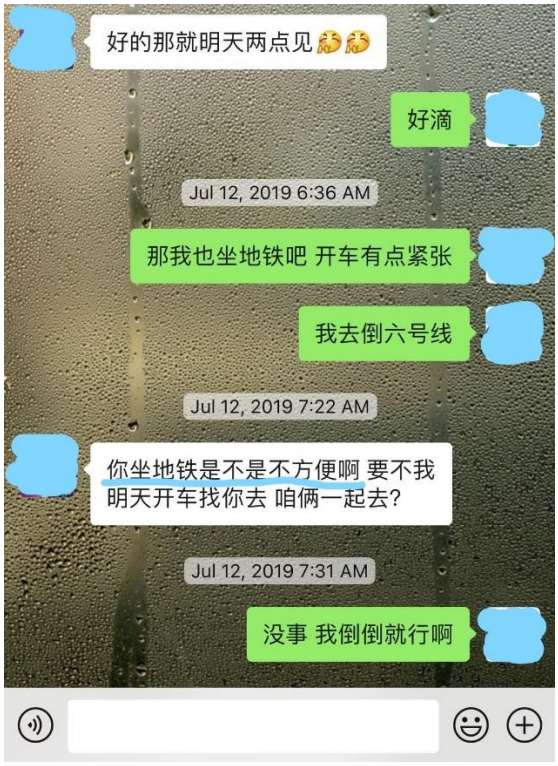
6.5. Constructions of I-*a* in WeChat Data

Table 26: Distribution of interrogative *a*-tagged utterances in WeChat data

Function	Form	Frequency
Genuine Question (Inquiry about Information)		**12**
	• [WHQ + a]	10
	• [YNI + a]	1
	• [AQ + a]	1
Confirmation Request		**10**
	• [YNI + a]	9
	• [AQ + a]	1
Genuine Question (Inquiry about Reason)	**[WHQ + a]**	**4**
Expression of Surprise		**2**
	[WHQ + a]	1
	[YNI + a]	1
Showing Disagreement	**[WHQ + a]**	**1**
Expression of Frustration	**[WHQ + a]**	**1**
Grand Total		**30**

The WeChat findings for I-*a* are very similar to those for I-*ne*, as is shown in Table 26. It occurs in both WHQ tokens and YNI ones. 14 out of 17 WHQ tokens are genuine questions, meaning they are used as questions to ask for either information or reason

for something. One thing that is significant is that 9 out of 11 YNI tokens are used as a confirmation request, meaning that they are used by the speaker to solicit a confirmation of what has been previously chatted about from other interlocutors. There are two tokens of alternative question (shortened as "AQ" in the table), which is an "either-or" type of question. Let's see one example of *a*-tagged YNI as solicitation of confirmation in Figure 28.



B: OK, see you tomorrow at 2 then. [Two emojis meaning applause]	
	A: Okay.
	A: Then I'll take the subway. It might be challenging to drive there.
	A: I'll transfer at Line 6.
B: Is taking subway inconvenient for you? How about I drive and pick you up?	
	A: Oh no, I'm fine with subway. I'm comfortable with the transfer.

Figure 28: *ni zuo ditie shi bu shi bu fangbian a* 'Is taking subway inconvenient for you?'

In this conversation, Speaker A and B were talking about the imminent meeting between them. Speaker B said that he or she could take the subway to reach the place for the meeting while A thought that taking subway might be relatively inconvenient and requested B to confirm the possible inconvenience by saying 你坐地铁是不是不方便啊 *ni zuo ditie shibushi bu fangbian a* 'Is taking subway inconvenient for you?', and to perhaps reconsider the choice of taking the subway.

6.6. Interim Summary of I-*a*

The corpus data findings show that I-*a* constructions are mainly used to insistently

interrogate the hearer. Specifically, they are used as a way to show concern to the hearer or as mock questions which serve as grumbling or complaint.

Note that I-*ne* and I-*a* constructions tend to attract words from different registers. I-*ne* constructions notably host what I call "nouns of complexity", which is reproduced here as in Table 27,

Table 27: Nouns of Complexity used in I-*ne* constructions

Freq	Stat (MI)	Collocate
17	3.31599	关系 guanxi 'relationship'
16	3.7006	意义 yiyi 'significance'
13	3.0333	原因 yuanyin 'reason (noun)'
12	3.5959	后来 houlai 'later'
11	3.41378	想法 xiangfa 'thought (noun)'
8	3.84744	看法 kanfa 'opinion'
8	3.7479	作用 zuoyong 'impact (noun)'
7	3.76171	特征 tezheng 'feature (noun)'
6	3.15229	现象 xianxiang 'phenomenon'
5	4.16936	处置 chuzhi 'to handle'
5	3.97672	区别 qubie 'difference'
5	3.39176	角色 juese 'role'
5	3.06983	命运 mingyun 'destiny'

As analyzed, answers to such interrogatives are likely to be sophisticated, discourse-level responses. By contrast, I-*a* constructions do not show this tendency.

CHAPTER VII

DISCUSSION

The discussion covers the semantics and pragmatics of the three target SFPs. As terms in linguistics, "semantics" is "the general and more abstract meaning of a form" while "pragmatics" is "what speakers do with that meaning in context, which has to do with the accomplishment of social goals, as well as the organization of information in discourse" (Jing-Schmidt, 2017, p. 34).

7.1. Discussion of the Findings of *bei*

Semantics

Schemas of *bei* are represented in Figure 29:

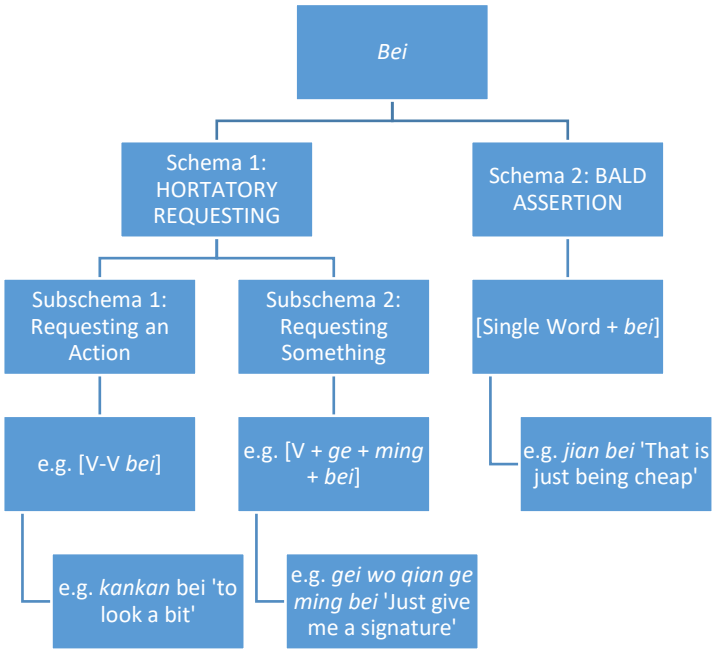


Figure 29: Schemas of *bei*

Both the HORTATORY REQUESTING construction and the BALD ASSERTION construction are semantically derived from the notion of restrictivity. The notion of restrictivity is defined by Xiang (2011) as "a subjective representation of a 'constraint reality' where nothing else is possible" (p. 1378). She maintains that the linguistic

encoding of restrictivity is common cross-linguistically. One example cited in Xiang (2011) is English *just* as a discourse particle, discussed in Aijmer (2002):

b: you got a cold

a: no, ***just*** a bit sniffy. Cos I'm- I am cold and I'll be all right once I've warmed up.

The "just" restricts hearers' interpretation of the speaker's having a cold to that the cold is only a mild one. *Just* here has a down toning function.

Contra the English *just*, Xiang (2011) points out that, in some other Southeast Asian languages, this notion is encoded by SFPs. Although previous studies on these SFPs pinpoint various imports conveyed by them, Xiang insightfully believes that all these imports are derived from the notion of restrictivity. Such particles include:

- Mandarin *me*

Chappell (1990) argues that *me* has two main functions. The first, and the major function, is that *me* can be used by the speaker to point out the self-evident logical cause-and-effect connection. *me* under such use expresses obviousness. For example,

(66) Yinwei xin… xin huang *me.* Ta tou -le dongxi
because heart heart anxious ME 3sg steal LE thing
'Because he was feeling upset, after all. He'd stolen something.' (p. 48)

This sentence is from the author's data elicited by the Chinese Pear [or Guava] Stories. The speaker was providing reason for the boy's behavior that the boy was feeling upset because he had stolen something. The use of *me* indicates that his feeling upset is self-evident since theft naturally leads to feeling upset.

The second function is that *me* can be used to convey disagreement on the part of the speaker based on a self-evident situation and this use of *me* has more emotional element in it. For example,

(67) B: Xianzai shou zhei zhong chuguochao yingxiang de ren tai duo le
now suffer this kind go:abroad:trend influence DE people too many LE
'There are far too many people being influenced by the trend to go abroad'

C: Zhe ye shi hao shi me!
this also be good matter ME

'That's something good too!' (p. 55)

B and C were talking about "Chinese Education System." B thought a lot of Chinese people were influenced by the trend of studying overseas. C replied with a *me*-tagged utterance and showed his or her attitude that this was actually a good thing. C's assumption was that many people studying abroad should be self-evidently good. He or she uses *me* to indicate the disagreement with B in questioning the positiveness of this.

We can see that despite *me* in the two cases differ slightly in the import it conveys, the two functions of it are both based on the speaker's opinion of the obviousness (the quality of being self-evident) of the state-of-affair. As Xiang maintains, the notion of "obviousness" is a derivative of the underlying notion of restrictivity.

- Cantonese *je¯* (Chan, 1996, also discussed in Xiang 2011)

Chan argues that the *je¯*-tagged utterances typically have "delimiting, diminutive, or downplaying functions" (p. 14), as in *Ngh baak gan **je¯** ma, haih maih a?* 'It's only 500 catties, isn't it?' and *Yat cheung nghwuih **je¯*** 'It's just a case of misunderstanding' (p. 14).

I think the notion of restrictivity applies to the interpretation the HORTATORY REQUESTING construction, which restricts the addressee's interpretation of the request to that fulfilling the request just needs minimal efforts. Using Chan's words, this construction underscores the "triviality" of the situation (Chan, 1996, p. 14). This sense of triviality is corroborated by previously discussed significant collocates including *yixia*, *ge*, and *dian(er)*, which all revolve around a small quantity and the cooccurring verb reduplication structure which also means "to do something a little bit." We therefore can say that the "triviality" semantics "rubs off onto" (Bublitz, 1995, p. 13, cited from Stempel, 2019, p. 22) the semantics of *bei*, through the lens of the Semantic Prosody theory.

I then further argue that this sense of triviality associated with *bei*-related constructions gives rise to the "being hortatory" overtone. The overtone implies that the speaker of such *bei*-tagged utterances is trying to maximize "the chances of compliance"

from the recipient (HAYANO, 2013, p. 409). Put it differently, the speaker uses the *bei*-related constructions to downplay the efforts that the hearer needs to take so that the hearer will be maximumly compliant with the speaker's request or suggestion.

The same analytical approach can also be applied to *bei* in the BALD ASSERTION construction. Recall that the form of this construction is that one lexical item appears before *bei*. By using this construction, the speaker makes an assertion by juxtaposing a single word with *bei* and therefore restricts the hearers' interpretation of the state-of-affair to what that single word describes.

Here the quantity principle of iconicity proposed by Givón (Givón, 1995; Ji, 2007) can further support my analysis. The quantity principle maintains that:

(a) "A larger chunk of information will be given a larger chunk of code".

(b) "Less predictable information will be given more coding material".

(c) "More important information will be given more coding material". (Givón, 1995, p. 49)

A logical deduction of Axiom (a) will be that "A smaller chunk of information will be given a smaller chunk of code". In the BALD ASSERTION construction, only one word appears before *bei*. A single word is a less chunk of code than a phrase, a clause, etc. The structural simplicity of one word therefore suggests a limitation of information.

Connected with previous studies, the notion of restrictivity can explain why Liu et al (2001) come to the conclusion that *bei* "常常表示'道理简单'、'无须多说'的语气" (*Bei* usually conveys the overtone that "the issue is straightforward", "there is no need to bring it up anymore"). This can also strengthen Zhao & Shi (2015)'s argument that the core meaning of *bei* is 应而不愿 'a speaker uses *bei* when he or she doesn't feel like replying but has to do so' (p. 75), thus expressing negative sentiment. The *bei* speaker believes that the state of affair has already been obvious while other conversationalists still bring up the issue. The speaker then may be impatient or reluctant or in whatever similar emotional state. No wonder Liu et al (2001) conclude that 说话者用'呗'时多不太满意 '*Bei* speaker usually uses this SFP when he or she is unsatisfied)'.

The [XP + *jiu* + XP + *bei*] construction attested from the response tokens in WeChat data also fits in the notion of restrictivity. Recollect that the form of this construction is that same or extremely similar structure appears before and after the adverb *jiu*. The meaning of this adverb in such use is 放在两个相同的成分之间，表示容忍 'To be put between two identical components to mean tolerance', as provided by the DMC Dictionary, wherein the speaker uses this construction to restrict addressee's interpretation to just one option which is denoted by the structure appearing before or after the adverb *jiu*.

For the manipulative tokens in the WeChat data, those *bei*-tagged utterances that express a suggestion connotes a self-perceived high epistemic stance by the speaker vis-à-vis other interlocutors. My interpretation is that by saying so the speaker offers his or her suggestion as the only one plausible solution of the state of affair, which is another instance of restrictivity.

However, what distinguishes my analysis from Xiang's analysis and what I need to emphasize again is that the notion of restrictivity is not expressed by *bei* alone, but by its constructions. *Bei* should be always considered as part of constructions, according to the usage-based constructionist approach.

Pragmatics

My data reveal that *bei*-related constructions can be used not only in declarative utterances as responses (as is maintained in all previous *bei*-related studies to the best of my knowledge), but also in manipulative utterances, notably in suggestion and request.

In constructions under the HORTATORY REQUESTING schema, *bei* as well as the significant collocates, *yixia*, *ge*, *dian(er)* and reduplicated verbs are illocutionary force indicating devices (IFIDs), which are linguistic devices that can indicate the illocutionary force a speaker is performing in an utterance (Searle, 1969, p. 30). An example provided by Searle is "I promise" as the illocutionary force indicator of the sentence "I promise that I will come" whereas "that I will come" is the propositional

content indicator (p. 30). For the case of *bei*, this SFP and those cooccurring collocates are also IFIDs, which indicate that mitigated request is the illocutionary force. *Bei* and these of its collocates significantly downtones the imperious tone in requests.

Here are more details. The two classifiers *ge* and *dian* are actually not compulsory in the construction [V + CL + N + *bei*]. It is legitimate to say V + *ming* (instead of V + *ge* + *ming* + *bei*) and *gei* + *jianyi*/*yijian* (instead of *gei* + *ge*/*dian(er)* + *jianyi*/*yijian* + *bei*) in utterances. Indeed as the experiments in Erbaugh (2013) show, almost half of the sentences generated in real speech by their research participants did not have a classifier beside them, which goes against the assumption that a Mandarin noun must be preceded by a classifier.

Erbaugh then concludes from her experimental findings that "discourse, not the noun, controls the classifiers" (p. 101). Arguably then the two classifier's case in my study shows us that to use a classifier or not, is a question considered by speakers taking into account discursive factors.

Let's begin with the classifier *ge*. The DMC Dictionary explains its function as "用于约数的前面，使句子显得语气轻松、随便" 'used in front of a number which is not that exact so as to make the sentence sound lighthearted and casual'. The *ge* in *bao ge ming* 'to register', *zhu ge ming* 'to register', *qian ge ming* 'to give a signature', etc., is an instance of such usage. For example, for the examples regarding "to request signature", in real-life situations the speaker may request just one signature, let's say from a celebrity, or multiple signatures from a person when that person may have to sign his or her name in multiple places for some paperwork. The same is true for *bao ming* 'to register', *zhu ming* 'to register', etc. because myriad possibilities exist for the quantity of items being requested in real-life scenarios. When registering, a person may fill in his or her name for just once or for many times in one setting. Therefore, the use of *ge* indexes an uncertain yet small quantity. Those utterances with *ge* are arguably intended to soften the tone of the whole sentence in that it conveys the import that the speaker is only requesting something in small quantity and therefore the request is not very imposing.

Similarly, *dian* is explained by The DMC Dictionary as 表示少量 'to indicate small quantity.' The use of this classifier indexes small quantity of the items being requested. It is clear then the speaker deliberately uses the two classifiers to highlight the smallness of quantity of items being requested in those utterances. Clearly, both *ge* and *dian* are diminutives that inherently have the function of mitigation. Therefore, the existence of the two classifiers as IFIDs in constructions indicates that mitigated request is the social action achieved by those constructions. They are functional equivalents of the reduplicated verb structure that marks the delimitative aspect. Overall, constructions under the HORTATORY REQUESTING schema express mitigated requests that make it easier for the hearer to accept the requests. In other words, the speaker is implying that "since what I am requesting is a small favor, please just accept the request."

My WeChat data especially suggests that *bei*'s presence signals a termination of an ongoing sequence. This means its presence has this discourse organizational and textual function. This finding sheds some light on the discourse management functions of *bei*, which has been overlooked by previous literature.

Previous studies predominantly focus too much on the quality of the message itself in *bei*-tagged utterances, such as "obviousness", "not-a-big-deal attitude", etc. My findings of *bei* shed new lights on its interpersonal and discourse organization functions because I have identified constructions that express nuanced illocutionary forces.

7.2. Discussion of the Findings of *ne*

Semantics

My argument is that *ne*'s presence registers a need to update epistemics of interlocutors. This argument is more elegant and effective than the treatment of this SFP in prior literature. In my theory, there is no need to postulate that *ne* has several variants such as topic marker *ne* or evaluative marker *ne* (Li, 2006, pp. 7-21). Instead, *ne* is always part of constructions. The high-frequency patterns ("pathbreakers") of those constructions are used to register a need to update interlocutors' epistemics. In other

words, we cannot say *ne* itself has some certain quality of message because the message is conveyed by the utterance preceding *ne*. With ne tagged, such utterances draw the hearer's attention to epistemics in conversations.

Let's return to the NGAUPs (now-generally-agreed-upon positions) reached by previous studies and this comparison will show my new contribution to the understanding of *ne*.

- NGAUP 1: *Ne* is used in conversations

My data corroborates this NGAUP again. As my literature review shows, this is also true of SFPs in general.

- NGAUP 2: *Ne* can be used in both declaratives and interrogatives

My data supports this belief, as this SFP is attested in both sentence types. Also, the identified schema INSISTENT INTERROGATION confirms one of the DMC Dictionary entries that ne is "used in the sentence-final position of interrogatives (wh-questions, alternative questions, A-not-A questions) to convey an overtone of reminding and strong enquiry", as an insistent interrogation is a strong enquiry.

- NGAUP 3: *Ne* is used when the speaker wants to say something vis-à-vis what the hearer has said or has in mind

My data suggest that actually a speaker uses *ne*-related constructions to update interlocutors' epistemics. Epistemics is the common denominator of all previous accounts of *ne* semantics, be it claim, expectation, or belief on the part of the hearer, or the unusualness in the speaker's words. Also, this belief actually evinces again that PPs are hearer-oriented.

- NGAUP 4: *Ne* marks the progressive aspect

This observation made by previous scholars is correct in that this particle does appear in example sentences that are about interlocutors' current activities. Consider the fact that one notable function of NI-ne utterances in WeChat data is "drawing attention to speaker current situation." However, my analysis reveals that the *ne*-tagged utterances are about the epistemics of current activities. It should not be treated as an aspect marker because this progressiveness can be expressed without the SFP and by

the pre-SFP utterance. For example, the progressiveness can still be conveyed by the without-*ne* counterparts of all three sentences provided in the DMC Dictionary (counterpart sentences without *ne* are presented in the brackets following the free English translations):

(68) ta zai jing bian dashui ne
she at well side fetch.water NE
'Look, she is fetching some water at the well.'
(ta zai jing bian dashui)

(69) bie zou le, waimian xia zhe yu ne
NEG walk LE, outside down IMP rain NE
'Don't leave now. You see, it's raining outside.'
(bie zou le, waimian xia zhe yu)

(70) Lao Zhang, men wai you ren zhao ni ne
Old a.surname door outside have person look.for you NE
'Zhang Sir, someone is looking for you. S/he is waiting outside, right now.'
(Lao Zhang, men wai you ren zhao ni)

This of my viewpoint is corroborated by Han (1988)'s observation. Han likewise maintains that *ne* is not a progressive aspect marker but "adds (-s added by me) an aura of conviction to the tone of the sentences" (p. 26). Han used the word "conviction", which is about speaker's firm belief and is therefore well compatible with my theory that *ne* revolves around epistemics. My viewpoint challenges Zhu (1984)'s proposal that *ne* has two variants: ne_1 marks aspect and ne_2 expresses interrogation. I have shown that there is no need to make this distinction; *ne* is not an aspect marker. The understanding that *ne* is an aspect marker is a result of overly emphasizing the message of pre-SFP utterances while overlooking *ne*'s interpersonal functions.

Pragmatics

Schemas of *ne* are represented in Figure 30 and Figure 31, NI-*ne* and I-*ne* respectively:

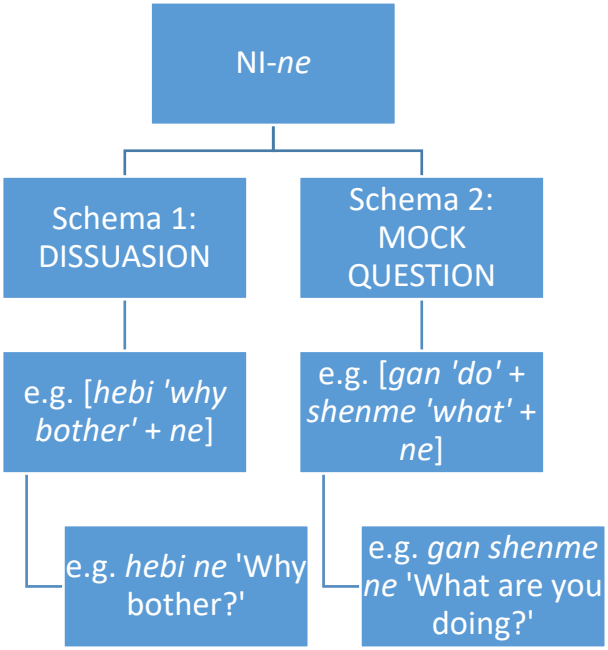


Figure 30: Schemas of NI-*ne*

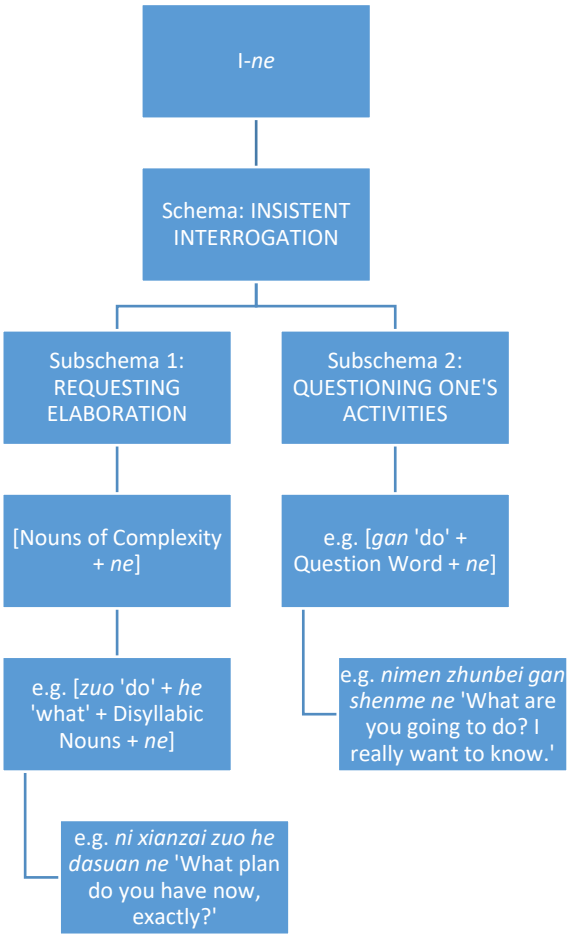


Figure 31: Schemas of I-*ne*

The pragmatic functions of both NI-ne and I-ne parallel with Wu (2005)'s belief

that the core function of ne is to engage the hearer to negotiate Common Ground. This is particularly true of I-*ne*-tagged utterances, which invite elaboration or explanation on the part of other interlocutors. This illocutionary force is best indicated by question words and what I call Nouns of Complexity as IFIDs. On the other hand, my findings do not provide any support of the belief upheld by previous scholars that *ne* marks significance or contrast of the messages expressed in discourse.

As for the NGAUP 5, i.e. *Ne* is used due to a consideration of politeness by softening the tone, my corpus-based findings cannot support this belief. Corpus cannot empirically define and test the notion of "soften."

7.3. Discussion of the Findings of *a*

Semantics

NI-*a* constructions fulfill much more functions than any other SFPs discussed in the current study, as it has the most significant collocates revealed by the corpus analysis. Such a wide range usage of NI-*a* has been has already been noticed by Li (2006).

A greatly embodies the "flexibility and multifunctionality" of SFPs, as previously mentioned in its literature review. Constructions relevant with *a* show the highest type frequency, as I have identified the most amount of significant collocates for it from the corpus. It seems impossible to pin down a central, abstract meaning, or at least some central tendencies of *a*-related constructions, as its semantics. This finding corroborates Li (2006)'s statement that "As for *a*, we have shown that it does not express any specific meaning, but is mainly pragmatically driven" (p. 53). This vagueness in semantics might also be attributed to its maximal compatibility with all kinds of utterances, as has been mentioned that it can be used in all five major sentence types in Mandarin.

Schemas of ne are represented in Figure 32 and Figure 33, NI-*a* and I-*a* respectively:

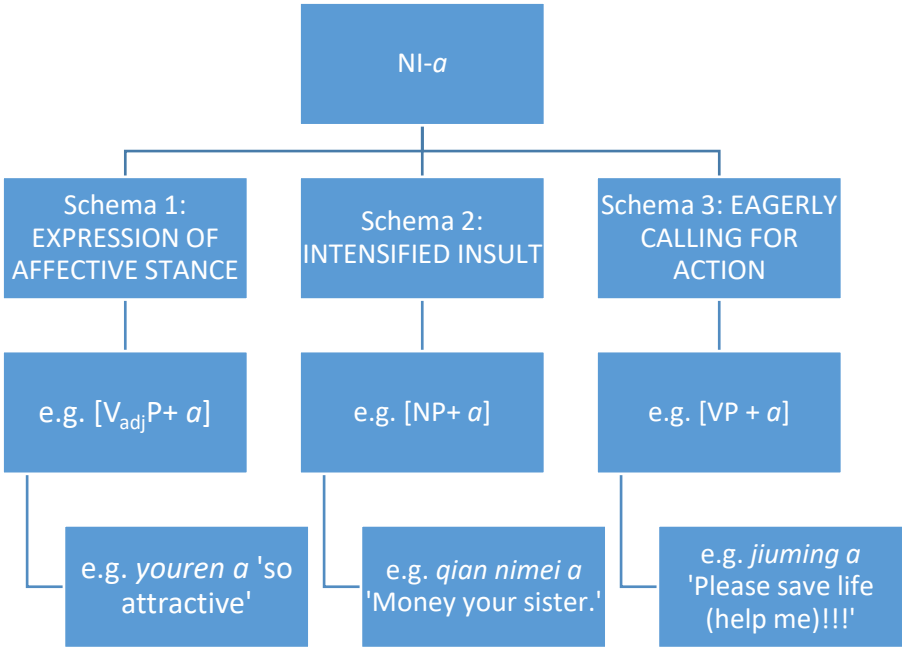


Figure 32: Schemas of NI-*a*

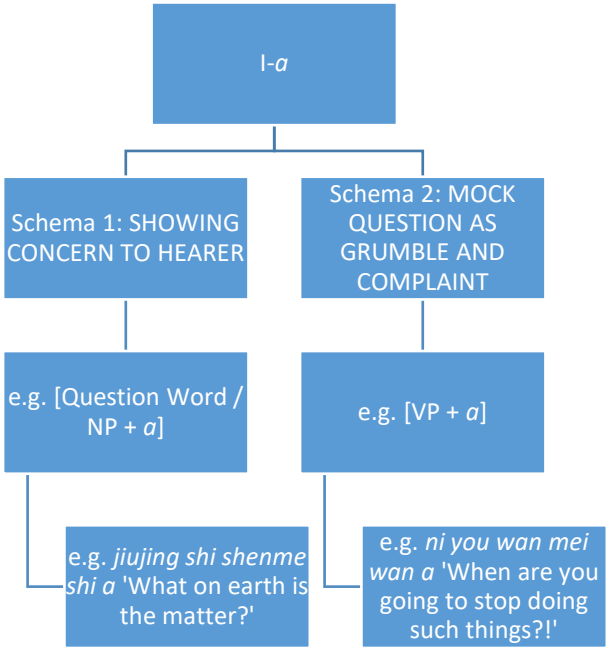


Figure 33: Schemas of I-*a*

Still, it is well-neigh impossible to find any common denominators, but at least we have captured some high-frequency exemplars as "pathbreakers." This can be a good starting point for future studies adopting a usage-based constructionist approach. This

actually demonstrates the merit of this approach. Because it is usage-based, I can still find exemplars for *a*, even though it is so widely used in conversations in such a flexible and multifunctional way.

Perhaps one pertinent study is Lee-Wong (1998). Her study focuses on SFPs *ba*, *a*, and *ne* used in requests. She argues that in requests *a* has the function of signaling "informality and casualness" (p. 396). This theory can probably explain why *a* in my data notably cooccurs with neologisms such as 你妹 *nimei* 'your younger sister (an insulting word)' 你妹夫 *nimeifu* 'your younger sister's husband (an insulting word)'' 坑爹 *kengdie* 'to deceive father (generally means "to deceive")' in NI-*a*-tagged utterances. These words are not likely to appear in formal register. Lee-Wong and my observations thus point to the direction that SFP *a* likely to occur in informal setting. This observation is supported by that fact a is three times more frequently used in the Conversation subcorpus of BCC while ne is three times more frequently used in the Newspaper and Magazine subcorpus, as is shown in Table 28,

Table 28: Frequencies of *ne* and *a* in subcorpora of BCC

SFP	Freq in the Conversation Subcorpus	Freq in the Newspaper and Magazine Subcorpus
ne	1214454	215769
a	3973951	70098

Such insights can only be gained from a corpus-based study while cannot be realistically gained from studies that are limited by dataset sizes, such as conversation analysis studies.

Also, recall the five schemas of *a* that I have identified from corpus data:

- Expression of Affective Stance
- Intensified Insult
- Eagerly Calling for Action
- Showing Concern to Hearer
- Mock Question as Grumble and Complaint

All five of them are significantly affectively loaded, some of them being highly

emotional, such as the Intensified Insult schema. This finding endorses the DMC Dictionary entry that *a* can "used at the end of a declarative sentence to add a flavor of emotion to the sentence." This amounts to say that a-tagged utterances are suitable for daily colloquial uses. This stands in sharp contrast with I-*ne*, which tends to cooccur with nouns of complexity or disyllabic nouns, the two being a feature of formal register. This of my finding can explain why it is well-neigh impossible to capture the semantics of *a*: it contributes the affective stances expressed by its host utterances while does not contribute to messages therein. This corroborates one property of pragmatic particles (PPs) in general, as I have reported in my literature review, that PPs do not contribute to the propositional content of an utterance.

Furthermore, my findings of *a* echo previous accounts of *a*'s polarized pragmatic functions that have indeed been reported by previous research. Scholars have found that it can interestingly appear in utterances that either reduce the forcefulness or register a surprise or unexpectedness. I find that *a* occurs in utterances that may encourage addressees to take further actions that are beneficial to the speaker, such as call for action while also occurs in ones that may possibly incur non-beneficial actions, such as insult.

7.4. General Discussions

This study offers insights into the usage and functions of SFPs. The case studies in my dissertation show that the meanings of SFPs are not elusive, but empirically observable and tractable as long as we analyze substantial data within theoretical frameworks that can adequately explain the data.

For Research Question 1, viz. "what are the distribution patterns of the three SFPs?", my analyses show that *bei*, *ne*, and *a* significantly cooccur with collocates that can be categorized into different schemas based on functional similarities.

For Research Question 2, viz. "are the functions of SFPs specific to the constructional contexts in which they are used? If so, what are those functions?", my analyses confirm that the functions of SFPs are construction-specific. Recall that the

usage-based constructionist approach uncovers the fact that when NI-*ne* is used in construction [Question Word + *ne*], the function is dissuasion, while when it is used in [嗯 en 'Yes/OK' + *ne*] or [是 shi 'Yes/OK' + *ne*], the function is confirmation. The two opposite-valenced functions can only be associated with constructions. The same is true of the discovery that a can be used in both the EAGERLY CALLING FOR ACTION schema and the INTENSIFIED INSULT schema.

Pragmatically speaking, SFPs are IFIDs that indicate illocutionary force in utterances. They do not stand alone in discourse but belong to constructions which have certain forms that can achieve certain social actions as functions. I think SFPs are actually IFIDs also because their identity as IFIDs conform to the fact that SFPs are hearer-oriented, which is a keyword I have found from my mini-corpus of previous literature on pragmatic particles. In other words, IFIDs are deployed to indicate the illocutionary force of an utterance for the hearer, thus hearer-oriented, so are SFPs.

Nonetheless, the presence of SFPs signals more nuanced illocutionary forces. For example, as my study reveals, *bei* appears in utterances which are not just requests, but mitigated requests; *ne* appears in utterances which are just questions, but questions inviting an elaborated response; *a* appears in utterances conveying an informal and casual overtone.

A question arises: why speakers need to deploy SFPs to indicate the nuances in the first place? The answer relates to one extraordinary property of human communication, as discussed by Searle (1969):

> "If I am trying to tell someone something, then (assuming certain conditions are satisfied) as soon as he recognizes that I am trying to tell him something and exactly what it is I am trying to tell him, I have succeeded in trying to tell it to him. Furthermore, unless he recognizes that I am trying to tell him something and what I am trying to tell him, I do not fully succeed in telling it to him. In the case of illocutionary acts we succeed in doing what we are trying to do by getting our audience to recognize what we are trying to do". (p. 47)

In other words, Searle tells us that the illocutionary force in a speaker's utterance is and

can only be achieved by a hearer's understanding it. His example is the expression "Hello" in English, whose illocutionary force as greeting can only be achieved when a hearer recognizes that "Hello" is a greeting (p. 49).

On the other hand, the constructionist approach can not only reveal the nuances, but also provide a much more unified account of SFPs. Take SFP *a* as an example. It can occur in utterances that perform opposite functions, such as to call for action vs to insult. The constructionist approach can comfortably explain that this paradoxical phenomenon is caused by the same SFP cooccurring with different collocates in different constructions. When it occurs with *jiuming* 'save life' in the [*jiuming* + a] construction, the constructional function is calling for help, which is beneficial. When it occurs with *nimei* 'your sister' in the construction [*nimei* + a], the constructional function is insult, which is non-beneficial to the speaker. This is the same benefit of adopting a constructionist approach to explain the renowned example sentence "John sneezed the napkin off the table." "sneeze" is no longer needed to be explained as being "transitive" just in this sentence while our general impression is that this verb should be "intransitive." The sentence is just another instance of the English caused-motion construction, i.e. [Subject + Verb + Object + Directional]. Its meaning is associated with the constructional meaning, not with the verb "sneeze." In addition, this constructional way of thinking can probably be applied to the explanation of *ba*, as I mentioned in the Overview section, which occurs in both utterances sounding confident and at other times utterances sounding less confident. My hypothesis is that there are Schema 1 "confident" and Schema 2 "less confident" for *ba*. When it cooccurs with collocates with a "confident" reading in constructions, the utterances sound confident. When it cooccurs with collocates with a "less confident" reading, the utterances sound less confident. This awaits future research to test it. However, such a constructional account of *a* and *ba* now certainly rules out the necessity to say that they are "elusive."

Also responding to Simpson's summarization of the tension surrounding SFP research that "whether it is necessary to **assume** (bolding mine) multiple, homophonous SFPs in certain instances, each particle having a different function, or whether a unique

particle can be posited to exist with an underlying, broad meaning that may be applied in a range of different contexts", I maintain that linguists do not need to assume any more if they adopt the usage-based constructionist approach because "what-you-see-is-what-you-get." As my findings show, we can find constructions as exemplars for each SFP by schematization of corpus data. We can therefore know how an SFP is prototypically used in natural language to fulfill what functions. No more theorization of its nature and usage is needed.

This approach can also avoid the tension between a maximalist approach and a minimalist approach[35]. The usage-based constructionist approach provides with exemplars as usage patterns for SFPs. The exemplars are high-frequency patterns and therefore reside in the center of speakers' cognitive representation. Therefore, exemplars do not have to cover each and every use of an SFP, as a maximalist would do. Linguists will not be confused by all the nuts and bolts of uses. On the other hand, exemplars do cover prototypical usage patterns of SFPs, thus saving linguists the ordeal to have to pick out the very "core function," as a minimalist would do. This approach can therefore address the multifunctionality issue with the description and analysis of SPFs. This said, the study also raises questions about Construction Grammar. First, if our linguistic knowledge is a continuum from the representation of concrete tokens to the representation of abstract schemas, what level of abstraction is appropriate for an analysis of constructions that are in between? Second, when it comes to a bound morpheme such as an SFP, what is the relationship between the meaning of the construction to which the SFP is attached and the larger holistic utterance meaning derived from combination of that construction and the SFP? Finally, Goldberg (2019) talks about memory traces of usage leading to cognitive representations of a construction, it remains unclear how to formalize the different levels of representation of a construction biased on single usage events.

This study also sheds some lights on speech acts studies in general. Sadock (2004)

[35] As discussed earlier, linguists adopting the maximalist approach list different senses of a single SFP, on the basis of various types of utterances, while linguists adopting the minimalist approach come up with a core, central function of a single SFP and advocate that other functions are derived from it.

points out that speaker's intention lies at the heart of speech acts, despite the fact that there are definitely conventional expressions that can be used to perform speech acts, for example, to greet some person by saying "Hi!" (p. 53) "Oh, I love chocolates" can be a thank-you token said to a guest, but Sadock tells us that it is not conventionally so. Therefore, the study of speech acts needs to strike "the proper balance between convention and intention" (p. 53). The balance, as is shown by my study, can be achieved by adopting a constructional view of SFP usage. On the one hand, we can notice that speakers use SFP-related schematic constructions as conventional templates to perform social actions. SFPs and their significant collocates fill some slots in constructions. On the other, to precisely express themselves, speakers also use specific lexical items to fill in those open slots in those SFP-related constructions. That is to say, by using such constructions, speakers simultaneously conform to conventions in daily exchanges and fully express their communicative intents under the given interactional circumstances.

7.5. Limitations

This study has multiple limitations.

First, the qualitative analysis of the data may be subjective because the annotation was done by the author alone as the single researcher of the present study. Therefore, no intercoder reliability could be obtained, which would have been possible if the research involved a second coder.

Second, the analysis of the sub-patterns and exemplars is laden with uncertainties as to what level of abstraction is ideal for a usage-based analysis. However, this is not a unique problem for my study. This is a challenge for construction grammar in general.

Third, WeChat data should have included sociolinguistic information such as gender, social hierarchy, etc. of the interlocutors so that a deeper understanding of contexts of SFP use can be achieved.

Fourth, future studies should more extensively investigate modal adverbs and other contextual cues collocating with these schemas identified in the corpus data, which

could reduce the uncertainty of the semantic coding.

CHAPTER VIII

CONCLUSION

The three case studies in my study show that SFPs are functionally dependent on the larger constructional context in which they are used, which is in line with the axiom "to know language is just to know its constructions" as the central tenet of Construction Grammar (Hilpert, 2014b). My study cautions against the understanding of "sentence-final particle" as having meanings independent of the constructions in which they are used. As a constructionist, I believe they are not isolated particles, but part of some larger constructions and their functions are best understood holistically as part of those constructions.

Theoretically speaking, my dissertation shows that Construction Grammar and Semantic Prosody Theory mutually complement each other. First, as discussed in Chapter 2, both theoretical traditions notice that constructionists and neo-Firthians both uphold that the basic units of language are "chunks." My SFP case studies reveal that the meaning and usage of SFPs are not "elusive" at all if viewed them as components of constructions. Second, both theories are usage-based. Following this theoretical guidance, my findings of the three SFPs all come from actual usage events and therefore not blurred by my own limited intuition of their meanings and usages. Third, both theoretical traditions emphasize the function side of linguistic expressions. Sinclair maintains that semantic prosody of a lexical item is its function in communication. Constructionists maintain construction as form-function pairing. My study shows that SFPs are indeed function words, as upheld by Lin (2007, p. 228), but more accurately, they are part of constructions that have those functions. Finally, the two theoretical traditions provide mutual support methodologically. Neo-Firthians teach me to treat collocation as a functional indicator of meaning. Collocation in this study make constructions "palpable and accessible" (Jing-Schmidt, 2015, p. 14) to me because it helps me to identify constructions. The usage-based theoretical frameworks therefore informed my methodology in that I combined corpus analysis with social media

conversation-based data analysis. The results are well organized thanks to construction's schematicity, ranging from schema/subschemas to constructs. My study of SFPs therefore benefits from this methodological compatibility, which yield collocation-driven, construction-in-shape, WeChat-data-further-elaborated results.

Pedagogically speaking, my study provides language teachers a way to elucidate the multifunctionality of SFPs. As Jing-Schmidt (2015) explains, language teachers can resort to corpus data and obtain constructions of SFPs, especially those high-frequency exemplars as "pathbreakers." Students can then "see" the patterns of use of SFPs and "get" it. This way, "teaching becomes less arbitrary, and more coherent and effective" (p. 13).

APPENDIX A

This section contains all collocate displays that are used for the three SFPs. The tables follow this order: *bei*, non-interrogative *ne*, interrogative *ne*, non-interrogative a, and interrogative a. Within each particle, the tables follow this order: 1L, 2L, and 3L.

Collocate displays of *bei*

Table 29: Collocate display of 1L of *bei*

Freq	Stat (MI)	Collocate
675	4.58165	背 *bei* 'to recite'
192	3.84145	一下 *yixia* 'one time'
100	3.36083	看看 *kankan* 'to look a bit'
69	4.06336	试试 *shishi* 'to try a bit'
34	3.48482	贱 *jian* 'cheap; ignoble'
28	4.68276	屄 *bi* 'cunt'
25	4.15669	玩玩 *wanwan* 'to play a bit'
23	4.56246	抠 *kou* 'stingy'
22	3.0761	名 *ming* 'name'
20	4.68276	贫 *pin* 'garrulous'
20	4.1467	钩 *gou* 'hook'

(continued)

Freq	Stat (MI)	Collocate
20	3.96029	建议 *jianyi* ‘advice’
18	3.42642	玩儿 *waner* ‘to play a bit’
17	3.41267	骗 *pian* ‘to cheat’
16	4.0389	意见 *yijian* ‘opinion’
13	4.29573	得了 *dele* ‘That is a deal’
13	3.85963	说说 *shuoshuo* ‘to talk a bit’
13	3.29573	测测 *cece* ‘to test a bit’
12	3.80829	聊聊 *liaoliao* ‘to chat a bit’
12	3.18026	得瑟 *dese* ‘to be cocky’
12	3.13844	追 *zhui* ‘to chase’
12	3.05827	修 *xiu* ‘to fix’
11	4.05473	听听 *tingting* ‘to listen a bit’
11	3.33483	造 *zao* ‘to build’
11	3.18799	脱 *tuo* ‘to take off’
11	3.01291	混 *hun* ‘to muddle through’
10	3.48112	火锅 *huoguo* ‘hotpot’
9	3.94579	转转 *zhuanzhuan* ‘to walk around a bit’
9	3.04533	熊 *xiong* ‘to be coward’
9	3.04533	再说 *zaishuo* ‘to talk about it again’
8	4.36083	膀 *bang* ‘limb’

(continued)

Freq	Stat (MI)	Collocate
8	4.22333	拉倒 *ladao* ‘Let’s don’t talk about it anymore.’
8	3.22333	想想 *xiangxiang* ‘to think a bit’
7	4.49011	飚 *biao* ‘to contest with each other’
6	4.46036	直说 *zhishuo* ‘to speak it out honestly’
6	4.26772	慢慢来 *manmanlai* ‘to take your time’
6	3.46036	唠嗑 *laoke* ‘to chat’
6	3.09779	抵制 *dizhi* ‘to boycott’
6	3.01979	懒 *lan* ‘lazy’
5	4.68276	掉价 *diaojia* ‘to lower one’s status’
5	4.41972	走着瞧 *zhouzheqiao* ‘Let’s wait and see.’
5	4.41972	走人 *zouren* ‘to leave’
5	4.41972	试试看 *shishikan* ‘to try a bit’
5	4.41972	梵 *fan* ‘sancrit’
5	4.41972	唬 *hu* ‘to scare’
5	4.19733	走走 *zouzou* ‘to walk a bit’
5	4.19733	上街 *shangjie* ‘to shop around’
5	4.00469	降 *jiang* ‘to drop; to decline’
5	3.68276	赚钱 *zhuanqian* ‘to make money’
5	3.68276	瞧瞧 *qiaoqiao* ‘to look a bit’
5	3.54525	离婚 *lihun* ‘to divorce’

(continued)

Freq	Stat (MI)	Collocate
5	3.41972	奋斗 *fendou* ‘to endeavor’

Table 30: Collocate display of 2L of *bei*

Freq	Stat (MI)	Collocate
30	3.26772	认识 *renshi* ‘to know each other’
27	3.01138	爹 *die* ‘daddy’
20	3.33226	慢慢 *manman* ‘slowly’
19	4.53837	幽默 *youmo* ‘humorous’
19	4.0727	勇敢 *yonggan* ‘brave’
17	3.91224	温柔 *wenrou* ‘tender’
13	3.29573	测 *ce* ‘test’
12	4.68276	勤俭 *qinjian* ‘thrifty’
12	4.36083	仗义 *zhangyi* ‘generous’
12	3.94579	谦虚 *qianxu* ‘modest’
9	4.15224	勾搭 *gouda* ‘to seduce’
8	3.43483	艺术 *yishu* ‘art’
7	3.24218	互相 *huxiang* ‘mutual’
7	3.03068	秀 *xiu* ‘show’
6	3.68276	耗 *hao* ‘consumption’
6	3.01979	大不了 *dabuliao* ‘Even if something bad happens, we can just …’

(continued)

Freq	Stat (MI)	Collocate
5	4.68276	信念 *xinnian* ‘firm belief’
5	4.19733	忽视 *hushi* ‘overlook’
5	3.41972	透露 *toulu* ‘to disclose’
5	3.09779	唠 *lao* ‘to repeatedly say something’
5	3.00469	老实 *laoshi* ‘honest’

Table 31: Collocate display of 3L of *bei*

Freq	Stat (MI)	Collocate
457	4.01896	背 *bei* ‘to recite’
31	3.68276	晒 *shai* ‘to show something online’
11	3.68276	凑合 *couhe* ‘to muddle through’
10	3.54525	举 *ju* ‘to lift’
10	3.36083	签 *qian* ‘to sign’
10	3.30425	冒 *mao* ‘to pop up’
9	4.68276	miui ‘unknown sequence of English letters’
7	3.03068	露 *lou* ‘to be disposed’
6	4.68276	同类 *tonglei* ‘people or things within the same category’
5	4.19733	手术 *shoushu* ‘surgery’
5	3.09779	软 *ruan* ‘soft’
5	3.09779	尽量 *jinliang* ‘to try ones best’
5	3.09779	反应 *fanying* ‘reaction’

Collocate displays of NI-*ne*

Table 32: Collocate display of 1L of NI-*ne*

Freq	Stat (MI)	Collocate
59	3.89832	何必 *hebi* ‘why bother’
23	3.2713	怎样 *zenyang* ‘how’
17	3.39592	怎么办 *zenmeban* ‘how to deal with something’
16	4.10682	何苦 *heku* ‘why bother’
16	3.92994	上网 *shangwang* ‘to surf the Internet’
16	3.58599	怎么样 *zenmeyang* ‘how’
12	3.35736	意义 *yiyi* ‘significance’
11	4.38937	稿子 *gaozi* ‘draft’
11	3.76789	意见 *yijian* ‘opinion’
11	3.08981	火锅 *huoguo* ‘hotpot’
9	3.34088	兔子 *tuzi* ‘rabbit’
9	3.09987	好听 *haoting* ‘nice sounding’
8	4.46046	干吗 *ganma* ‘What do you want to do?’
8	3.04542	兴奋 *xingfen* ‘excited’
7	4.26781	红烧 *hongshao* ‘braised’
7	3.97831	册 *ce* ‘brochure’
7	3.53085	尾 *wei* ‘tail’
7	3.11581	聊天 *liaotian* ‘chat’
6	4.63038	中餐 *zhongcan* ‘Chinese food’

(continued)

Freq	Stat (MI)	Collocate
6	3.89342	来不及 *laibuji* ‘too late to do something’
6	3.63038	受伤 *shoushang* ‘to get injured’
6	3.51491	作用 *zuoyong* ‘impact’
6	3.04542	责任 *zeren* ‘responsibility’
5	4.14496	较劲 *jiaojin* ‘to contest with each other’
5	4.14496	节操 *jiecao* ‘integrity’
5	4.14496	发愁 *fachou* ‘to worry about something’
5	3.63038	笨 *ben* ‘stupid’
5	3.63038	怀里 *huaili* ‘in one’s hug’
5	3.49288	捞 *lao* ‘to capture’
5	3.49288	开会 *kaihui* ‘to convene’
5	3.36735	亲人 *qinren* ‘relative’
5	3.25187	记忆 *jiyi* ‘memory’
5	3.04542	流星 *liuxing* ‘meteor’
5	3.04542	保暖 *baonuan* ‘to keep warm’

Table 33: Collocate display of 2L of NI-*ne*

Freq	Stat (MI)	Collocate
85	3.02296	干 *gan* ‘do’
12	3.04542	逛 *guang* ‘to shop around’
10	3.49288	何 *he* ‘what’

(continued)

Freq	Stat (MI)	Collocate
7	3.11581	文艺 *wenyi* ‘literature and art’
6	4.21534	对得起 *duideqi* ‘to live up to’
6	3.89342	乜 *nie* ‘what’
6	3.12788	夸 *kua* ‘to praise’
5	3.78239	膝盖 *xigai* ‘knee’
5	3.78239	海底 *haidi* ‘the bottom of a sea’
5	3.36735	周五 *zhouwu* ‘Friday’
5	3.25187	at ‘the English word “at”’

Table 34: Collocate display of 3L of NI-*ne*

Freq	Stat (MI)	Collocate
6	3.40799	国外 *guowai* ‘overseas’
5	3.04542	早早 *zaozao* ‘very early’

Collocate displays of I-*ne*

Table 35: Collocate display of 1L of I-*ne*

Freq	Stat (MI)	Collocate
519	3.06059	什幺 *shenme* ‘what’
67	4.19732	怎幺办 *zenme ban* ‘what to do’
63	3.21422	哪里 *nali* ‘where’
58	3.92781	怎幺样 *zenme yang* ‘how’
28	3.06983	何必 *hebi* ‘why bother’
27	3.32222	办 *ban* ‘do’

(continued)

Freq	Stat (MI)	Collocate
24	3.04993	然后 *ranhou* ‘afterwards’
21	4.65479	玩玩 *wanwan* ‘to play a little bit’
19	3.04474	礼物 *liwu* ‘gift’
17	3.31599	关系 *guanxi* ‘relationship’
16	3.7006	意义 *yiyi* ‘significance’
13	3.0333	原因 *yuanyin* ‘reason (noun)’
12	3.5959	后来 *houlai* ‘later’
11	4.02676	方舟 *nuoya fangzhou* ‘Noah's ark’
11	3.8663	何苦 *heku* ‘why bother’
11	3.41378	想法 *xiangfa* ‘thought (noun)’
10	4.65479	何乐而不为 *hele er buwei* ‘Why not’
10	3.65479	多久 *duojiu* ‘how long’
9	3.36529	变化 *bianhua* ‘change’
8	3.84744	看法 *kanfa* ‘opinion’
8	3.7479	作用 *zuoyong* ‘impact (noun)’
8	3.40686	干吗 *gan ma* ‘Do what’
8	3.40686	合适 *heshi* ‘appropriate’
8	3.33286	状态 *zhuangtai* ‘status’
7	3.76171	特征 *tezheng* ‘feature (noun)’
7	3.21422	反应 *fanying* ‘reaction’
7	3.06983	上网 *shangwang* ‘to surf the Internet’
6	3.78032	做梦 *zuomeng* ‘to dream / to daydream’
6	3.15229	现象 *xianxiang* ‘phenomenon’
5	4.65479	益处 *yichu* ‘advantage’
5	4.65479	心动 *xindong* ‘to start to feel interested in’
5	4.16936	处置 *chuzhi* ‘to handle’

(continued)

Freq	Stat (MI)	Collocate
5	3.97672	区别 *qubie* ‘difference’
5	3.39176	角色 *juese* ‘role’
5	3.27628	女朋友 *nyupengyou* ‘girlfriend’
5	3.06983	命运 *mingyun* ‘destiny’

Table 36: Collocate display of 2L of I-*ne*

Freq	Stat (MI)	Collocate
71	3.22711	干 *gan* ‘do’
41	3.14198	些 *xie* ‘a few’
24	3.33286	何 *he* ‘what’
21	3.83766	屏 *ping* ‘screen’
11	4.02676	诺亚 *nuoya* ‘Noah’
9	3.12428	校 *xiao* ‘school’
8	3.65479	伤害 *shanghai* ‘to hurt’
6	3.91783	意味着 *yiweizhe* ‘to have a certain significance’

Table 37: Collocate display of 3L of I-*ne*

Freq	Stat (MI)	Collocate
43	3.60532	什幺样 *shenmeyang* ‘what outlook’
22	3.59066	刷 *shua* ‘to cover’
14	3.33286	答案 *daan* ‘answer’
11	4.20733	登上 *dengshang* ‘to reach the top of some place’
7	3.21422	嫁 *jia* ‘to marry’

Collocate displays of NI-a

Table 38: Collocate display of 1L of NI-a

Freq	Stat (MI)	Collocate
181	3.19225	起 *qi* ‘up’
143	3.57024	你妹 *nimei* ‘your younger sister (an insulting word)’
109	3.45514	力 li ‘force’
104	4.34978	啊啊 *a a* ‘a reduplication of SFP *a*’
91	3.61524	爹 *die* ‘daddy’
61	3.20811	累 *lei* ‘tiresome’
54	3.57282	无聊 *wuliao* ‘boring’
45	3.55623	抢 *qiang* ‘to grab’
41	3.21171	容易 *rongyi* ‘easy’
28	3.02059	不行 *buxing* ‘Not OK’
27	4.36241	凶 *xiong* ‘fierce’
23	3.76851	救命 *jiuming* ‘to save life’
23	3.58794	堪 *kan* ‘to be able to bear’
22	3.41487	小心 *xiaoxin* ‘small heart (to watch out)’
21	4.04166	了不起 *liaobuqi* ‘great’
20	3.20075	激动 *jidong* ‘excited’
19	4.03053	哪样 *nayang* ‘What’
18	3.26287	难受 *nanshou* ‘’unbearable
17	3.69498	人才 *rencai* ‘talent’
16	4.21324	不易 *buyi* ‘not easy’
15	3.18426	无奈 *wunai* ‘helpless’
14	3.1866	郁闷 *yumen* ‘depressed’
13	3.11531	麻烦 *mafan* ‘troublesome’
12	4.07384	苍天 *cangtian* ‘the grand sky’
11	4.61127	来之不易 *laizhibuyi* ‘hard-won’

(continued)

Freq	Stat (MI)	Collocate
11	3.61127	兴奋 *xingfen* ‘excited’
11	3.15184	证 *zheng* ‘certificate’
10	4.47377	没完 *meiwan* ‘endless’
10	4.25137	诱人 *youren* ‘attractive’
10	4.15184	朱丽叶 *zhuliye* ‘Juliet’
10	3.41487	顺 *shun* ‘smooth’
10	3.20075	形象 *xingxiang* ‘image’
10	3.05873	受不了 *shoubuliao* ‘unbearable’
9	4.7368	洋务派 *yangwupai* ‘proponents of The Westernization Movement’
9	4.09937	鸡翅 *jichi* ‘chicken wing’
9	3.7368	差距 *chaju* ‘gap’
9	3.26287	漏 *lou* ‘leak’
9	3.26287	奇葩 *qipa* ‘strange flower (to be weird)’
8	4.03636	节操 *jiecao* ‘integrity’
8	3.82991	邮 *you* ‘mail’
8	3.7368	好笑 *haoxiao* ‘funny’
8	3.48887	料 *liao* ‘ingredients’
8	3.48887	吓人 *xiaren* ‘scary’
8	3.34448	福利 *fuli* ‘wellbing’
8	3.09294	丑 *chou* ‘ugly’
8	3.03636	热闹 *renao* ‘hussle and bustle’
7	4.22223	心声 *xinsheng* ‘the voice from the bottom of the heart’
7	3.84372	天理 *tianli* ‘fairness’
7	3.7368	畜生 *chusheng* ‘animal (in a derogatory sense)’
7	3.54416	悲哀 *beiai* ‘sad’
7	3.54416	厚道 *houdao* ‘honest’

(continued)

Freq	Stat (MI)	Collocate
7	3.45669	王道 *wangdao* ‘The Kingly Way’
7	3.45669	打折 *dazhe* ‘to have discount’
7	3.37423	妻 *qi* ‘wife’
7	3.29623	奢侈 *shechi* ‘luxurious’
7	3.15184	威武 *weiwu* ‘grand’
7	3.08472	蛇 *she* ‘snake’
7	3.02059	烦躁 *fanzao* ‘irritated’
6	4.7368	funny ‘the English word “funny”’
6	4.7368	dian ‘an unknown sequence of pinyin in Mandarin’
6	4.51441	辈 *bei* ‘generation’
6	4.51441	搬家 *banjia* ‘to move’
6	4.51441	四声 sisheng ‘four tones’
6	4.51441	佃 *dian* ‘tenant’
6	4.32176	没错 *meicuo* ‘No mistakes’
6	4.15184	好强 *haoqiang* ‘so strong’
6	4.15184	停车位 *tingchewei* ‘parking lot’
6	4.15184	sorry ‘the English word “sorry”’
6	3.86233	泪流满面 *leiliu manmian* ‘Tears are all over the face.’
6	3.86233	喜庆 *xiqing* ‘festive’
6	3.7368	无误 *wuwu* ‘no mistakes’
6	3.62132	惬意 *qieyi* ‘comfortable’
6	3.51441	孽 *nie* ‘evil’
6	3.41487	自杀 *zisha* ‘suicide’
6	3.2343	尴尬 *ganga* ‘embarrassment’
6	3.15184	良心 *liangxin* ‘conscience’
6	3.15184	心疼 *xinteng* ‘frustrated’
6	3.07384	美味 *meiwei* ‘delicious food’

(continued)

Freq	Stat (MI)	Collocate
5	4.7368	蒙娜丽莎 *mengnalisha* 'Monalisa'
5	4.7368	淫荡 *yindang* 'lewd'
5	4.7368	后妻 *houqi* 'step wife'
5	4.7368	交汇点 *jiaohuidian* 'intersection'
5	4.7368	dokidoki 'an unknown sequence of English letters'
5	4.47377	空穴来风 *kongxuelaifeng* 'not without reasons'
5	4.47377	妹夫 *meifu* 'younger sister's husband (an insulting word)'
5	4.05873	难看 *nankan* 'ugly'
5	4.05873	惭愧 *cankui* 'guilty'
5	3.8888	民间 *minjian* 'folks'
5	3.8888	萌 *meng* 'cute'
5	3.7368	真理 *zhenli* 'truth'
5	3.5993	效率 *xiaolyu* 'efficiency'
5	3.47377	班车 *banche* 'shuttle'
5	3.47377	流泪 *liulei* 'to shed tears'
5	3.47377	回归 *huigui* 'to return'
5	3.35829	猥琐 *weisuo* 'to be obscene'
5	3.35829	浅 *qian* 'shallow'
5	3.35829	对比 *duibi* 'comparison'
5	3.25137	爆发 *baofa* 'outbreak'
5	3.25137	废话 *feihua* 'nonsense'
5	3.25137	到位 *daowei* 'adequate and appropriate'
5	3.25137	亲切 *qinqie* 'friendly and amiable'
5	3.15184	阴 *yin* 'dark and wet'
5	3.15184	精辟 *jingpi* 'insightful'
5	3.15184	好处 *haochu* 'advantage'

(continued)

Freq	Stat (MI)	Collocate
5	3.15184	rp
5	3.05873	得意 *deyi* ‘happy’

Table 39: Collocate display of 2L of NI-a

Freq	Stat (MI)	Collocate
71	3.46028	坑 *keng* ‘a hole on the ground’
23	3.80093	何以 *heyi* ‘to rely on what’
9	3.6588	兵器 *bingqi* ‘weaponry’
7	4.54416	边城 *biancheng* ‘cities near borders’
7	3.29623	贤 *xian* ‘sagacious’
7	3.84372	清晰 *qingxi* ‘clear’
7	4.22223	有钱 *youqian* ‘wealthy’
6	4.32176	小于 *xiaoyu* ‘less than’
6	3.2343	冻死 *dongsi* ‘frozen to death’
5	3.05873	略 *lue* ‘a little bit’
5	4.7368	无理取闹 *wuliqunao* ‘disruptive’
5	4.47377	bobby ‘an English name’

Table 40: Collocate display of 3L of NI-a

Freq	Stat (MI)	Collocate
149	3.97298	伤 *shang* 'hurt'
14	3.63727	人人 *renren* 'everyone'
7	3.37423	江 *jiang* 'river'
6	4.51441	辈 *bei* 'generation'
6	3.32176	要不 *yaobu* 'If … then …'
6	4.15184	洋务派 *yangwupai* 'proponents of the Westernization Movement'
6	3.99984	处处 *chuchu* 'every place'
6	3.7368	俺家 *anjia* 'my home'
6	4.51441	佃 *dian* 'tenant'
6	4.7368	dian 'a sequence of pinyin in Mandarin'
5	3.5993	物质 *wuzhi* 'material'
5	3.05873	复习 *fuxi* 'review; to go over'

Collocate displays of I-a

Table 41: Collocate display of 1L of I-a

Freq	Stat (MI)	Collocate
197	3.16331	事 *shi* 'matter; incidence'
115	3.3018	哪里 *nali* 'where'
95	3.37909	样 *yang* 'appearance; situation'
94	3.40772	办 *ban* 'do'
66	3.83751	意思 *yisi* 'meaning'
46	3.31668	未 *wei* 'not yet'
44	3.25255	哪儿 *naer* 'where'
39	4.08925	啊啊 *a a* 'a reduplication of SFP a'

(continued)

Freq	Stat (MI)	Collocate
29	3.15806	结婚 *jiehun* ‘to get married’
29	3.22012	情况 *qingkuang* ‘situation’
29	3.09856	怎样 *zenyang* ‘how’
29	3.4991	不行 *buxing* ‘not OK’
28	4.18544	堪 *kan* ‘to be able to bear’
23	3.48661	节目 *jiemu* ‘program’
19	3.14058	关系 *guanxi* ‘relationship’
17	3.37244	多久 *duojiu* ‘how long’
17	3.25696	事儿 *shier* ‘thing’
13	4.04681	什么人 *shenmeren* ‘what people’
12	4.20697	了不起 *liaobuqi* ‘great’
10	3.6683	毛病 *maobing* ‘problem(s)’
10	4.73243	歌颂 *gesong* ‘to sing highly of’
10	4.38451	哪样 *nayang* ‘which’
8	3.86994	自杀 *zisha* ‘suicide’
8	3.86994	神经病 *shenjingbing* ‘neuropathy (an insulting word)’
8	3.01195	料 *liao* ‘ingredients’
8	3.01195	播 *bo* ‘to broadcast’
8	3.01195	宁 *ning* ‘peace’
8	3.96304	区别 *qubie* ‘difference’
7	3.7704	软件 *ruanjian* ‘software’
7	4.21786	没完 *meiwan* ‘endless’
6	3.4549	錢 *qian* ‘money’
6	3.99547	报警 *baojing* ‘to tell the police’
6	3.20697	孽 *nie* ‘evil’
6	3.06258	升级 *shengji* ‘upgrade’
6	3.13297	刺激 *ciji* ‘incentive’

(continued)

Freq	Stat (MI)	Collocate
6	3.13297	何在 *hezai* ‘where’
6	3.75446	仇 *chou* ‘hatred’
6	3.64754	东东 *dongdong* ‘a neologism to refer to things’
5	3.86994	难看 *nankan* ‘ugly’
5	3.6069	烟花 *yanhua* ‘firework’
5	3.6069	心疼 *xinteng* ‘frustrated’
5	3.86994	心事 *xinshi* ‘a thing to worry about’
5	3.02194	偶遇 *ouyu* ‘chance encounter’
5	4.86994	什么歌 *shenmege* ‘what song’
5	4.86994	什么书 *shenmeshu* ‘what book’
5	3.73243	丢人 *diuren* ‘embarrassment’
5	3.38451	上映 *shangying* ‘to be on the air’

Table 42: Collocate display of 2L of I-a

Freq	Stat (MI)	Collocate
39	3.14411	假 *jia* ‘false’
25	4.34387	何以 *heyi* ‘to rely on what’
19	3.11786	坑 *keng* ‘a hole on the ground’
16	3.31535	几时 *jishi* ‘when’
9	3.03986	玩意 *wanyi* ‘gadget’
7	3.28497	折磨 *zhemo* ‘torture’
6	3.36743	舍不得 *shebude* ‘reluctant’

There is no significant collocate on the 3L position of I-a.

APPENDIX B

These are the 34 documents that constitute my mini-corpus of studies on pragmatic particles. Note that the full reference list for this dissertation is in the end.

Aijmer, K. (2002). Introduction. In *English Discourse Particles: Evidence from a Corpus* (pp. 1–56). John Benjamins.

Alleton, V. (1981). FINAL PARTICLES AND EXPRESSION OF MODALITY IN MODERN CHINESE. *Journal of Chinese Linguistics*, *9*(1), 91–115. https://www.jstor.org/stable/23753519

Chappell, H. (1991). Strategies for the assertion of obviousness and disagreement in Mandarin: A semantic study of the modal particle me. *Australian Journal of Linguistics*, *11*, 39–65.

Chu, C. C. (2006). A Contrastive Approach to Discourse Particles ——A Case Study of the Mandarin UFP Ne 呢. *Journal of Foreign Languages [外国语]*, *163*, 7–29.

Constant, N. (2011). On the independence of Mandarin aspectual and contrastive sentence-final ne. *The 23rd North American Conference on Chinese Linguistics (NACCL-23)*.

Deng, D. (2015). The syntacticization of illocutionary forces and the root vs. non-root distinction: Evidence from the sentence-final particle ba in Mandarin. *Lingua*, *162*, 32–55.

Erbaugh, M. S. (1985). Sentence Final Particles as an Asian Areal Feature. In S. DeLancey & R. S. Tomlin (Eds.), *the First Annual Meeting of the Pacific Linguistics Conference*.

Fang, H. (2018). Mirativity in Mandarin: The Sentence-Final Particle Le (了). *Open Linguistics*, *4*, 589–607. https://doi.org/https://doi.org/10.1515/opli-2018-0029

Goddard, C. (2005). Sentence-final (illocutionary) particles. In *The Languages of East and Southeast Asia : An Introduction* (pp. 144–147). Oxford University Press.

Han, Y. S. (1995). A PRAGMATIC ANALYSIS OF THE BA PARTICLE IN MANDARIN CHINESE. *Journal of Chinese Linguistics*, *23*(2), 99–128. https://www.jstor.org/stable/23756541

Hancil, S., Post, M., & Haselow, A. (2015). Introduction: Final particles from a typological perspective. In S. Hancil, A. Haselow, & M. Post (Eds.), *Final Particles* (pp. 3–35). De Gruyter, Inc. https://ebookcentral.proquest.com/lib/uoregon/detail.action?docID=1867198

Haselow, A. (2011). Discourse marker and modal particle: The functions of utterance-final then in spoken English. *Journal of Pragmatics*, *43*, 3603–3623. https://doi.org/10.1016/j.pragma.2011.09.002

Haselow, A. (2012). Subjectivity, intersubjectivity and the negotiation of common ground in spoken discourse: Final particles in English. *Language & Communication*, *32*, 182–204. https://doi.org/http://dx.doi.org/10.1016/j.langcom.2012.04.008

Heine, B. (2013). On discourse markers: Grammaticalization, pragmaticalization, or something else? *Linguistics*, *51*(6), 1205 – 1247. https://doi.org/10.1515/ling-

2013-0048

Hudson, M. E., & Lu, W. (2003). On Japanese ne and Chinese ba. In K. M. Jaszczolt & K. Turner (Eds.), *Meaning Through Language Contrast. Volume 2* (pp. 197–212). John Benjamins Publishing Company. https://doi.org/https://doi.org/10.1075/pbns.100.13hud

Jing-Schmidt, Z. (2021). Sentence-final particles: Sociolinguistic and discourse perspectives. In C.-R. Huang, Y.-H. Lin, & I.-H. Chen (Eds.), *Cambridge Handbook of Chinese Linguistics*. Cambridge University Press. (Forthcoming)

LAM, P. W. Y. (2009). Discourse Particles in Corpus Data and Textbooks: The Case of Well. *Applied Linguistics*, *31*(2), 260–281. https://doi.org/10.1093/applin/amp026

Lee-Wong, S. M. (1998). Face Support-Chinese Particles as Mitigators: A Study of BA A/YA and NE. *Pragmatics: Quarterly Publication of the International Pragmatics Association*, *8*(3), 387–404.

Li, B. (2013). INTEGRATING TEXTUAL AND PROSODIC FEATURES IN THE INTERPRETATION OF CHINESE UTTERANCE-FINAL-PARTICLES: A CASE OF A AND NE. *Journal of Chinese Linguistics*, *41*(1), 145–169. https://www.jstor.org/stable/23753857

Lin, W. (1984). WHAT DOES THE MANDARIN PARTICLE NE COMMUNICATE? *Cahiers de Linguistique-Asie Orientale*, *13*(2), 217–240. https://doi.org/https://doi.org/10.3406/clao.1984.1157

Ljungqvist, M. (2010). Zhongwen Yuyong Biaojiyu Ba de Guanlian Lilun Fenxi [A RELEVANCE-THEORETIC ANALYSIS OF THE PRAGMATIC MARKER BA IN MANDARIN CHINESE]. *Journal of Chinese Linguistics*, *38*(2), 261–287.

Luke, K. K. (1990). Introduction: Utterance Particles in Cantonese. In *Utterance Particles in Cantonese Conversation* (pp. 1–16). John Benjamins Publishing Company.

Morita, E. (2015). Japanese interactional particles as a resource for stance building. *Journal of Pragmatics*, *83*, 91--103. https://doi.org/http://dx.doi.org/10.1016/j.pragma.2014.12.008

Norrick, N. (2009). Interjections as pragmatic markers. *Journal of Pragmatics*, *41*, 866–891. https://doi.org/10.1016/j.pragma.2008.08.005

Sato, S. (2017). On establishing I think as a final particle in interactions: Some comparisons with sentence-final particles in Japanese. *Journal of Pragmatics*, *110*, 83--98. https://doi.org/http://dx.doi.org/10.1016/j.pragma.2016.11.013

Strauss, S., & Xiang, X. (2009). Discourse particles: Where cognition and interaction intersect—The case of final particle ey in Shishan dialect. *Journal of Pragmatics*, *41*, 1287–1312. https://doi.org/10.1016/j.pragma.2009.02.006

Sybesma, R., & Li, B. (2007). The dissection and structural mapping of Cantonese sentence final particles. *Lingua*, *117*, 1739–1783. https://doi.org/10.1016/j.lingua.2006.10.003

Tang, S.-W. (2015). A Generalized Syntactic Schema for Utterance Particles in Chinese. *Lingua Sinica*, *1*(3), 1–23. https://doi.org/10.1186/s40655-015-0005-5

Tantucci, V. (2017). An evolutionary approach to semasiological change: Overt influence attempts through the development of the Mandarin 吧-ba particle.

Journal of Pragmatics, *120*, 35–53. https://doi.org/http://dx.doi.org/10.1016/j.pragma.2017.08.006

Wu, R.-J. (2004). Introduction. In *Stance in Talk : A Conversation Analysis of Mandarin Final Particles* (pp. 1–24). John Benjamins Publishing Company.

Wu, R.-J. R. (2005). '"There is more here than meets the eye!"': the use of final ou in two sequential positions in Mandarin Chinese conversation. *Journal of Pragmatics*, *37*, 967–995. https://doi.org/10.1016/j.pragma.2004.12.006

Xiang, X. (2011). Constraint reality: Linguistic expressions of restrictivity and emotive stances. A discourse-pragmatic study of utterance-final lāh in Shishan (Hainan Island, China). *Lingua*, *121*(8), 1377–1400. https://doi.org/10.1016/j.lingua.2011.03.002

Yang, X., & Wiltschko, M. (2016). The confirmational marker ha in Northern Mandarin. *Journal of Pragmatics*, *104*, 67–82. https://doi.org/http://dx.doi.org/10.1016/j.pragma.2016.09.004

Yap, F. H., Wang, J., & Lam, C. T.-K. (2010). Clausal integration and the emergence of mitigative and adhortative sentence-final particles in Chinese. *Taiwan Journal of Linguistics*, *8*(2), 63–86.

APPENDIX C

Table 43: 50 terms extracted from my own mini-corpus of pragmatic particle studies

Ranking	Term
1	modal particle
2	use subject
3	final particle
4	illocutionary force
5	final ou
6	final position
7	speech act
8	propositional content
9	discourse particle
10	english discourse
11	conversation analysis
12	discourse marker
13	sentence-final particle
14	utterance particle
15	main clause
16	pragmatic pretext
17	pragmatic marker
18	ne
19	topic marker

(continued)

Ranking	Term
20	preceding discourse
21	core meaning
22	yes-no question
23	preceding utterance
24	internal conjunct
25	use of final ou
26	mandarin particle
27	relevance theory
28	clausal integration
29	particle ba
30	boundary tone
31	contrastive topic
32	completive aspect
33	initial position
34	discourse segment
35	default vowel
36	discourse function
37	responsive position
38	modality page
39	default tone

40	pragmatic analysis
41	question particle
42	rising intonation
43	typological perspective
44	reference grammar
45	expression of modality page
46	ba-tagged sentence
47	preceding proposition
48	word order
49	discourse analysis
50	syntactic structure

REFERENCES CITED

Aijmer, K. (2002a). Introduction. In *English Discourse Particles: Evidence from a Corpus* (pp. 1–56). John Benjamins.

Aijmer, K. (2002b). The interpersonal particle just. In *English Discourse Particles: Evidence from a Corpus* (pp. 153–174). John Benjamins.

Alleton, V. (1981). FINAL PARTICLES AND EXPRESSION OF MODALITY IN MODERN CHINESE. *Journal of Chinese Linguistics*, *9*(1), 91–115. https://www.jstor.org/stable/23753519

Andersen, G. (2017). A corpus study of pragmatic adaptation: The case of the Anglicism [jobb] in Norwegian. *Journal of Pragmatics*, *113*, 127--143. https://doi.org/http://dx.doi.org/10.1016/j.pragma.2016.12.015

Anthony, L. (2020). *AntConc* (3.5.9). Waseda University. https://www.laurenceanthony.net/software

Ashby, M., & Maidment, J. (2005). *Introducing Phonetic Science*. Cambridge University Press.

Berkenfield, C. (2001). The role of frequency in the realization of English that. In J. Bybee & P. Hopper (Eds.), *Frequency and the Emergence of Linguistic Structure* (pp. 281–307). John Benjamins Publishing Company.

Bublitz, W. (1995). Semantic prosody and cohesive company: somewhat predictable. *Leuvense Bijdragen*, *85*(1), 1–32.

Bybee, J. (2013). Usage-based Theory and Exemplar Representations of Constructions. In T. Hoffmann & G. Trousdale (Eds.), *The Oxford Handbook of Construction Grammar* (pp. 49–69). Oxford University Press.

Callier, P. (2007). *A Corpus Study of Sentence-Final me: "Cosmopolitan" Mandarin?* Stanford University.

CASS. (2019). *Xiandai Hanyu Cidian [Dictionary of Modern Chinese]* (7th ed.). The Commercial Press.

Chan, M. (2002). Gender-related Use of Sentence-final Particles in Cantonese. In M. Hellinger & H. Bußmann (Eds.), *Gender Across Languages: The Linguistic Representation of Women and Men* (pp. 57–72).

Chan, M. (1998). Sentence particles je and jek in Cantonese and their distribution across gender and sentence types. *Engendering Communication: Proceedings of the Fifth Berkeley Women and Language Conference*, 117–128.

Chan, M. K. M. (1996). GENDER-MARKED SPEECH IN CANTONESE: THE CASE OF SENTENCE-FINAL PARTICLES JE AND JEK. *Studies in Linguistic Sciences*, *26*, 1–38.

Chao, Y. R. (1926). Beijing, Suzhou, Changzhou Yuzhuci De Yanjiu [Studies of SFPs in the Regional Varieties of Beijing, Suzhou, and Changzhou]. *Qinghua Daxue Xuebao (Ziran Kexue Ban) [Journal of Tsinghua University (Science and Technology)]*, *2*.

Chao, Y. R. (1968). *A grammar of spoken Chinese.* University of California Press.

Chappell, H. (1991). Strategies for the assertion of obviousness and disagreement in Mandarin: A semantic study of the modal particle me. *Australian Journal of Linguistics*, *11*, 39–65.

Chu, C. C. (2006). A Contrastive Approach to Discourse Particles ——A Case Study of the Mandarin UFP Ne 呢. *Journal of Foreign Languages [外国语]*, *163*, 7–29.

Croft, W. (2001). *Radical Construction Grammar : Syntactic Theory in Typological Perspective*. Oxford UP.

Deng, D. (2015). The syntacticization of illocutionary forces and the root vs. non-root distinction: Evidence from the sentence-final particle ba in Mandarin. *Lingua*, *162*, 32–55.

Desagulier, G. (2017a). Association and Productivity. In *Corpus Linguistics and Statistics with R: Introduction to Quantitative Methods in Linguistics* (pp. 197–238). Springer International Publishing. https://doi.org/10.1007/978-3-319-64572-8

Desagulier, G. (2017b). *Corpus Linguistics and Statistics with R : Introduction to Quantitative Methods in Linguistics*. Springer.

Dong, H. (2014). *A History of the Chinese Language*. Routledge.

Enfield, N. J. (2005). Areal Linguistics and Mainland Southeast Asia. *Annual Review of Anthropology*, *34*, 181–206. https://doi.org/10.1146/annurev.anthro.34.081804.120406

Erbaugh, M. (2013). Classifier choices in discourse across the seven main Chinese dialects. In Z. Jing-Schmidt (Ed.), *Increased Empiricism: Recent Advances in Chinese Linguistics* (pp. 101–126). John Benjamins.

Erbaugh, M. S. (1985). Sentence Final Particles as an Asian Areal Feature. In S. DeLancey & R. S. Tomlin (Eds.), *the First Annual Meeting of the Pacific Linguistics Conference*.

Fang, H. (2018). Mirativity in Mandarin: The Sentence-Final Particle Le (了). *Open Linguistics*, *4*, 589–607. https://doi.org/https://doi.org/10.1515/opli-2018-0029

Feng, S. (2015). Shengdiao, yudiao, yu hanyu de jumo yuqici [Tone, Intonation and Sentence Final Particles in Chinese]. *Essays on Linguistics*, *1*, 52–79.

Fillmore, C. (1985). Syntactic Intrusions and The Notion of Grammatical Construction. *Proceedings of the Eleventh Annual Meeting of the Berkeley Linguistics Society*, 73–86.

Fillmore, C., Kay, P., & O'Connor, M. C. (1988). Regularity and Idiomaticity in Grammatical Constructions: The Case of Let Alone. *Language*, *64*(3), 501–538. https://www.jstor.org/stable/414531?seq=1#metadata_info_tab_contents

Firth, J. R. (1957). A synopsis of linguistic theory 1930–1955. In F. Palmer (Ed.), *Selected Papers of J. R. Firth*. Longman.

Fraser, B. (1999). What are discourse markers? *Journal of Pragmatics*, *31*, 931–952.

Fraser, B. (2015). The combining of Discourse Markers -- A beginning. *Journal of Pragmatics*, *86*, 48--53. https://doi.org/http://dx.doi.org/10.1016/j.pragma.2015.06.007

Givón, T. (1995). Isomorphism in the Grammatical Code: Cognitive and Biological Considerations. In R. Simone (Ed.), *ICONICITY IN LANGUAGE* (pp. 47–76). John Benjamins Publishing Company.

Givón, T. (2001). Non-declarative speech-acts. In *Syntax: An Introduction* (pp. 287–325). John Benjamins Publishing Company.

Goddard, C. (2005). Sentence-final (illocutionary) particles. In *The Languages of East and Southeast Asia : An Introduction* (pp. 144–147). Oxford University Press.

Goldberg, A. (1995). *Constructions: A Construction Grammar Approach to Argument Structure*. University of Chicago Press.

Goldberg, A. (2003). Constructions: a new theoretical approach to language. *Trends in Cognitive Sciences*, *7*(5), 219–224.

Goldberg, A. (2006). *Constructions at work: The nature of generalization in language (Oxford linguistics)*. Oxford University Press.

Goldberg, A. (2019). *Explain Me This: Creativity, Competition, and the Partial Productivity of Constructions*. Princeton University Press.

Gries, S. (2009). *Quantitative Corpus Linguistics with R: A Practical Introduction*. Routledge.

Gries, S., & Newman, J. (2013). Creating and using corpora. In R. Podesva & D. Sharma (Eds.), *Research Method in Linguistics* (pp. 257–287). Cambridge University Press.

Guo, S., & Li, S. (2016). Supporting Collocation Learning and Teaching with a Chinese Collocation Profile Database. *Journal of Technology and Chinese Language Teaching*, *7*(2), 38–57.

Han, Y. (1988). *A pragmatic study of some sentence-final and post-verbal particles in Mandarin Chinese*. University of York.

Hancil, S., Post, M., & Haselow, A. (2015). Introduction: Final particles from a typological perspective. In S. Hancil, A. Haselow, & M. Post (Eds.), *Final Particles* (pp. 3–35). De Gruyter, Inc. https://ebookcentral.proquest.com/lib/uoregon/detail.action?docID=1867198

Haselow, A. (2011). Discourse marker and modal particle: The functions of utterance-final then in spoken English. *Journal of Pragmatics*, *43*, 3603–3623. https://doi.org/10.1016/j.pragma.2011.09.002

Haselow, A. (2012). Subjectivity, intersubjectivity and the negotiation of common ground in spoken discourse: Final particles in English. *Language & Communication*, *32*, 182–204. https://doi.org/http://dx.doi.org/10.1016/j.langcom.2012.04.008

Hayano, K. (2013). Question Design in Conversation. In J. Sidnell & T. Stivers (Eds.), *Handbook of Conversation Analysis* (pp. 395–414). Wiley-Blackwell.

Heine, B. (2013). On discourse markers: Grammaticalization, pragmaticalization, or something else? *Linguistics*, *51*(6), 1205 – 1247. https://doi.org/10.1515/ling-2013-0048

Heritage, J. (2013). Epistemics in Conversation. In J. Sidnell & T. Stivers (Eds.), *Handbook of Conversation Analysis* (pp. 370–394). Wiley-Blackwell.

Hilpert, M. (2014a). *Construction Grammar and Its Application to English*. Edinburgh UP.

Hilpert, M. (2014b). Introducing Construction Grammar. In *Construction Grammar and Its Application to English* (pp. 1–24). Edinburgh UP.

Hoey, M. (2005). *Lexical Priming: A New Theory of Words and Language*. Routledge.

Hoffman, S., Fischer-Starcke, B., & Sand, A. (2015). Introduction. In S. Hoffman, B. Fischer-Starcke, & A. Sand (Eds.), *Current Issues in Phraseology* (pp. 1–5). John Benjamins Publishing Company.

Hu, Y. (1995). *Xiandai Hanyu [Modern Chinese]* (B. Xu, S. Zhang, H. Zhang, B. Zhang, Q. Yang, X. Yan, K. Fan, Z. Zhou, & H. Yuan (eds.)). Shanghai Educational Publishing House.

Huang, Y. (2014). *Pragmatics* (2nd ed.). Oxford University Press.

Ji, S. (2007). A textual perspective on Givón's quantity principle. *Journal of Pragmatics*, *39*(2), 292–304. https://doi.org/10.1016/j.pragma.2006.01.010

Jing-Schmidt, Z. (2015). The Place of Linguistics in CSL Teaching and Teacher Education: Toward a Usage-Based Constructionist Theoretical Orientation. *Journal of the Chinese Language Teachers Association*, *50*(3), 1–22.

Jing-Schmidt, Z. (2017). What are they good for? A constructionist account of counterfactuals in ordinary Chinese. *Journal of Pragmatics*, *113*, 30–52. https://doi.org/http://dx.doi.org/10.1016/j.pragma.2017.03.004

Jing-Schmidt, Z. (2019). Sentence-final particles: Sociolinguistic and discourse perspectives. In C.-R. Huang, Y.-H. Lin, & I.-H. Chen (Eds.), *Cambridge Handbook of Chinese Linguistics*. Cambridge University Press.

Jing-Schmidt, Z., Jun, L., Steffi, H., & Lin, Z. (2021). Aspect construal in Mandarin: A usage-based constructionist perspective on LE (under review). *Linguistics*.

Johnson, D. E. (2013). Descriptive statistics. In R. Podesva & D. Sharma (Eds.), *Research Method in Linguistics* (pp. 288–315). Cambridge University Press.

Jurafsky, D., & Martin, J. H. (2019). Vector Semantics and Embeddings. In *Speech and Language Processing (Draft of October 2, 2019)*.

Kilgarriff, A., & Rychlý, P. (2003). *Sketch Engine*. Lexical Computing Limited. https://www.sketchengine.eu/

Kilgrarriff, A. (2005). Language is never, ever, ever, random. *Corpus Linguistics And Linguistic Theory*, *1*(2), 263–276.

King, B. (1986). Ne--A Discourse Approach. *Journal of the Chinese Language Teachers Association (JCLTA)*, *XXI*(1), 21–46.

Kishner, J. M., & Gibbs, R. W. (1996). How "'just'" Gets Its Meanings: Polysemy and Context in Psychological Semantics. *Language And Speech*, *39*, 19–36.

LAM, P. W. Y. (2009). Discourse Particles in Corpus Data and Textbooks: The Case of Well. *Applied Linguistics*, *31*(2), 260–281. https://doi.org/10.1093/applin/amp026

Lee-Wong, S. M. (1998). Face Support-Chinese Particles as Mitigators: A Study of BA A/YA and NE. *Pragmatics: Quarterly Publication of the International Pragmatics Association*, *8*(3), 387–404.

Lee, S. (2018). Seeking Academic Help: A Case Study of Peer Brokering Interactions. *Transitions: Journal of Transient Migration*, *2*(2), 149–173. https://doi.org/10.1386/tjtm.2.2.149_1

Levinson, S. C. (2013). Action Formation and Ascription. In J. Sidnell & T. Stivers (Eds.), *The Handbook of Conversation Analysis* (pp. 103–130). Wiley-Blackwell.

Li, Bin. (2013). INTEGRATING TEXTUAL AND PROSODIC FEATURES IN THE INTERPRETATION OF CHINESE UTTERANCE-FINAL-PARTICLES: A CASE OF A AND NE. *Journal of Chinese Linguistics*, *41*(1), 145–169. https://www.jstor.org/stable/23753857

Li, Boya. (2006). *Chinese Final Particles and the Syntax of the Periphery*. Netherlands Graduate School of Linguistics.

Li, C., & Thompson, S. (1981). *Mandarin Chinese: A Functional Reference Grammar*. University of California Press.

Li, S., & Guo, S. (2016). Collocation Analysis Tools for Chinese Collocation Studies/ 可用于汉语搭配研究的搭配分析工具. *Journal of Technology and Chinese Language Teaching*, *7*(1), 56–77.

Lin, W. (1984). WHAT DOES THE MANDARIN PARTICLE NE COMMUNICATE? *Cahiers de Linguistique-Asie Orientale*, *13*(2), 217–240. https://doi.org/https://doi.org/10.3406/clao.1984.1157

Lin, Y.-H. (2007). *The Sounds of Chinese*. Cambridge University Press.

Liu, Y., Pan, W., & Gu, L. (2001). *Shiyong Xiandai Hanyu Yufa (Zengding Ben) [A Practical Grammar of Modern Chinese (extended edition)]*. Commercial Press.

Ljungqvist, M. (2010). Zhongwen Yuyong Biaojiyu Ba de Guanlian Lilun Fenxi [A RELEVANCE-THEORETIC ANALYSIS OF THE PRAGMATIC MARKER BA IN MANDARIN CHINESE]. *Journal of Chinese Linguistics*, *38*(2), 261–287.

Louw, B. (1993). Irony in the Text or Insincerity in the Writer? The Diagnostic Potential of Semantic Prosodies. In M. Baker, G. Francis, & E. Tognini-Bonelli (Eds.), *Text and Technology: In Honour of John Sinclair* (pp. 157–176). John Benjamins.

Lu, W.-Y. (2005). *SENTENCE-FINAL PARTICLES AS ATTITUDE MARKERS IN MANDARIN CHINESE*. University of Illinois at Urbana-Champaign.

Luke, K. K. (1990). *Utterance particles in Cantonese conversation (Pragmatics & beyond ; new ser. 9)*. J. Benjamins Pub. Co.

Lyu, S. (1974). *Zhongguo Wenfa Yaolue [Outline of Chinese Grammar]*. Wenshizhe Publishing House.

Lyu, S. (1999). *Xiandai Hanyu Babai Ci [Eight Hundred Words of Chinese]*. Commercial Press.

Matthews, P. H. (2014). *The Concise Oxford Dictionary of Linguistics* (3rd ed.). Oxford University Press. https://www.oxfordreference.com/view/10.1093/acref/9780199675128.001.0001/acref-9780199675128

McEnery, T., & Hardie, A. (2012). *Corpus linguistics : Method, theory and practice*. Cambridge University Press.

McEnery, Tony, & Hardie, A. (2012). Neo-Firthian corpus linguistics. In Tony McEnery & A. Hardie (Eds.), *Corpus Linguistics: Method, Theory and Practice*. Cambridge University Press.

Morita, E. (2015). Japanese interactional particles as a resource for stance building. *Journal of Pragmatics*, *83*, 91--103. https://doi.org/http://dx.doi.org/10.1016/j.pragma.2014.12.008

Norrick, N. (2009). Interjections as pragmatic markers. *Journal of Pragmatics*, *41*, 866–891. https://doi.org/10.1016/j.pragma.2008.08.005

Okamoto, S. (1996). Pragmaticization of meaning in some sentence-final particles in Japanese. In M. Shibatani & S. Thompson (Eds.), *Essays in Semantics and Pragmatics : In honor of Charles J. Fillmore* (pp. 219–246). John Benjamins Publishing Company.

Östman, J.-O. (1981). *You know: a discourse functional approach*. John Benjamins Publishing Company.

Östman, J.-O. (1982). The symbiotic relationship between pragmatic particles andimpromptu speech. *Impromptu Speech: A Symposium*, 147–177.

Partington, A. (2017). Evaluative clash, evaluative cohesion and how we actually read evaluation in texts. *Journal of Pragmatics*, *117*, 190--203. http://dx.doi.org/10.1016/j.pragma.2017.06.008

Qi, H. (2002). *Yuqici Yu Yuqi Xitong [SPFs and Modal Systems]*. Anhui Jiaoyu Chubanshe.

Rhee, S. (2012). Context-induced reinterpretation and (inter)subjectification: the case of grammaticalization of sentence-final particles. *Language Sciences*, *34*, 284–300. https://doi.org/10.1016/j.langsci.2011.10.004

Sadock, J. (2004). Speech Acts. In L. Horn & G. Ward (Eds.), *The Handbook of Pragmatics* (pp. 53–73). John Wiley & Sons.

Sandel, T., Ou, C., Wangchuk, D., Ju, B., & Duque, M. (2019). Unpacking and describing interaction on Chinese WeChat: A methodological approach. *Journal of Pragmatics*, *143*, 228–241. https://doi.org/https://doi.org/10.1016/j.pragma.2018.08.011

Sato, S. (2017). On establishing I think as a final particle in interactions: Some comparisons with sentence-final particles in Japanese. *Journal of Pragmatics*, *110*, 83--98. https://doi.org/http://dx.doi.org/10.1016/j.pragma.2016.11.013

Schegloff, E. (2007). *Sequence organization in interaction: A primer in conversation analysis*. Cambridge University Press.

Schourup, L. (2011). The discourse marker now: A relevance-theoretic approach. *Journal of Pragmatics*, *43*, 2110–2129. https://doi.org/10.1016/j.pragma.2011.01.005

Searle, J. (1969). *Speech Acts: An Essay in the Philosophy of Language*. Cambridge University Press.

Shao, J. (2016). *Xiandai Hanyu Tonglun [General Theory of Modern Chinese]* (J. Shao, W. Wu, J. Fei, C. Zhao, & X. Fang (eds.); 3rd ed.). Shanghai Educational Publishing House.

Shi, D., & Huang, C.-R. (2016). Syntactic overview. In C.-R. Huang & D. Shi (Eds.), *A Reference Grammar of Chinese* (pp. 14–66). Cambridge University Press.

Sidnell, J., & Stivers, T. (2013). *The Handbook of Conversation Analysis*. Wiley-Blackwell.

Simpson, A. (2014). Sentence-Final Particles. In *The handbook of Chinese linguistics (First ed., Blackwell handbooks in linguistics)* (pp. 156–179). Wiley-Blackwell.

Sinclair, J. (2004a). The lexical item. In J. Sinclair & R. Carter (Eds.), *Trust the Text: Language, Corpus and Discourse* (pp. 132–148). Routledge.

Sinclair, J. (2004b). The search for units of meaning. In J. Sinclair & R. Carter (Eds.), *Trust the Text: Language, Corpus and Discourse* (pp. 34–58). Routledge.

Stempel, P. (2019). *A Constructional Reanalysis of Semantic Prosody*. Rice University.

Stewart, D. (2010). *Semantic prosody : A critical evaluation (Routledge advances in corpus linguistics ; 9)*. Routledge.

Stivers, T., & Sidnell, J. (2013). Introduction. In J. Sidnell & T. Stivers (Eds.), *The Handbook of Conversation Analysis* (pp. 1–8). Blackwell Publishing Ltd.

Stubbs, M. (2001). *Words and Phrases: Corpus Studies of Lexical Semantics*. Blackwell.

Sun, C. (2006). *Chinese: A Linguistic Introduction*. Cambridge University Press.

Sybesma, R., & Li, B. (2007). The dissection and structural mapping of Cantonese sentence final particles. *Lingua*, *117*, 1739–1783. https://doi.org/10.1016/j.lingua.2006.10.003

Tang, S.-W. (2015). A Generalized Syntactic Schema for Utterance Particles in Chinese. *Lingua Sinica*, *1*(3), 1–23. https://doi.org/10.1186/s40655-015-0005-5

Tantucci, V. (2017). An evolutionary approach to semasiological change: Overt influence attempts through the development of the Mandarin 吧-ba particle. *Journal of Pragmatics*, *120*, 35–53. https://doi.org/http://dx.doi.org/10.1016/j.pragma.2017.08.006

Ten Have, P. (1990). Methodological Issues in Conversation Analysis. *BMS: Bulletin of Sociological Methodology/Bulletin De Méthodologie Sociologique*, *27*, 23–51.

Traugott, E. C., & Trousdale, G. (2013). The Framework. In E. C. Traugott & G. Trousdale (Eds.), *Constructionalization and Constructional Changes* (pp. 1–44). Oxford University Press. https://doi.org/10.1093/acprof:oso/9780199679898.001.0001

Traugott, E., & Trousdale, G. (2013). *Constructionalization and Constructional Changes*. University of Oxford Press.

Wang, H., Shi, H. H., & Jing-Schmidt, Z. (2021). Affective stance in constructional idioms: A usage-based constructionist approach to Mandarin [yòu X yòu Y]. *Journal of Pragmatics*, *177*, 29–50. https://doi.org/https://doi.org/10.1016/j.pragma.2021.02.004

Wu, G. (2005). The Discourse Function of the Chinese Particle NE in Statements. *Journa of the Chinese Language Teachers Association (JCLTA)*, *40*(1), 47–82.

Wu, R.-J. (2004). *Stance in Talk : A Conversation Analysis of Mandarin Final Particles*. John Benjamins Publishing Company.

Wu, R.-J. R. (2005). '"There is more here than meets the eye!"': the use of final ou in two sequential positions in Mandarin Chinese conversation. *Journal of Pragmatics*, *37*, 967–995. https://doi.org/10.1016/j.pragma.2004.12.006

Xiang, X. (2011). Constraint reality: Linguistic expressions of restrictivity and emotive stances. A discourse-pragmatic study of utterance-final lāh in Shishan (Hainan Island, China). *Lingua*, *121*(8), 1377–1400. https://doi.org/10.1016/j.lingua.2011.03.002

Xu, J. (2018). Putonghua Kouyu Zhong "a, ya, na, wa" De Fenbu [The Distributions of "a， ya， na， wa"in Mandarin Chinese]. *Yuyan Wenzi Yingyong [Applied Linguistics]*, *2*, 62–71. https://doi.org/10.16499/j.cnki.1003-5397.2018.02.007

Xun, E., Rao, G., Xiao, X., & Zang, J. (2016). Dashuju Beijing Xia BCC Yuliaoku De Yanzhi [The construction of the BCC Corpus in the age of Big Data]. *Yuliaoku Yuyanxue [Corpus Linguistics]*, *3*(1), 93–118.

Yang, X., & Wiltschko, M. (2016). The confirmational marker ha in Northern Mandarin. *Journal of Pragmatics*, *104*, 67–82. https://doi.org/http://dx.doi.org/10.1016/j.pragma.2016.09.004

Yap, F. H., Wang, J., & Lam, C. T.-K. (2010). Clausal integration and the emergence of mitigative and adhortative sentence-final particles in Chinese. *Taiwan Journal of Linguistics*, *8*(2), 63–86.

Zhan, W., & Bai, X. (2016). Sentence types. In C.-R. Huang & D. Shi (Eds.), *A Reference Grammar of Chinese* (pp. 401–450). Cambridge University Press.

Zhang, H. (2012). *NLPIR Chinese Lexical Analysis System*. NLPIR-ICTCLAS. https://github.com/NLPIR-team/NLPIR

Zhao, C., & Shi, D. (2015). "bei" de taidu quxiang ji qi yuyi jichu [Attitudinal Orientations and Semantic Source of Sentence-Final Particle Bei]. *LANGUAGE TEACHING AND LINGUISTIC STUDIES*, *4*, 68–78.

Zhao, C., & Sun, L. (2015). Jumo Zhuci "ba" De Fenbu Yanzheng Yu Yuyi Tiqu [On the Distributional Verification and Semantic Extraction of the Sentence-final Particle Ba]. *Zhongguo Yuwen [Studies of the Chinese Language]*, *2*(365), 121–132; 191.

Zhu, D. (1984). *Yufa Jiangyi [Lectures on Grammar]*. Commercial Press.