Spring 2015

Sentence final particles in Shanghainese: Navigating the left periphery

Ethan C. Myers
Purdue University

Graduate School Form 30
(Revised 08/14)

PURDUE UNIVERSITY
GRADUATE SCHOOL
Thesis/Dissertation Acceptance

This is to certify that the thesis/dissertation prepared

By Ethan C Myers

Entitled
SHANGHAINESE SENTENCE FINAL PARTICLES: NAVIGATING THE LEFT PERIPHERY

For the degree of Master of Arts

Is approved by the final examining committee:

Elaine Francis

Ronnie Wilbur

Lori Czerwionka

Hong Wei

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

Approved by Major Professor(s): Elaine Francis

Approved by: Felicia Roberts 04/21/2015

Head of the Department Graduate Program Date

SENTENCE FINAL PARTICLES IN SHANGHAINESE:

NAVIGATING THE LEFT PERIPHERY

A Thesis

Submitted to the Faculty

of

Purdue University

by

Ethan C Myers

In Partial Fulfillment of the

Requirements for the Degree

of

Master of Arts

May 2015

Purdue University

West Lafayette, Indiana

This thesis goes out to my family, for having supported me all these years

While I pursued my dreams of traveling and studying.

For my brothers and sisters,

I hope that I can be a good example for you to follow as you emerge into the world;

Of optimism in the face of risk;

Perseverance during hardships;

And humility in my successes.

For my parents and grandparents, especially to my mom, Alice,

Whom I wish could have seen me finish.

I hope that I never lose that urge to make you proud

And I hope that I can find your strength

To recover my way when I get lost.

And finally this thesis is dedicated to my girlfriend, Jue,

Whose immeasurable patience, humor, and grace

Ensures that I never want for a better role model or companion.

Thanks!

ACKNOWLEDGMENTS

Of course, this thesis would have never come into being without the support of so many generous friends, family, and faculty at Purdue, where I've spent the best 6 years of my life (so far). To all of my colleagues in the Indigenous and Endangered Language Lab, and to Elena Benedicto, who have given me a purpose in more ways than one over four years; to right injustices, and to be a creative problem solver, to relax once in a while, and to always be in awe of language. To my major professor, Elaine Francis, and my other committee members Ronnie Wilbur, Hong Wei, and Lori Czerwionka, I can never repay you for all of the help you've given me during the year (or semester?) I spent writing this thesis. All of your words of advice when I was stuck, even if I didn't know at the time, to your words of consolation when I was stressed out. I will carry these lessons with me always. To my close friends from undergrad, who are breaking new ground across the country, as I pursue the academic path. Thanks to all of you for your constant companionship. Even though we've graduated and moved on, the communication is always open, and always pleasant. Thanks to my undergraduate advisor, Marianne Gupta, for letting me go to study in UNAM, and for introducing me to the field of linguistics. Thank you to the countless others for whom I lack the space and good sense to mention by name. I have had so much help over the years.

TABLE OF CONTENTS

Page

Page

ABSTRACT

Myers, Ethan C. M.A., Purdue University, May 2015. Shanghainese Sentence-Final Particles: Navigating the Left Periphery. Major Professor: Elaine Francis.

The purpose of the present study was to provide a modern syntactic analysis of Shanghainese, the regional language spoken in Shanghai, China. While Shanghainese has received significantly less exposure in the published literature compared to other regional counterparts, there is much that the language can contribute to the overall study of human language. In spite of the fact that Shanghainese has been experiencing a decline in first language learners in recent generations, this thesis aims to increase its exposure in the academic discourse and enforce the legitimacy of its standing as a distinct, living language and as a valuable piece of cultural identity that belongs to the people of Shanghai. This exposure will be specifically limited to the language's inventory and treatment of a class of lexical items known as sentence-final particles, which encode discourse material and information about the speaker's subjective states (e.g. surprise or assumption). These particles have been analyzed in detail in several languages by postulating the presence of rich functional fields at the upper periphery of a language's syntactic structure. In a joint effort to introduce Shanghainese to the academic discourse and to test claims about hypothetical universal structures in an empirical fashion, this study developed two tasks to test some hypotheses about the functional field.

The first study was a qualitative analysis of several Shanghainese sentence-final particles to develop a basis of comparison for this class of items and the analyses that have been offered for other languages. Secondly, testing the assumption that semantic features are reliably encoded in a predictable hierarchy, a controlled acceptability judgment experiment was devised to test whether certain two-particle permutations of sentence-final particles were preferred over their mirrored counterparts. The qualitative study found that the sentence-final particles in Shanghainese share many semantic features in common with those in Mandarin and Cantonese. The acceptability task showed that semantic features of various particles are a statistically reliable predictor for that set's acceptability. The discussion goes on to conclude that the same hierarchy which has been proposed as a description of sentence-final particle behavior in other languages is active in Shanghainese.

CHAPTER 1. INTRODUCTION

1.1 Introduction

Language allows a large number of possibilities for expressing similar ideas which may be encountered by the majority of human societies. Understanding how speakers represent these concepts both internally and externally is a key motivation in modern linguistic research. As an example of interesting phenomena at the forefront of current research are morphemes at clausal boundaries which denote information specific to the speaker or to the informational content of an utterance (i.e. discourse markers). One question to ask is this: If syntax, the process by which sentences are constructed and organized, occurs largely internally to the mind and discourse, the process by which internal ideas can become diffused across individuals and social groups, occur at near opposite ends of the linguistic ‘cycle’ then how much interplay is there between these two linguistic procedures? An example of speakers taking advantage of discourse features to provide meta-contextual content in English (1) and (2) Shanghainese, a language spoken in Shanghai, China, can be seen below.

(1) You're going off to Barcelona, **eh**?

(2) Jing ma-leq yaq ze seubiao **laa** **va**?
Jing buy-PRF one-CL watch SFP SFP
Hasn't Jing bought a watch yet? (We've been expecting it)
(Personal Communication)

These examples depict how speakers of a language can specify additional information, such as their expectations, surprise, or disapproval. While examples like these are very common, they have not been featured as prominent features of the syntactic discussion at large in decades past.

1.2 Context

In one specific subfield of linguistic research, the discourse-syntax interface (discosyntax for short), or the interface between syntactic processes and discourse-related processes, researchers continue to look into behaviors like those seen above, though many subtle effects of discourse markers continue to elude explanation, or the analyses have been exposed to few experimental contexts.

Still more troubling is the fact that many entire languages are underrepresented in the larger corpus of syntactic literature. Shanghainese, a Northern Wu dialect spoken in Shanghai, China is one such language that has been traditionally underrepresented by 'Western' linguistics as a field, and much of what has been written about it has been comparative phonological or sociological studies, leaving very few traces of morphosyntactic or discosyntactic research on the language to ever feel the heat of the press. This thesis will be one attempt at revitalizing interest in the scientific pursuit of understanding Shanghainese better as a language, and as a cultural symbol of a vibrant metropolis. Much of the work on Shanghainese up to this point has been descriptive and

primarily centered on the languages general morphological and phonological structure (Qian 1997, Zhu 2006). This leaves an unfortunate gap in the understanding of Shanghainese grammar, which includes both a formal description of the language's discourse markers, referred to as sentence final particles (or SFPs throughout), and descriptions of the language's syntax.

Certain syntactic descriptions of discourse particles rely partly on work done by Luigi Rizzi(1997, 2008, 2010), and put the features responsible for these discourse properties into the functional projections which compose a rich field at the uppermost periphery of the structure. With the introduction of many functional projections, the Cartography movement (Cinque 1998) came into being which was motivated by the hypothesis that structures may be 'mapped' by tracing movements relative to positions of certain modifiers. This hypothesis, while born from observations of similar ordering behaviors in various typologically diverse languages, has yet to settle a number of questions such as how much variation is permitted, whether parametric variation can account for ordering variation that does occur and whether the feature inventories are universal.

1.3 Shanghainese Discourse Particles

As described above, syntactic descriptions of many Shanghainese phenomena are sorely lacked. Of these phenomena, discourse particles appear as a way to investigate open questions about the structure of the human language capacity as a whole, such as how the syntax interfaces with the discourse components of language. Seizing the opportunity to pursue two simultaneous solutions - namely that of investigating the syntax of the Shanghainese language, and discovering more about the syntax of discourse

markers in general, a research idea was conceived to measure speakers' knowledge of how discourse markers are used in their language.

More explicitly, this thesis aims to fill two gaps in the literature by 1) looking analytically into the syntactic behaviors of Shanghainese discourse particles and extending modern analyses to fit with its behaviors and 2) by expanding the coverage of current models of discourse particles to demonstrate the validity of the model. Utilizing previous syntactic analyses of discourse markers in Italian, Cantonese, and Mandarin in which each discourse particle acts as the head of an 'adverbial' projection in a rich functional field, (as seen in Rizzi 1997, Cinque 1999, Paul 2014 and Sybesma and Li 2007), this thesis aims to reinforce current syntactic models of discourse particles by testing some of the predictions made by the active models against usage data of SFPs in Shanghainese. This thesis also seeks to answer whether the syntactic positions of SFPs in Shanghainese pattern similarly to their counterparts in Cantonese and Mandarin, as the Cartographic enterprise would predict.

This was done using two methods. First, a series of interviews was conducted with native speakers of Shanghainese to compare their interpretations of these particles to the descriptions in an aging literature (Chen 2007, Qian 1997, Zhu 2006), which is already beginning to show surprising differences in the speech of older generations and speakers born in the past three decades (Gilliland 2006, Liu 2012). Secondly, using feedback from this fieldwork and elicitation study, a survey was constructed to test the acceptability of some two-particle permutations versus their mirrored ordering. The scope of the present study is only concerned with a syntactic, language internal analysis of eight Shanghainese discourse particles, though more than simply these eight are discussed in

Chapter 2 to acquaint the reader with the varied uses that speakers can apply discourse particles to in a communicative setting.

1.4 Outline of Thesis

Chapter 2 will be a discussion of the larger context of both Shanghainese language and culture, and its status as a minority language in a city that is industrializing at an unprecedented pace. It will be shown that Shanghainese is actually under the threat of language extinction in a few generations at current rates of immigration and attrition in the community. Following the discussion of the status of Shanghainese, a summary of the state of the art with respect to discourse particles in Sinitic languages is provided to show the reader what gaps still remain which need to be addressed in years to come. Accompanying this section on discourse particles in general, definitions of twelve discourse particles are provided with contributions from grammars of Shanghainese (Qian 1997) and insights from native speakers obtained from interviews and personal communication. Next, the reader will be introduced to the syntax of the Left Periphery (also called Split-CP hypothesis) as described by Luigi Rizzi (1997), and the motivations of the Cartography movement (Cinque 1998, 2008), as well as the relevance of these two frameworks to the discussion of SFPs. Chapter 2 closes with a presentation of the main hypothesis of this thesis, which is that there is a predictable ordering of Shanghainese discourse particles. This ordering can be predicted via direct semantic comparison to particles from Cantonese and Mandarin occupying the heads of functional projections used by the Split-CP hypothesis and Cartography. A potential structural model for Shanghainese is provided based on the above comparisons before beginning Chapter 3.

Chapter 3 provides a comprehensive discussion of the methodologies used in the two sub-components of the project; the fieldwork/ elicitation procedure and the online survey/ experiment. First, a more methodologically oriented summary of the main hypothesis is presented to ground the chapter in the larger discussion about Shanghainese particles. Section 2 goes into detail about the subjects, materials, and procedure of the fieldwork task. Following this outline of the fieldwork task, Section 3 is a thorough explanation of all subjects, materials, tasks that were part of the online survey experiment. The online survey experiment provides an empirical account of the data in support of the hypothesis, so the majority of Chapter 3 will be centered on the construction, distribution and analysis of the survey experiment.

As the last substantive chapter, Chapter 4 provides an organized array of all results obtained from the tasks outlined in Chapter 3 before describing the results of statistical analysis that was used to corroborate the patterns predicted by the hypotheses in Chapters 2 and 3. Beginning with brief recollections of the methodology, the results are presented across two sections. The first wave of reporting centers on the results of the qualitative fieldwork portion and describes some nuances of the particles that confirm certain aspects of these particles' descriptions in the literature while disconfirming others. Since an important assumption throughout this paper is that SFPs are constituted by semantic features, the ability of SFP clusters (e.g. *aa laa*) to occur in either order would be a serious challenge to the proposal of a universal functional (i.e. adverbial) hierarchy (Cinque 1998, 2008). To test if this is the case, an acceptability judgment experiment was designed to test speakers' acceptance of one or either permutations of Shanghainese SFP clusters. In Chapter 4, the results of the survey experiment are presented, beginning first

with general properties of the singleton particles, then showing how the mirrored pairs of two particle clusters fared in the acceptability rating task. With all results disclosed, the original hypothesis will be compared to the findings, and a few modifications will have to be proposed.

Due to the narrow scope of this project, there remain multiple opportunities to expand the current research in the future, such as investigating new combinations of particles, controlling for subject effects, or distributing across a larger subject pool to validate the finer distinctions that did not result in statistically significant conclusions.

CHAPTER 2. LITERATURE REVIEW

2.1 Introduction

This chapter contains a brief overview of the developing body of knowledge surrounding the syntax of discourse particles of Shanghainese. Section 2 provides a basic description of the Shanghainese language, its modern status, and a history of its use. Section 3 introduces the reader to a special class of discourse particles common across Sinitic dialects, the sentence final particle. This section also includes descriptions of 12 common Shanghainese SFPs, their meanings, and comparisons to similar particles in Cantonese and Mandarin. Section 4 discusses two possible approaches to the syntax of discourse particles, in the form of the Split-CP (Rizzi 1997) and the Cartographic approach to syntax (Cinque 1998). Section 5 describes the specific questions this thesis sets out to answer within the frameworks discussed in this chapter. Section 5 outlines the hypothesis and basic methodology used in the formulation of the research question, which is whether or not SFPs in Shanghainese pattern similarly in the syntax with respect to their semantic counterparts in Cantonese and Mandarin

2.2 History and Status of Shanghainese

Shanghainese, or Shanghaihua, has had a relatively limited exposure within the field of linguistics, at least compared to other regional or familial counterparts such as Mandarin or Cantonese. Originally spoken in small fishing villages on the Yangtze River Delta, this Northern Wu dialect was conceived at the mouth of the Huangpu River, where the metropolis-to-be, Shanghai, eventually became a major hub for international business and trade. Though still largely unintelligible to speakers of the standard Chinese dialect, Mandarin, modern Shanghainese has not been unaffected by the influence of more globally dominant languages. While Mandarin boasts some 960 million speakers as of 2010 (Nationalencyklopedin, 2010), as a comparatively small language, Shanghainese claims a population estimated to be between 10 and 14 million speakers.

While this may seem above the range generally considered to be 'at risk', due to economic and social tensions introduced by a rapidly evolving economy replete with large scale urban redistricting, and augmented by the effects of modern globalization, Shanghainese speaking communities are being dispersed and many of the city's youth are opting not to learn the language in favor of more globally marketable languages such as English, which is needed for enrollment in many universities. This is added to the common perception of regional languages as mere 'dialects', which have diminished significance compared to Standard Mandarin, which is used to spread policy in China. (Gilliland 2006) Additionally, many native speaking parents are not using it as frequently in the home, resulting in fewer youth acquiring the language. As a result, many of the young generation of speakers who do use the local tongue speak a markedly different variant of Shanghainese compared to speakers who lead them by as little as a single

generation. Striking differences such as a loss of velar onset in first person-singular pronouns over the course of a generation shows the effects of the newly 'Mandarinized' Shanghainese, or 'New Shanghainese' which refers to the variety of the language spoken by speakers born in or after 1985 C.E. (Chen 2007, Wellman 2012). As a case in point, middle aged speakers (those born in the 1960-70s) will often use the first person singular pronoun *ngu* (pronounced with a velar nasal in the onset), while phonological studies of younger speakers indicate loss of initial velars, whereby the speaker will say *wu*, analogous to Mandarin 'wo' (Liu 2012). This is not unique to Shanghainese, however, as similar alterations are being observed in other languages as well, such as Cantonese (Matthews and Yip 2011).

Combined with controversial municipal policy since 1985 (though now reversed) preventing the use of Shanghainese in schools (Wang 2014), several trends which restrict the number of new Shanghainese learners put the language at risk for extinction within as many as two generations, according to Shanghai native, linguist and activist, Qian Nairong (Reuters 2012).

There is, fortunately, a growing movement in the city led by older speakers, including a few Shanghainese speaking celebrities to preserve the language and encourage youth to learn, as it is considered a a powerful token of the Shanghainese identity (Gilliland 2006). Some city bus lines announce their destinations in Shanghainese in addition to Mandarin, and a popular radio program uses only Shanghainese (China Daily 2012). Shanghainese Opera still enact their performances using a more traditional variety of the language, as well. While many television programs in the region are broadcast in standard Mandarin, one exception is an older Shanghainese

program ‘Lao Niang Jiu’ (Mandarin: 老娘舅), which is still presented occasionally in the native tongue with Mandarin subtitles. However, this program mainly appeals more to the nostalgia of older, already proficient, population rather than passing interest in the language to new generations. (China Daily 2012; Wellman 2012).

While the language does have a number of published grammars (Qian 1997; Zhu 2006) published mostly in Mandarin, the predominant focus of the authors has been comparative and diachronic study of the language’s phonology and morphology with respect to other Sinitic dialects (Qian 1997; Zhu 2006; Chen 2007). Surprisingly little has been done to assuage the need for a thorough syntactic analysis of the language, which elaborates beyond the surface word order and topicalization structures. This thesis aims to contribute an effort toward filling a significant gap in the literature by investigating a relatively new topic in formal syntax, discourse particles, in a language that has long lacked rigorous syntactic description of its behavior.

Shanghainese possesses approximately 10-12 sentence particles as they have been recorded in the Grammars of Shanghainese (Qian 1997). Some speakers do not use certain particles due mainly to sociolinguistic effects, for example, the sentence final particle *tzi* [dz:] is commonly used by older speakers while it is unattested by the younger speakers of ‘New Shanghainese’. As a testament to the influence of Mandarin on regional languages, Shanghainese *leq* is a borrowing of Mandarin le. Qian provided semantic and distributional descriptions of 12 other particles in Shanghainese in his *Grammar of Shanghainese (*上海話語法) (Qian 1997: 214-217).

2.3 Treatment of Sentence Final Particles

2.4 Sentence Final Particles in Shanghainese

This section will outline 12 commonly used Shanghainese particle, along with certain semantic analogues to similar particles used in Cantonese and Mandarin. Although the present study is concerned with only a syntactic analysis of 8 SFPs in Shanghainese, more information is provided here to show the breadth of uses that these particles can take. The particles which are excluded from the present study are those which either do not occur sentence finally, or are known to have restrictive selectional biases for the sentences they modify (Qian 1997)

2.4.1 Leq

As for *leq*, the behavior of this particle will be most familiar to speakers of Mandarin, since it is itself a borrowing from said language. This particle bears a very similar meaning and use as Mandarin *le*, which is to 'emphasize the current state of a scenario'. Particle *leq* is also believed to be a relatively recent borrowing into Shanghainese, and behaves very similarly accordingly. Qian (1997:214) describes it as possessing a tense-like meaning, as an inference, and an expression asserting the current state of events.

(6) *g'eq-g'eq* *tang* *lao* *niq* ***leq***
DEM soup good hot SFP
This soup is really hot!
(Personal Communication)

However (and also observed in Mandarin), there is a potential homophonous relationship between the standard perfective aspect marker and the SFP, which both have the same phonetic form, [leʔ]. It is not believed that these are the same word, as there are

examples of grammatical utterances containing both occurrences. Since the language would not have two instantiations of an Aspect head, it is believed that this second *leq* may be a different morpheme.

See also in example (7) above where the particle *leq* occurs with a stative predicate, which are not commonly known to occur with perfective aspect markers. Mandarin *le* also follows this pattern (8).

(7) *Ming zang tsuo* ***leq*** *gong-jia tsuo* ***leq*** ***nao***
Ming board PERF bus SFP SFP
Ming boarded the wrong bus ("you see?"/ "Let that be a warning to you").
Shanghainese (Personal Communication)

(8) Ta chi le fan le
3s eat PRF rice
He has eaten.
(Mandarin, Taken from Example 5a, Paul 2014, p. 86)

2.4.2 Aa

This particle is used as an interrogative particle, but can also be used as an expression of affectedness, such as intense surprise or pain, or discomfort. This particle also occurs with topicalizations, but while doing so occurs sentence medially, wedged between the topicalized phrase and the rest of the utterance. Qian (1997:215) describes the meaning of *aa* as being highly versatile, with as many as seventeen attested meanings and uses including 'exclamation, appraisal, asking for choices, rhetorical questions, and denoting suspicions and confirmations'.

(9) *Fang doe li sen ti* ***aa****?*
Fang exercise SFP
Does Fang do exercise?
(Personal Communication)

The versatility of Shanghainese *aa* resembles the description of Cantonese *aa3* (Sybesma and Li 2007), who cite that it is used for a variety of purposes, though many of these serve to make the sentence sound 'more natural'. *Aa*3 also reinforces the relevance of an utterance or is used to 'soften' or mitigate the effect of an utterance.

2.4.3 Nao

The particle *nao* is spoken with certainty, and is also used as a reminder to the listener that he or she should be aware of the information in the host utterance, or that it is part of the common ground between the pair (or within the group). *Nao* is listed by Qian (1997: 215) as having meanings such as carrying a threat, a warning, an assertion, a reminder, blame, and emphasis. This particle is also used as an "I told you so", such that the speaker is reminding the hearer that he has been warned in the past, and now it can be seen why the speaker brought the issue up. A common interpretation of *nao*, using (8) as a guide, is that the speaker has repeatedly told Mao to be careful and to pay attention. Now he has been in an accident, and the speaker is pointing out his past warnings.

(10) *Mao sang leq yiq eq tsuo-zi* ***nao***
Mao crash PRF 3s CL car SFP
'Ugh, Mao crashed his car (today).' ("I told him to be careful")
(Personal Communication)

Cantonese analogues to Shanghainese *nao* are *le5* and *wo5*. *Le5* has been described as denoting the fact that the speaker has brought the information up at an earlier time, and that the hearer must not have remembered or acknowledged that they were aware. It is said to carry a similar "I told you so" meaning as Shanghainese *nao* (and occasionally *lao)*, with a hint of reproachfulness. Cantonese *wo3* is also analogous to *nao* due to its adhortative sense. Sybesma and Li (2007) list the 5th and 3rd tones in

Cantonese as ‘minimal meaningful units’ that contribute an epistemic meaning to a word or phrase, as it is a common feature of these particles with that meaning.

2.4.4 Va

This particle is most analogous to the Mandarin particle mǎ (吗), due to its similar role in licensing Y/N questions. *Va* is often used when the speaker wants to turn a neutral declarative statement, and turn it into a question. The questions are ‘neutral’ in the sense that the speaker often doesn’t necessarily have a specific answer in mind, or they are not asking a leading question.

(11) *nong* *hao* ***va***?
2s good SFP
How are you? (Lit. “Are you good?”)
(Personal Communication)

(12) *yiq* *eq* *fu* *mu* *dao* *leq* ***va****?*
3s Cl father mother arrive PRF SFP
Have his parents arrived yet?
(Personal Communication)

It also functions similarly to Sybesma and Li’s depiction of Cantonese *maa3,* due to its transformation of normal declaratives into neutral interrogatives. Other authors (Matthews and Yip, 2011) claim that *maa3* carries notes of formality as well, however this has not been noted in studies of Shanghainese *va*. Other potential analogues to *va* are Cantonese *me1,* though it is also claimed (Sybesma and Li 2007; Law 1990:18) that *me1,* in addition to denoting a YN question, signifies a kind of surprise or disbelief. Unlike Cantonese *me1,* Shanghainese *va* is considered a neutral way of asking questions, though it is not uncommon to use the particle in a surprised or exasperated way. It is not clear

that this is also part of the meaning, and not simply emotive shading from the speaker, since *va* marks the most common way of forming questions.

2.4.5 Laa

This particle is commonly seen as a question marker. This particle also is used as a way of checking a speaker's presupposed knowledge on behalf of the listener. However, *laa* does not always act as an interrogative particle, allowing for the possibility that this particle is potentially a contraction of *leq* and *aa*. The use of this particle resembles the description of Cantonese *laa4* (SandL 2007), in that it seeks confirmation, rather than seeking newer information.

(13)	*Ji*	*bao*	*yiq*	*eq*	*mama*	***laa***
	Ji	hug	3s	CL	mother	SFP
	Ji hugged his mother? (I think he did)					
	(Personal Communication)					

2.4.6 Wa

Wa expresses a mirative reaction to a stimulus, either something heard or observed. Speakers using this particle feel surprised by what they've experienced, though this surprise is not always necessarily negative. Other uses of the particle include use as a rhetorical question marker, such as "is or isn't it?", as attested by Qian (1997:214). *Wa* is another particle that seems to be used only by a subset of speakers. Several younger speakers claim that they do not use *wa*, and it is sometimes confused with *va* by these speakers. It has been indicated by some speakers that this variation may be a result of dialectical differences in certain districts within Shanghai.

(14)	*Jiang*	*chiq leq*	*saq*	*weu*	*wang-deng*		***wa***
	Jiang	eat PRF	3	CL	dumpling		SFP
	Jiang ate three plates of dumpling!						

(Personal Communication)

2.4.7 Lao

Lao is a particle which confirms some information, either clarifying a previous statement or checking common ground knowledge. Qian (1997:214) describes it as marking ‘rebuttals, confirmations and explanations’. It can also be spoken with a kind of ‘sassy’ intonation, in which case it denotes an exasperated sort of reminder. For example, as the subject in (12) below, Ming could be seen as an individual who never keeps his cars for very long, to the chagrin of his friends and family. (12) might be spoken in response to a question inquiring about Ming’s old car, to which the speaker replies that it has been sold, as per usual. The content of expressions accompanied by *lao* are generally well accepted by the public or expected by the speaker, based on past experiences or rumors.

(15) *Ming ma leq yiq eq tsuo-dzi* ***lao***
Ming sell PRF 3s CL car SFP
Ming sold his car (as you know).
(Personal Communication)

Lao matches Sybesma and Li’s description of Cantonese *gaak3,* in that it makes an assertion, though it can be spoken in disagreement as a counterpoint, or as a reminder to the speaker himself.

2.4.8 Ya

Ya often seems to denote a state of affectedness in the speaker. The speaker using this particle often describes themselves as being ‘happy’ while saying the particular sentence. Other speakers claim to use this particles when ‘demanding an explanation’, or when explaining something that the speaker assumed would be obvious. This is

confirmed by Qian (1997:215), who lists an entry for *ya* as something that is restated, rebutted, or exclaimed. (14) below demonstrated the mirative, or happily surprised, use of *ya*.

(16) *WeiSuang men Wufi jihuen* ***ya!***
WeiShuang ask HuFei marry SFP
Wei Shuang asked Wufi to marry!
(Personal communication)

While few Cantonese particles have been described as being an ideal match to particles *wa* and *ya*, Sybesma and Li do mention a class of particles which bear the emotive, surprised response that is found in common uses of the two particles. This class of Cantonese particles was derived by comparing subsyllabic units of certain particles, such as the coda [*-k*], which Sybesma and Li (2007: 1750) claim is a 'minimal meaningful unit' which captures the speakers affectedness and mirative response.

2.4.9 Mo

Qian (1997:215) explains *mo* very briefly as meaning "of course", and that it is in persuasion or negotiation. For other speakers, *mo* is a particle which is most commonly interpreted as meaning 'speaking of...'. This particle never occurs sentence finally, and prefers to license topicalized phrases or other discourse-relevant propositions.

(17) *Gang do yiq teu angli* ***mo****, yiq h'ae lagei laogae li*
Speak 3s steal necklace PRT 3s still in jail P
Speaking of him stealing a necklace… He's still in jail.
(Personal Communication)

2.4.10 Ne

Shanghainese ne can be seen as analogous to Mandarin ne (呢) as Paul describes (2014: 91), as it is also used often in wh-questions and in A-not-A questions. Qian (1997:215) describes it use as an interrogative by claiming it is used to denote 'questions

with choices', rhetorical questions, and questions which the speaker may ask while carrying certain assumptions.

(18) *Nong eq yezi lagei sa difang* ***ne****?*
2s CL keys in what place dSFP
Where did you put your keys, huh?
(Personal Communication)

2.4.11 Lei

Lei is used to emphasize or exaggerate the significance of an utterance. It is used to point out or emphasize some bit of information. Qian's (1997:214) description of *lei* contains several contradictory senses, such as expressing the 'past', a 'process', or the 'future'. Qian also offers the following meanings; a confession, stated assumption, and anger. Other speakers confirm that this particle can occasionally carry some negative emotive connotations, such as "unfortunately", or "I can't believe it".

(19) *Gong ma leq ya zi mae ju eq seu-biao* ***lei***
Gong buy PRF one CL very costly CL watch SFP
Look! Gong bought a really expensive watch!
(Personal Communication)

2.4.12 Clusters

The particles *mo* and *ne* will be excluded from this study on the basis of their conflicting nature with the rest of the class. *Mo* never occurs sentence finally, so it is unlikely to provide insight into the relative position of Shanghainese's SFPs. Likewise, *ne* is excluded due to its restrictive environment of wh-questions, which also limit the particles potential to form clusters.

Shanghainese particles, like those in Cantonese, do possess a limited ability to occur in clusters or 'stacks', though not quite to the extent that the latter enjoys. While

Cantonese has examples cited in the literature of four particles occurring in a single utterance (18), Shanghainese is largely limited to a very small set of two-particle combinations (19).

(20) *keoi5* *lo2-zo2* *dai6-jat1* *ming4* *tim1* *ge3* *laa3* *wo3*
3S take-PRF first-one place SFP SFP SFP SFP
'and she got first place too, you know'
Cantonese (Matthews and Yip 2013: 395)

(21) *XiaoLi* *yong* *leq* *yiq vae* *kuae* ***laa*** ***va?***
XiaoLi spend PRF one-ten thousand RMB SFP SFP
Has XiaoLi spent that 10, 000 RMB yet?
(Personal Communication)

This ability for particles to co-occur in clusters or stacks provides an interesting venue for investigating the placement of semantic features in the syntax, especially in an area of the structure that remains elusive in many dominant languages in the field (e.g. English or Spanish). The morphological expression of numerous functional features is a valuable language trait, and one that permits interesting conclusions to be made about the existence, or lack thereof, of a rich functional field that constitutes the highest 'region' in a derived hierarchical structure.

2.5 The Left Periphery

Rizzi's Left Periphery (As depicted in Figure 1) is based on adverbial modification and particles in Italian. Formed in the aftermath of Pollock's (1989) division of the IP into more specifically motivated functional heads, Rizzi proposed a compound structure made up of projections of a Force head, two Topic heads, and a Finiteness head. In conceiving this analysis of the CP (which is now synonymous with this general 'field'

of syntactic computation), Rizzi drew inspiration from Italian's topicalization and focus structures, as well as its wh-movement.

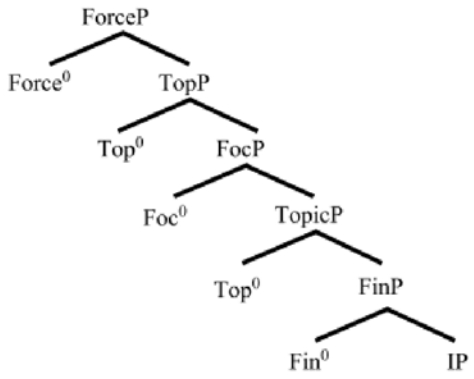

Figure 1: Split-CP (From Rizzi 1997: 297)

The Force head is claimed to be the source of illocutionary force in a language. Force⁰ also acts as a potential rest-area for wh-movement. Topic⁰ is a position that is incredibly useful for describing many phenomena observed cross-linguistically, from topic-comment structures, to topicalization, focus-raising, interjections, etc. Constituents preposed in the Spec-Topic position generally carry a particular intonation ("Boston? No, I've never been."). Focus⁰, however, contrasts with Topic⁰ in presenting items which bear contrastive focus, by being incompatible with wh-expressions, and by being less prolific (in a sense, there may be many topics within a clause (Rizzi 1997: 290). Finite⁰ is the lowest portion of the CP field, where CP interfaces with, and selects the IP, and the finiteness of a clause is determined. This framework, often called the Split-CP, or the Left Periphery, has been incredibly influential in the literature revolving around discourse particles (c.f. Haegeman 2014, Law 2002, Sybesma and Li 2007).

In a similar effort of depicting the numerous possibilities of the syntactization of grammatical concepts and functional features, linguists such as Cinque (1998, 2008)

began attempting to ‘map’ universal functional categories which seemed to appear in many languages, even those as typologically diverse as English, Japanese, French, and Yoruba (Cinque 2008). Efforts to provide a consistent, logical, and universal mapping to an ever increasing multitude of functional projections (again, in the wake of Pollock’s influential *Verb movement, Universal Grammar, and the structure of IP* (1989)) became known as the Cartographic movement. As more and more functional features became known to syntacticians, these features’ distributions defied what one would expect to be a ‘random’ system, as the relative positions of grammaticalized features often occurred in consistent orderings. Where one would, *prima facie*, expect these positions to be governed by Chomsky’s (2005) Second Factor of language design (experience, usage, and historical accident), essentially random, placed in orderings idiosyncratic to a particular language, the cross-linguistic models seems remarkably parallel, despite the mathematical fact that for a language which grammaticalized K from a total of N features (where N is the total inventory of possible features), there is potential for $N! \div (N-K)!$ permutations. As large as N has been speculated to be, there is a much larger design space that seems to be used by the majority of languages, which asserts the need for a principled explanation of why this is the case. What principles of language design orchestrate these correspondences, and what features are bound by the restrictions? To follow Cinque’s strong version of the cartographic hypothesis (2008), a functional head which is realized syntactically in one language must be available (at some point) to all other languages. This hypothesis results in a common maxim for many cartographic investigations, which is the “one property - one feature - one head” (Cinque 2008), such that the position of a particular feature is highly regular and any vacillation of an element

relative to that feature is a symptom of movement. That is, if an element, such as a verb, were to occur in positions both proceeding and following an adverb denoting some property, then the two positions of the verb are distinct, and one may theorize that it is the verb which has been moved.

This relates to current research on Shanghainese particles, as particles have been known to occur in clusters (see section 3.1.12, examples (18-19)), although not all clusters are judged equally by native speakers. This is demonstrated in examples (20-21), where the grammatical combination *leq aa* is preferred to its mirror **aa leq*.

(22) Ming zang-tsuo leq gong-jia tsuo leq aa?
Ming board PRF bus SFP SFP
Did Ming really board the wrong bus?!
(Personal Communication)

(23) *Ming zang-tsuo leq gong-jia tsuo aa leq?
Ming board PRF bus SFP SFP
Did Ming really board the wrong bus?!
(Confirmed via Personal Communication)

In languages like Cantonese, which have as many as 40 specialized particles for particular functions in the discourse (Kwok 1984; Matthews and Yip 2011), it is fairly interesting to attribute various particles to corresponding functional heads that have been proposed by proponents of the Split-CP framework or in the Cartography movement. Languages like English, however, which lack dedicated particles and instead 'recycle' common adverbials and interjections to serve discourse purposes ('You know?", "Yeah", "Please") find it more challenging to match discourse markers to their appropriate functional heads. This problem is compounded by the fact that many discourse particles can be said to change meaning depending on the context. If a particle is beholden to a

certain functional projection for its standard position, what can be said about particles which adapt their semantics to the scenario, especially when a primary factor in determining the proper functional head to house a particle is based on the particle's semantics? This becomes an issue with "one feature - one place", as researchers may be unable to categorize a particle or determine its most salient feature, compounding the issue of having too many places.

2.6 Hypothesis

2.6.1 Comparative Analysis

In order to determine the structure of Shanghainese particles, it is first necessary to determine their semantic features (described first in Qian 1997) and compare these to similar lexical items in other languages which have already been thoroughly analyzed and placed within the functional hierarchy. Once the semantic properties of the individual particles have been established, then it is possible to formulate a hypothesis about their ordering and structure on the basis of their relative ordering and co-occurrence. This also provides a valuable test to validate typological studies of semantic particles which assume universal behavior of functional heads within the Left Periphery.

Using as a guide Sybesma and Li's (2007, p. 1779) proposed structure for Cantonese particles in Figure 4, itself based on Cartographic work by Cinque (1998, 2004) and semantic work on Cantonese particles by Fang (2003) and Fung (2000), the semantic features of Shanghainese particles can be more easily compared to Cantonese and translated into a testable hierarchy.

Figure 4 below depicts a range of minimal meaningful units (MMUs) in Cantonese, which are subsyllabic units believed to carry the semantic features observed

in Cantonese particles. These were derived by comparing particles phonological structures and their semantics and drawing parallels between them via reduction. For example, particles *laak3*, *zek1*, and *gaak3* all convey an idea of affectedness and emotivity, thus causing Sybesma and Li (2007) to infer that it is the {-k} morpheme which carries this meaning.

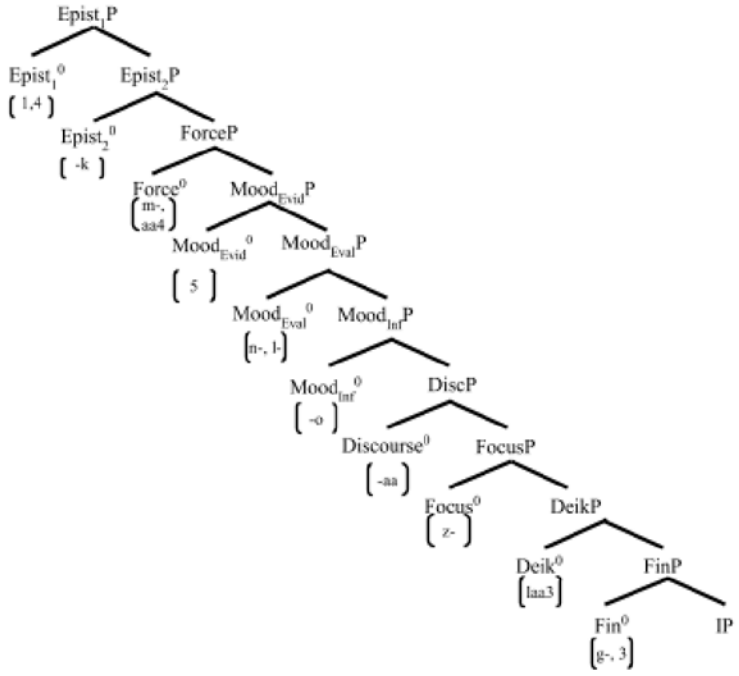

Figure 2: Cantonese MMUs (From Sybesma and Li 2007: 1779p)

2.6.2 Featural Comparison

Using features as a guide for comparison, the hypothetical ordering of Shanghainese particles (lowest/first to highest/last) is; tense and aspect related particles (*leq*); mood and emotive particles (*wa*, *ya*); force particles (*va*, *laa*); and epistemic particles (*nao, lao*). This ordering is more clearly represented by the diagram in Figure 2 and Figure 3.

In section 3.2 above, semantic comparisons were given between Shanghainese particles and their Mandarin and Cantonese analogues, where information was available in the literature. No native speakers of Cantonese were directly consulted to verify the detailed descriptions of particles given in the literature, though most published research about particles has been done *by* native speakers. The approach to comparing the particles of the two languages was to find the closest semantic correspondences, and to use the minimal meaningful units attributed to those particles in Cantonese by Sybesma and Li (2007) to determine their probable position. In cases where multiple particles with different components seem to match, but the units correspond to different positions, then the unit which contributed most to the Shanghainese particles' general meaning was used. As an example, *nao* was likened to both *le5* and *wo3*, which use the units {5, 3, -l, o}. To resolve this issue, it was deemed that all of {o, -l, 5} fall within the Mood field, and that these three capture the largest portion of the meaning, while the {3} is not a significant factor and is in fact hypothesized by Sybesma and Li to be a generic or default tone (2007: 1778)

Table 1: Cross-linguistic Comparison of Particles

Shanghainese	**Cantonese**	**Mandarin**	**Position**
leq	*laa3*	*le*	Deik/Fin
aa	*aa3*		Discourse
nao	*le5, wo3*		$Mood_{Inf}$/$Mood_{Evid}$
va	*me1*	*ma*	Force
laa	*laa4*		Force
wa	(*-k*)		Epist
lao	*gaak3*		Epist
ya	(*-k*)		Epist

Using data from an interview task and an online acceptability judgment survey, this thesis will present an analysis of 8 Shanghainese discourse particles and compare them in both syntax and semantics to those particles from other, more studied languages and see if the hypothetically universal ordering still holds true for Shanghainese.

Starting from lower in the tree and moving up (the top of Table 1), Cantonese *laa3* most closely corresponds to *leq* because of their signaling of finiteness, and the expression of a current state of affairs. Cantonese *aa3* and Shanghainese *aa* are alike on one respect for their overall versatility, but more specifically for their portrayal of a very 'natural', emphatic, or mitigative mind state on behalf of the speaker. Shanghainese *nao* has been likened to both *le5* and *wo3*, although no single Cantonese particle matches all uses of *nao*. *Le5* matches the previous acknowledgment conditions of *nao* while *wo3* shares its use as an adhortative. The similarities between Cantonese *laa3* and Shanghainese *laa* consist of their use in posing interrogative or confirmatory utterances,

which bely the speaker's pre-existing assumptions. *Va* and *me1* share the basic function of forming neutral Yes-No questions. *Lao* and its Cantonese counterpart *gaak3* are endowed with the sometimes indignant, sassy, reminder, and rebuttal meanings, at least as has been described by their uses in the literature. Shanghainese *wa* and *ya* were the most difficult to match; *wa* especially, as a result of its not being used by a portion of the population. However, as speakers described it, no single particle matched well enough to draw a comparison. The closest match to *wa* and *ya* is the minimal meaningful unit {-k}, which expresses the emotivity and affectedness that is shared by several Cantonese particles which use a checked coda. Interestingly, *ya* shares a sense of being a repeated tellings or reminders with *lao,* which was compared to *gaak3*. Note the {-k} in *gaak3,* as well. A chart has been provided in Table 1 which summarizes the comparisons being made between Shanghainese and Cantonese, as well as the structural positions of the most salient semantic units according to Figure 2 above. Finer distinctions between the respective particles is given in Section 3.2.

2.6.3 Predictions

Two important things to note about the relative simplicity of the diagram in Figure 3 compared to Sybesma and Li's in Figure 2: First, since Shanghainese has many fewer particles than Cantonese, the positions of the particles are not likely to fill all functional heads claimed to be used in Cantonese. Secondly, Sybesma and Li's main goal in devising the diagram in Figure 2 was to decompose the Cantonese particles into their 'minimal meaningful units' (MMU), or subsyllabic morphemes which agglutinate during spell out yet base generate in specific functional heads.

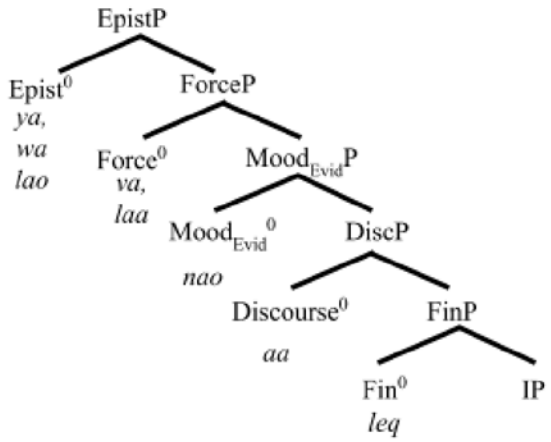

Figure 3: Preliminary Hypothetical Structure

Thus, the morphemes in Figure 2 do not represent whole particles. This analysis of Shanghainese does not yet venture so far as to decompose the particles into their minimal meaningful units (their existence thus far undetermined). More importantly, this being a preliminary analysis of Shanghainese, it is more important to focus on validating the most simple, descriptively adequate model before introducing further complexity. In this respect, it must first be determined whether there are word-level patterns to be observed in Shanghainese particles. All of the names of the functional heads in Figure 3 were borrowed from Sybesma and Li (2007: 1779p)

Using the notion of positional transitivity, in which one may predict that some particle, A, precedes particle C if A always precedes B and B always precedes C. Thus we predict the following, based on Figure 3:

(22) a. **Predicted ordering** (First to Last / Low to High)
leq > aa >nao> va, laa> lao, wa, ya

(23)

a.	*leq aa*	BUT	**aa leq*
b.	*aa ya*	BUT	**ya aa*
c.	*nao va/ nao laa*	BUT	**va nao/ *laa nao*
d.	*va lao/ laa wa*	BUT	**lao va/ *wa laa*
et cetera...			

To briefly conclude the hypothesis and goals of the study; after verifying the semantic features of the particles in an independent setting, combinations of SFPs will be measured in an acceptability judgment experiment to determine valid and invalid combinations, which by the transitive property will determine a relative ordering for all particles and shed light on the validity of the hypothetical ordering in (22).

While there are 8 particles of interest to the current project, it was impossible to test all 48 two particle permutations of the 8 particles (omitting items such as *aa aa*), due to concerns of subject fatigue. In a further reduction of scope, only 18 permutations were tested which were expected to shed light on the transitivity of the functional heads depicted in Figure 3. These 18 permutations consist of 9 pairs expected to be acceptable, and their mirrors, which are predicted to have a degradation in acceptability when rated by native speakers of Shanghainese.

2.7 Summary

This chapter began with some information about the Shanghainese language; about its history and about its current status as an endangered and undocumented language. While it has long been underrepresented in many linguistic subfields, such as semantics or syntax, there is a need to give the language exposure and legitimacy as both an engine for the transmission of a cultural legacy, and as a contributor to the understanding of the human language faculty in general. Following this, the concepts of discourse markers were introduced, allowing for Shanghainese's rich inventory of

discourse particles to be individually discussed first within the context of the language itself, then in the context of other languages with similar discourse morphology. With these semantic and pragmatic uses accounted for, it was essential to provide some theoretical background on the methods of analysis used when approaching similar elements in other languages. Two theoretical models which propose potential explanations of this behavior were explained. Luigi Rizzi's Split-CP and Guglielmo Cinque's Cartography offer the most reasonable approach to the study of discourse markers in Sinitic languages, particularly Shanghainese, due to their explanatory power shown via application to several languages and their rich inventory of features which can serve as carriers of the common SFP meanings. The application of these ideas was demonstrated in Section 5, wherein the hypothesis was presented. This paper seeks to test the hypothetical universal ordering of certain discourse features against usage data in Shanghainese. It is hypothesized that particles which share certain semantic features with Cantonese and Mandarin particles will occur in the same order that their semantic counterparts are used.

This prediction allows for the construction of a falsifiable model of the universality of discourse features and the ability of the Split CP model and Cartographic features to be extended to all languages. If the predicted ordering is indeed preferred by native speakers in an acceptability rating task, then it will be concluded that these innate structures do hold true of Shanghainese, and that the model is shown to have additional empirical coverage.

CHAPTER 3. METHODS

3.1 Introduction

It was discussed in Chapter 2 that, before this point, there has been no formal structural analysis of SFPs for the Shanghainese language. In attempting to provide such an analysis, several assumptions are made. Most importantly, it is assumed that all human language makes use of hierarchical syntactic structures that can be detected and described using various tests. Secondly, it is assumed that this structure is formed by the iterative merger of syntactic heads. This relates to the current study, as much like analyses of Cantonese and Mandarin particles (Sybesma and Li 2007; Paul 2014), SFPs are assumed to be overt manifestations of syntactic heads. Since the assumption of iterative merger necessarily implies a process in which some operations must take place before others, that is to say, that not all operations happen at once, then there is expected to be an asymmetrical relationship among various particles with an effect on the hierarchical structure of an utterance such that one particle must be dominated by another. Also, it is assumed that these asymmetries manifest in the linearization or phonologization of a structure, which permits a researcher to investigate the hierarchy on the basis of the surface structure.

3.2 Research Question and Hypothesis

Therefore, using Kayne's Linear Correspondence Axiom (1994), the linear transitivity of Shanghainese particles is assumed to be directly related to the particles' relative positions in the structure. Note that there is a hidden assumption implicit in the notion of linear transitivity: Namely it is assumed that there is no additional movement after the particles have been merged into their positions. Also related to the predictions made, and tested experimentally in this chapter, is the notion of a universal functional field, such as that proposed in Rizzi 1997 and Cinque 2008. This field is speculated to be utilized by many (if not all) languages for organizing complex hierarchies of functional features. Research Question and Hypothesis. A reasonable assumption in attempting to sketch a basic outline of the structure is the assumption of "one particle-one position". This is related to the assumption listed above that there is no covert movement; however there is a subtle difference. One particle-one position also assumes that the relative ordering of the particles in the surface structure remains stable, permitting the two-part hypothesis that SFPs 1) occur in a predictable ordering, and 2) that some orderings are better formed than others. Hence, assuming that there is indeed a structure to be detected, it remains to be proven what shape this hypothetical structure would look like in an analysis of Shanghainese particles. A series of two experiments were run to investigate the relative positions of SFPs in Shanghainese.

Due to the wide empirical coverage enjoyed by Cantonese SFPs, and the aforementioned hypothetical functional field used to organize functional features across languages, it was possible to test the predictions of such wide reaching hypotheses using a language which likewise exhibits a rich field for inflection of discourse, such as

Shanghainese. Matching Shanghainese particles to their semantic and functional analogues in Cantonese and Mandarin, predictions were made about the ordering among pairs of SFPs in Shanghainese such that structures in both languages which bear a particle with Meaning A would be preferred by native speakers when they precede a particle with Meaning B. As a corollary, structures in which the particle bearing Meaning B would be dispreferred when it precedes particle with Meaning A.

More precisely, following from the given Shanghainese Cantonese analogues given in Chapter two it was hypothesized that the following pairs in Group 1 would be preferred to the mirrors exhibited in Group 2, as shown in Table 2.

Table 2: Predicted Permutations

Preferred	Dispreferred
aa ya	ya aa
leq lao	lao leq
leq aa	aa leq
leq va	va leq
leq wa	wa leq
nao laa	laa nao
laa va	va laa
laa wa	wa laa
ya lao	lao ya

These hypotheses were tested using two tasks involving native speakers of Shanghainese: an acceptability judgment task administered as an interview task conducted in person (Section 3) and an online Qualtrics survey (Section 4) controlled to elicit judgments about all 18 of the pairs in Table 2.

3.3 Task I – Fieldwork Exercise

This project was undertaken in two experimental stages. The first of these experiments was a qualitative elicitation and field study of SFPs in Shanghainese. The objective was to recruit participants for informal interviews and obtain basic semantic, pragmatic, and distributional information about the particles. These interviews served two purposes. Firstly, this initial fieldwork ironed out any lexical or semantic issues with all of the host sentences to be used in the survey experiment below. Secondly, the aim of the fieldwork was to learn more about the usage of these SFPs through a more qualitative analysis of the particles.

3.3.1 Subjects

Approval for Human Subjects research was granted by an Institutional Review Board, under Protocol #1410015420. Compensation was approved for participants in Experiment 1. Participation in Experiment 2 was voluntary, due to the logistical challenges of compensating large numbers of participants in multiple geographic locations. To make up for the lack of compensation, the survey was designed to require much less input and effort from the participants. It was estimated that the survey would take on average less than thirty minutes, and in fact many subjects took twenty or less to complete the survey.

Subjects in the first stage of the project were recruited via flyers posted around the Purdue University West Lafayette Campus, offering compensation for participation in an interview that would not take any longer than a single hour. Participants were compensated $8 USD for their enrollment in the interview portion.

Five native speakers of Shanghainese agreed to participate in the interview process. The ages of the interviewed parties ranged from 19 to 32, and were a mix of undergraduate and graduate students. All of these participants were bilingual in Mandarin and Shanghainese, in addition to fluency in English. All five of the speakers fall in the age range considered to be speakers of New Shanghainese. The gender distribution of all interviewed participants included three males and 2 females. Four subjects spent their whole adolescence in Shanghai and spoke Shanghainese in the home, the fifth spent a smaller portion of their childhood in Shanghai, though also used Shanghainese regularly in the household in the United States.

3.3.2 Materials

All dialogue between the participants and the researcher was recorded on a SONY ICD-PX333 Digital Audio Recorder for later reference. With the goal of finding out as much about each particle as possible from a native speaker's perspective as possible, as well as pilot a few of the combinations suspected to be grammatical, the participants in the fieldwork portion were shown all 18 host sentences used in the survey portion. A list of all questions included in the survey is found in Appendix A. These were randomly assigned to all 9 singleton particles twice each, as well as six hypothetical permutations, which were also exposed twice. This makes for a total of 30 test sentences, which is not evenly divisible by 18. Thus, not all sentences received equal exposure in the fieldwork, though this is not seen as a major limitation of this portion since the primary goal of the fieldwork exercise was semantic elicitation.

3.3.3 Procedure

The procedure for the fieldwork portion began with the obtaining of informed consent. All participants were shown an IRB approved informed consent form which included a detailed description of the goals of the research and procedures to be followed. Afterward, subjects were informed that the interviews would be recorded and that all participation was voluntary, and that they may back out or refuse to answer a question at any time. If they signed the consent form, the interview process began.

The procedure was as follows: Firstly, the researcher (not a native speaker of Shanghainese) read the sentence aloud to the participant. It was then asked whether it was found comprehensible or acceptable by the speaker. This done, the researcher then read the same sentence aloud with either a singleton particle or a permutation of particles. Then the speaker was asked whether this was an acceptable formation, and what the difference between the bare sentence and the particle bearing sentence was. They were asked to describe how the meaning of the bare sentence had been affected, as well as how the speaker may feel or what the speaker believes as they say the utterance. Each sentence was repeated as many times as was requested by the participant. Any commentary provided by the participant was also noted in a journal.

With respect to time, the fieldwork portion was ongoing during the creation of the survey. Four of the five participants in the fieldwork exercise had been interviewed before the survey was finished, with the final participant being interviewed post-construction. This may be seen as a potential limitation for the construction of the survey as the original motivations for the fieldwork were to provide a better guide to which of

the permutations to include in the survey portion, as it was not feasible to test such a large number of potential permutations of the nine selected particles.

3.4 Task II – Acceptability Judgment Survey

3.4.1 Subjects

Participants in the Qualtrics survey were recruited through personal email connections and through online social media. By the nature of the recruitment methods, there was a bias in the recruitment toward younger speakers, which by and large fall into the generation of speakers speaking a more Mandarinized variety. Emails describing the survey, complete with basic instructions and an estimated time limit were sent to several peers currently living in Shanghai, who were then asked to forward the email to their own families and friends. Where this survey differs from many linguistic studies is that no subjects were recruited on a university campus, and several participants did not attend a four year institution. While all participants were native speakers, there is a believed to be greater variety in age, income, and education that is not found in other linguistic samples, such as those where subjects are recruited exclusively on a university campus. However, income and education were not explicitly inquired about in the background questionnaire shown to respondents.

All demographic information was self-reported in a background questionnaire at the beginning of the survey. The questionnaire inquired about the age, sex, home district, other languages spoken, a self-reported measure of the subject's confidence using Shanghainese, the amount of time that the speaker has either lived or spent in Shanghai during their lifetime, and the subjects' estimated frequency of use of the language (e.g.

every day, a few times per week, monthly, seldom). A copy of the Questions asked in the Demographic questions can be found in Appendix AI.

Across all 36 participants, the mean participant age was 34.8 years with a range of 50 years. The youngest participant was 21 years old and the oldest was 71 years old. Regarding the sex of all participants, there were 23 reported females and 13 reported males. This age was used to divide the participants into subgroups. Age group one consisted of speakers born after 1980, while age group 2were those born after. 1980 marks major shifts in attitudes considering what is a proper or useful language to use in schools. (Wang 2014, Gilliland 2006)

As for language, 34 of the 36 indicated knowledge or regular use of Mandarin. 22 of the participants report knowledge of English. 2 speakers report regular use of Cantonese. 4 participants have knowledge of Japanese, and 2 participants have knowledge of Korean. One participant has reported knowledge of four non-Shanghainese languages.

No participants report spending fewer than 10 years in Shanghai over their lifetimes. One respondent reports living in Shanghai for between 11 and 15 years. Three participants report living in Shanghai for between 16 and 20 years. Ten participants report residency in Shanghai lasting between 21 and 25 years. Nine other participants report a figure between 26 and 30. For those residing in Shanghai between 31 and 35 years, there are three participants. Only seven participants claim to have lived in Shanghai for a duration of longer than 41 years.

Participants were also asked to list their primary district of residence. Of all participants; Four claim Yangpu as their home; Ten from Hongkou; Six from Zhabei;

Three from Putuo; Three from Changning; Five from Xuhui; Three from Jingan; Two from Huangpu; Five from Pudong; Four from Baoshan; One from Luwan; Two from Minhang; and One from Jiading. Participants were allowed to select multiple responses for this question. Only 9 participants selected two districts for this question, while two participants selected as many as 3 districts.

It was also requested that subjects give a self-reported estimate of their comfort and experience using the language using a seven point scale. For the comfort scale, 1 marked the minimal possible comfort, equivalent to being unable to speak while 7 marked maximal comfort, equivalent to being a confident, fluent speaker. One potential confound with this measurement is the inclination for a speaker to be modest. However, this effect, if present, did not appear to have a strong influence, as the mean comfort rating reported by participants was 6.09 out of 7. Of all 36 participants, 18 participants rank themselves at a 7, while 15 participants rated themselves at a 5 or 6. Only three subjects identified their comfort with the language at either a 3 or 4 on the scale. No participants rated themselves at a 2 or lower in terms of comfort using the language.

As a related measure to comfort using Shanghainese, respondents were asked to estimate how frequently they find themselves using Shanghainese as a mode of communication. This question allowed for four categorical responses; ‘Everyday’, ‘Often’ (a few times a week), ‘Occasionally’ (A few times per month), and ‘Seldom’ (a few times per year). Of all 31 responses, only 5 participants rated their frequency of use as ‘Often’ and only a single participant listed their frequency of use as ‘Occasionally’. All 30 of the remaining participants claim to use Shanghainese on a daily basis.

Location information was recorded by the survey software for all survey participants. Participants were recorded taking the survey from terminals in the following locations; 27 in Shanghai, China; 1 in Dayton, Ohio; 2 in Lafayette, Indiana; 2 in Atlanta, Georgia; 1 in Chicago, Illinois; 1 in Pittsburgh, Pennsylvania; 1 in TaoYuan, China; and 1 in HeFei, China. All participants involved in the taking of the survey listed their home districts in an earlier question. It is assumed that the answer to the district question marks the district in Shanghai in which the participant has lived in at some point.

Since the survey was taken anonymously online, no method of contacting the participants for a follow-up was included, such as including a contact information question in the demographic portion of the survey. The contact information of the principal investigators was provided at the beginning of the survey for participants to request further instruction, a follow-up, or to obtain a copy of the final results. However, only a single participant made use of the contact requested the final results, and no subject requested additional instruction.

3.4.2 Materials

To investigate the co-occurrence and ordering restrictions of SFPs in Shanghainese, an experiment was designed to test their acceptability in contexts where ordering of certain two particle combinations was reversed. The independent variables in this experiment were the number and relative ordering of Shanghainese particles, as well as the specific particles which are used. Stimuli in this experiment contained either a singleton particle or a two particle cluster. The Number condition had two levels; singleton particle sentences and dual particle sentences. The ordering of particles was defined as the relative ordering within the pairs (e.g. AB and BA). As such, the Order

condition had two levels. Experimental block was also treated as an independent variable, with two levels. In terms of a Latin square design, there were six cells, since singleton particles have one level of ordering, and each combination occurred once in both blocks.

The dependent variable in this study was the perceived acceptability of each stimulus. Acceptability was defined as any native speaker's willingness to accept a given linguistic utterance as a well-formed sample of their language. This is related to the notion of grammaticality, in that grammaticality is defined as the likelihood that a given formed is a possible product of the language's grammar. Grammaticality, thus, is a theoretical notion and is not directly accessible while acceptability is easily measured in judgment tasks. In order to obtain information about a speaker's measure of acceptance toward a given stimulus, all stimuli were accompanied by a seven point Likert scale with integer values. The lowest value on this scale, a one, indicated that the sentence was wholly unnatural, while a seven, the highest point, described a sentence at its most natural. Participants were instructed to select the radio button located under the respective value that they felt best described their feelings for each stimulus.

This was a within subjects, repeated measure design, so each participant was exposed to each permutation, mirror, and singleton twice - once each per block, two total across contexts.

The format of the experiment was an online survey. This survey was hosted by Qualtrics, and designed using Qualtrics web-based survey software. This software was chosen due to the intuitive, minimalist aesthetic experience on behalf of the user. Mechanically, it also enabled embedded audio media to be presented along each stimulus.

Furthermore, the software offered built in randomization for all stimuli in the experimental blocks.

To limit the scope of the study, nine Shanghainese SFPs were chosen based on their semantic contribution, frequency of use, and sentence final status. Information on semantic content of the particles was obtained in Experiment 1 during one on one interviews with five native speakers of Shanghainese. This process is discussed in Section 2 above. Due to lack of a large corpus of spoken Shanghainese, frequency of use was not measured quantitatively, but was based on native speakers' intuitions of their frequency. The included particles are *va*, *aa*, *laa*, *nao*, *lao*, *leq*, *wa*, *ya*, and *lei*, because they have been known to form two particle clusters in natural speech.

In order to test co-occurrence of these particles, it was concluded that the minimal requirement for co-occurrence is a combination of at least two particles. Since relative ordering was hypothesized to be a significant factor affecting the grammaticality of these combinations, the optimal experimental units were drawn from the 63 possible two-item permutations of the nine particles (9 nPr 2=72, and excluding permutations consisting of duplicate particles such as 'l*aa laa'*). Additionally, in order to reduce the burden of the acceptability judgment task for participants, it was decided not to test all possible permutations and to restrict the scope to a few carefully selected permutations which contribute most to the validation or falsification of the hypothesis. Out of the 63 permutations, 9 unique pairs were chosen along with their respective 'mirrors' (the reverse of the permutation, e.g. '*leq va'* and '*va leq*'). All stimuli can be found in Appendix A, organized by block. It was also decided to use singleton particles as a control for the overall acceptability of the host sentences.

In this paper, ‘host sentence’ refers to a bare sentence to which a particle or combination thereof may be added. 18 host sentences were constructed by consulting a native speaker to ensure semantic congruity. To help ameliorate the effects of hearing these SFPs ‘out of the blue’, These 18 sentences were divided evenly into two imaginary scenarios, which were constructed to serve as a setting for the host sentences and provide the speakers a way of interpreting them. Nine of the sentences were about things one may hear while waiting at a bus stop, and nine sentences imitated potential utterances to be overheard in a crowded shopping mall. The scenarios can be seen in (1), as it was provided as part of the instructions for the acceptability judgment rating task.

(24) “You are waiting in line in a bus stop in Hongkou District. You are listening to what other passengers say around you until the bus arrives. The following are a few things you may hear them say. Please rate how natural each sentence sounds to your ears using your native intuitions. A rating of 1 marks the least natural sounding, while 7 marks something that sounds very natural.”

(25) “You are taking a rest in a shopping mall in downtown. You are listening to what people are saying around you. The following are some things that you hear spoken by the other people. Please rate how natural each sentence sounds to your ears using your native intuitions. A rating of 1 marks the least natural sounding, while 7 marks something that sounds very natural.”

The two contexts served as the primary semantic influence for all of the host sentences. To list a few examples, all nine host sentences from block 1 related to life in the shopping mall; withdrawing money from an ATM, buying a new watch, a broken down elevator, *et cetera*. These contexts also served as the main themes of each experimental block, each context being used for one block.

Each of the host sentences was used thrice in its respective block. As mentioned above, 9 two-particle permutations, 9 singletons, and 9 mirrors were selected as

experimental conditions. To ensure comparability among the permutations, the singletons, and the mirrors these were each randomly assigned to a host sentence in such a way that each host was not used more than once for each category. In other words, one of the 9 unique hosts was randomly paired with a unique member of the set of 9 permutations, the same host was randomly paired with a random member of the mirror set, and finally paired with a random member of the singleton set. None of the host sentences were repeated within any of the three conditions (Number and Order).

The host sentences were also controlled for tense, by using past tense or perfective examples. However, since Shanghainese syntax requires mention of aspect in order to maintain grammaticality, this often resulted in host sentences containing the perfective aspect marker, *leq*, which happens to be homophonous with one of the SFPs under investigation. During the randomized matching of particles to host sentences, it happened that some sentences generated were either ungrammatical or nonsensical, due to haplological occurrences of *leq*, as in (26) and (27) below. A demonstration of each type of sentence used in the experiment can be seen below in (28).

(26)	Host:	Fang	sanbei	sii	dou	**leq**	
		Fang	work	late	arrive	PRF	
		Fang came to work late.					
(27)	Host+Pair	*Fang	sanbei	sii	dou	**leq**	*leq wa*
		Fang	work	late	came	PRF	PRT PRT
(27)’	Corrected	Fang	sanbei	sii	dou	**leq**	*wa*
		Fang	work	late	came	PRF	PRT

Thus, each of the two experimental blocks consisted of 27 experimental stimuli, for a total of 54 items. Within each block, host sentences containing each of the two particle permutation, each of the respective mirrors, and one of each singleton particle

was represented once. This ensured that each of the twenty seven conditions was encountered by each subject no more than twice. Using this design, it was also possible to compare the permutations to their mirrors in two unique contexts. As mentioned above, the singleton particles were used in order to allow the estimation of a baseline acceptability for the host sentences.

To present all 54 unique stimuli to the survey's respondents, all were recorded by a native speaker. This was done due to the fact that Shanghainese shares its writing system with the standard dialect, and there was concern of potential influence from the orthography on the results of the acceptability task. 54 audio recordings were made using a SONY ICD-PX333 Digital Audio Recorder, and uploaded into a Qualtrics library to be embedded in the survey. It was decided to record the host sentences plus stimuli as single files (rather than recording all hosts and permutations separately to be put together randomly) in order to control for lexical effects that may have intervened otherwise (such as haplological *leq*).

(28)	Host:	Lei	nao leq	jiao gwei	tsaopiao!
		Lei	take-PERF	very many	cash
		Lei withdrew a lot of cash!			
	Singleton:	Lei	nao leq	jiao gwei	tsaopiao **laa**!
	Permutation:	?Lei	nao leq	jiao gwei	tsaopiao **laa va**?
	Mirror:	?Lei	nao leq	jiao gwei	tsaopiao **va laa**?

Acceptability of each stimulus was measured using a short instructional passage and a seven-point scale with radio buttons to prompt speakers to provide their judgments. Speakers were presented with a small passage that specified a context to interpret the utterance, which was contained in the short audio clip recorded by a native speaker. The instructions also requested that the participant use their native intuitions to provide an

estimation of the naturalness of the utterance. The subjects were provided a scale with clickable radio buttons for each integer value along the seven point Likert scale, where a value of one meant that the utterance was completely unnatural and a value of seven meant that the utterance was "very natural". The acceptability rating task is related to one form of noncompliance during the survey. A few participants provided anomalous ratings, such as rating no single item higher than a three, or by rating every item at the maximum value provided. The results of three participants were excluded for providing anomalous ratings such as these, since the average rating for each item demonstrated high degrees of variance within the group.

All instructions, contexts, consent forms, and the like used during the survey had been translated from English into Mandarin by a native speaker. Mandarin was chosen as the language of instruction due to the fact that both Shanghainese and Mandarin share writing systems, and it proved more efficient to use written instruction on a computer screen than to rely on audio clips for all non-stimulus materials.

The final data was exported into a Google Sheets spreadsheet from a report generated by Qualtrics, and analyzed using SAS (Statistical Analysis System). Demographic information was all calculated using summary statistics. Single-particle items were analyzed separately from two-particle items. For both analyses, a three-way ANOVA was used to test for effects of order, block, and age group. To examine the two-particle items in more detail, a two-way ANOVA was used to test for effects of the linguistically interesting variables, order and age group, within each particle pair. For both the single-particle and two-particle analyses, participant was run as a random variable in the script.

Since the order of the blocks was not randomized among participants, it was essential to measure the likelihood of block effects, such as those induced training and/or fatigue.

3.4.3 Procedure

The procedure followed was standard for that of a survey. A graphic has been provided below in Figure 4. All participants were given a link to follow during recruitment. Upon following the link, the participants were greeted by a copy of the digital consent form. This form relayed all necessary information about the goals, benefits, and risks inherent in the study. Respondents were made aware of the approximate duration, as well as the completely voluntary nature of the survey. Contact information for all researchers was provided, and it was made known to the respondents that they may request an anonymized version of the final results of the survey if they should like to have it. If respondents indicated that they consent to the survey, they were taken to the Demographic stage of the survey. Respondents failing to provide consent (i.e. by clicking the "No, I do not consent." option) were taken by the software directly to the end of the survey.

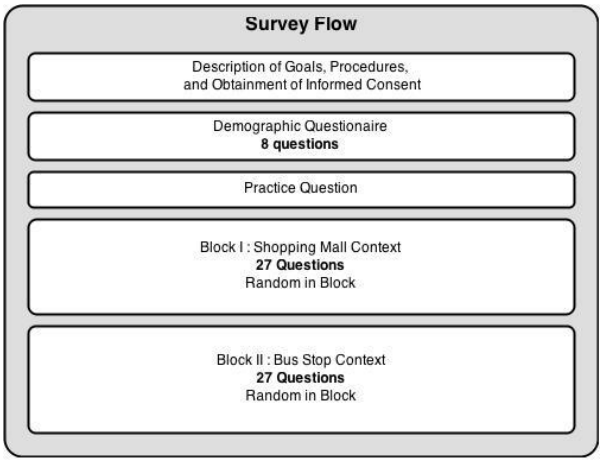

Figure 4: Survey Design Diagram

During the Demographic Questionnaire, subjects were asked to answer 8 questions about their background, including age, gender, languages spoken, time spent in Shanghai, home district in Shanghai, familiarity with the language, their frequency of usage.

Next, participants were given a single practice item which described in detail how to perform an acceptability judgment. It described that a seven point scale would be used, and indicated the extreme values of the scale, along with implications for how to implement the scale. (e.g. a quote from the instructions "You must select a rating on a scale between 1 and 7 to judge how natural the sentences sound to you. Seven is most natural. One is very unnatural sounding."). On the practice item, the speakers were asked to try the task by listening to a short transitive declarative clause with no particles. An identical scale was used in practice items as in the rest of the task. Upon answering, the survey progressed to the first block.

Block 1 and block 2 were presented in consecutive order for all respondents. Within each block, however, all stimuli were randomized in an order unique to the participant. There was no restriction on the randomized order except that the 27 Mall Context sentences were not mixed with the 27 Bus Stop Context sentences. The sentences were provided one-at-a-time, and participants were permitted to replay audio clips if necessary. This was permitted to control for potential technological failures.

When the final item in block 2 had been answered, all participants were thanked for their participation, and the results were collected in a report hosted in Qualtrics.

There was some attrition during the participation of this study. Several individuals began the study and did not complete the study. Two individuals failed to provide

consent when prompted, and were taken directly to the end of the procedure. Nearly thirty individuals failed to complete the survey for unknown reasons. These individuals completed between 3 and 58 percent of the total survey before failing to continue. While the distance between the researcher and the participants created by the use of a hosted online survey obscures the majority of causes resulting in attrition or noncompliance, there is one factor which is known about a small subset of cases. Due to the fact that Qualtrics surveys are also accessible by users browsing the internet on mobile devices, there is one relevant factor which impeded participants in this survey, which was not considered in early trials. Several users began the survey on their phones, however, since Qualtrics is a Flash-based interface and many mobile phones such as Apple and Android devices are incompatible for Flash media, the audio clips used in the survey were unavailable to participants. This factor is relevant for cases wherein the participants stopped after the eighth question, because the ninth item in the survey was a practice item used to familiarize the participants with the media playing interface and acceptability judgment task. Also, as many as four individuals stopped at precisely 58 percent of the survey, which happens to be exactly in between the first and second blocks. It is suspected that some technical glitch may have prevented them from accessing the rest of the survey. Since there was no common geographic, demographic, or technological factor which could be linked in all cases of attrition, with the exception of mobile phone users and the Flash errors, any range or combination of factors beyond the ability of the researcher to compensate for could have contributed to attrition in the study. Examples of relevant factors could be; lack of familiarity with Internet computing, interruptions as a

result of daily life, network errors, miscellaneous hardware compatibility issues, task comprehension, or patient fatigue and boredom.

Some limitations of this study include recruitment methods, patient fatigue, and side effects of the online survey modality. Another potential limitation of the online survey is that it was impossible for the researcher to accompany the participant during data collection. Thus, there was no ability of the researcher to answer questions or clarify instructions. It was also possible that some of the participants were unfamiliar with the Qualtrics interface, in addition to the common limitations of linguistically naïve subjects being exposed to uncommon or abnormal constructions in their own language. Several participants commented that they felt fatigued by the process of providing acceptability judgments. This could have also been a factor of the duration of the test. Average test time was near twenty minutes, although participants were asked to provide acceptability ratings for 54 samples over this time.

3.5 Summary

In summary, this chapter describes two tasks that were undertaken to probe Shanghainese discourse particles and discover the underlying structures that are believed to exist in Shanghainese, and to be responsible for the systematic orderings observed among discourse particles in several languages. The first of the two tasks was a semantically driven fieldwork and elicitation study to find the most common and prominent meanings conveyed by the particles contained in the task. The second of the two tasks, which was partially influenced by the fieldwork exercise, was a survey and acceptability judgment experiment. This experiment tested subjects on their acceptance of several pairs of SFPs, half of which were predicted to be acceptable, while their

mirrored counterparts were predicted to be unacceptable. While this survey does have a narrow scope, and some mechanical limitations, it is believed that this experiment has potential to allow a promising glimpse into the actual behavior and structure of Shanghainese SFP.

CHAPTER 4. RESULTS AND DISCUSSION

4.1 Introduction

This chapter is a presentation of all relevant results obtained from the experiments as described in Chapter 3. Section 2 describes the qualitative data obtained from the fieldwork experiment, in which 5 native speakers of Shanghainese were interviewed about the particles' meanings and usage. Section 3 depicts the output from the survey experiment, which was an acceptability rating task hosted online. The data used below was provided by 36 native speakers of Shanghainese in China and the United States. Information on attrition is provided in the Methodology section in Chapter 3.

The primary data analysis tools used in the processing of this data were Microsoft Excel, Google Sheets, and SAS. The methods used were combinations of simple summary statistics and two and three-way analysis of variance (ANOVA).

4.2 Fieldwork Results

4.2.1 Leq

The participants who participated in the interviews were flexible in their interpretations of *leq*, though there was noticeable overlap in their descriptions. While three of the five participants explicitly described *leq* as affecting the tense of the sentence, the two remaining subjects maintained that it could also be used to mean "finally" or that someone "has begun to do something".

These responses highlight the ambiguity that exists between the perfective marker and the SFP leq. Other parallels may exist due to the similar phonetic nature and history of the particle *leq* and Mandarin particle le.

Tense was not explicitly mentioned in all responses to sentences involving *leq*, as some participants described the semantic difference between a bare sentence and an identical utterance plus *leq* as being unnoticeable, showing that *leq* can in some cases fail to recall the perfective marker, although the more subtle meanings of the particle may be difficult to verbalize. As Qian describes (1997: 214) *leq*, it can often denote that 'narrative process becomes fact', or state a 'current process' or 'confession'.

4.2.2 Aa

Five participants in the interviews described particle *aa* as contributing an interrogative force to at least one utterance containing it. The relevant factor distinguishing cases in which it is interrogative and those in which it is exclamatory may lie in the prosody of the host/particle combination. Two speakers indicated that an utterance being spoken with a rising intonation and *aa* is more likely to be interpreted as a question. Other claimed uses of *aa* include emphatic or surprised uses, and one participant described the *aa* as denoting praise being given to the hearer.

4.2.3 Nao

In descriptions of the particle *nao*, speakers cite a common meaning of offering a warning to the listener. This response was offered by four of the five participants. However, even the singular participant who did not explicitly mention the word "warning" was quick to point out its "I told you so" meaning, which seems to indicate that advice or a warning was given in the past. To accompany this sometimes smug

meaning, it can indicate that an example is being provided at the time of the utterance containing *nao*, as in “Look there”, or “You see?”. Another speaker mentions that the particle may indicate that the user of *nao* feels a tinge of pity toward some individual, who may or may not have failed to heed some past advice.

4.2.4 Laa

Three participants cited an interrogative use of *laa*, which may or may not indicate that the speaker has some unspoken assumptions about the answer they expect. Other subjects indicate that use of *laa* reveals some surprised or excited affect on behalf of the speaker, signaling a kind of “oh no” or “finally”. The latter use may also indicate the same expectations or presuppositions that the speaker has when intending to use the particle in an interrogative register, although speakers seemed to agree that *laa* is not exclusively used to express questions. Indeed, one speaker claimed to never use the particle at all. In some cases, participants seemed to confuse particle *laa* with the phonetically similar *leq*.

4.2.5 Va

The largest degree of overlap among participants in this experiment, and simultaneously the least effort in defining the difference between bare sentences and those containing *va*, was observed in descriptions of this particle. It was unanimously agreed that *va* changes a bare utterance into a Yes-No question, though one speaker offered a statement that they still prefer bare A-not-A questions to those formed by the addition of *va*.

4.2.6 Wa

While interpretations of *va* were easily unanimous, speakers were much more uneasy about describing *wa*, especially because two of the five participants denied making use of the particle *wa* in their idiolect. For the remaining three, descriptions of its use were much more simplistic, such that *wa* was noted to mean that the speaker was 'surprised'. Only one of the five participants indicated something in addition to this finding, wherein they mentioned *wa* as pointing out some object or fact. Participants' depictions of the particle *wa* differ most strikingly from the few descriptions of it in the literature, at least when compared to how authors have defined the impacts of *va*, *lao*, and *nao*. Qian (1997:215) notes that *wa* is used as a rhetorical question, or to denote some expectation.

4.2.7 Ya

Much variety was observed when speakers were asked to describe uses of particle *ya* as well. Two speakers indicated that *ya* could be used when explaining something or providing information that had been mentioned in the past, as in repeated tellings. In a similar fashion to the 'repeated tellings' meaning, a third speaker cited its use as explaining something that 'wasn't the fault of the speaker'. Finally, two speakers recalled uses of *ya* that, like *wa*, expressed a kind of surprise or exclamation.

4.2.8 Lao

Descriptions of *lao* invoked equally sassy interpretations that were attributed to speakers' interpretations of *nao*. In the majority of responses, *lao* denoted a heavy air of "I told you so", and that a certain piece of information was well known by the speaker, but either the hearer or some third party under discussion has failed to consider it, to their

chagrin. Complements of the particle *lao* are often events of pieces of information that the speaker either expects to happen based on inference or past experience. Qian (1997) describes *lao* as being used with things that are ‘well accepted by the public’, or as being used to rebut some contradiction spoken by another person.

4.2.9 Lei

While only used by four of the five participants, the elicited descriptions of *lei* remain relatively consistent, with three of the four responses showing a marked dissatisfaction or disbelief toward the state of affairs, such as “unfortunately”. In a bit of overlap, two speakers also indicated that *lei* is used to point out something or to emphasize the expression.

As a summary of the particles’ comparison to their depictions in the Shanghainese literature is provided in Table 3 below. This table also indicates the features employed in the use of each particle.

Table 3: Comparison of Fieldwork to Qian 1997

SFP	Qian 1997	Fieldwork	Possible Features
leq	narrative process become fact, results of assumption, past state, confession,	past tense', 'finally', 'similar meaning'	Finite
aa	exclamation, ask choices, urge, appraise, persuade	surprise', 'interrogative' (with rising tone), 'emphasis', 'weak assumption', 'praise someone'	Mirative, +Q, Given
nao	warn, remind, rhetorical question, suggest, certainty. objection	"see?", 'told someone before', 'warning', 'give example", 'sympathy'	Adhortative, Given
va		Question', 'no assumptions', 'Like A-not-A'	+Q
laa		Question, "oh, no", 'believe to be true', 'casual'	+Q, Given
lao	well accepted by public, rebut, confirm, state reason	frequent', 'as expected", 'reminder', 'sassy', 'explanation', "Look",	Given, Adhortative
wa	"is or isn't", rhetorical question, infer,expect	surprise'	Mirative
ya	demand, exclamation	repeated tellings', "emotion', 'surprise', 'explanation'	Mirative, Adhortative
lei		not happy', 'emphasis', 'unexpected', 'point out'	Mirative, Adhortative, Given

4.3 Acceptability Judgment Survey Results

4.3.1 Singleton Particles

Stimuli containing singleton particles were included in the survey as a means of measuring whether each of these individual particles were interpreted favorably and to account for the event of one unused or otherwise sensitive particle which could compromise the acceptability of one of the pairs. The chart in Figure 5 presents summary data for the acceptability ratings of all singleton-containing stimuli across the two blocks. All data from the 36 respondents' measurements of all 18 singleton particle sentences were processed by with a three way analysis of variance (ANOVA). Using acceptability as a dependent variable, and age group, block, and ordering as independent variables.

Each of the nine particles were used twice. Recall that acceptability was measured by prompting respondents to use a seven-point Likert scale, on which 7 marked a fully natural utterance and 1 marked a totally unnatural one.

As can be seen in Figure 5, the singleton particles received fairly high ratings across the board, with a sample mean of 5.231 for all sentences using only a single particle. Particles *ya* and *wa* represent the extreme ends of the range of sample means at 6.36 and 3.28 out of 7.

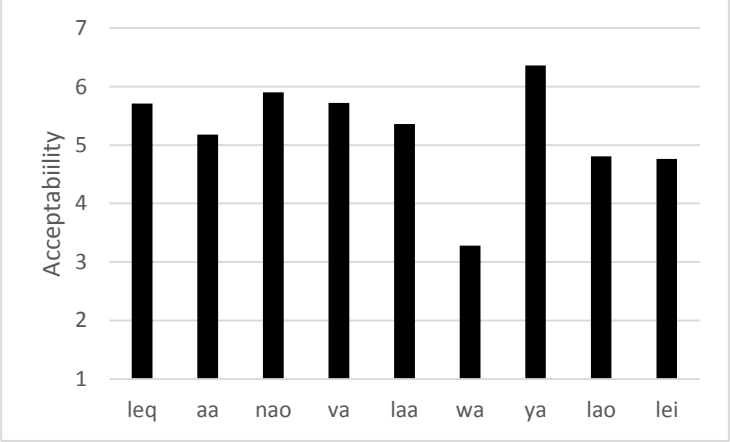

Figure 5: Mean Acceptability of Singleton SFP

A three-way ANOVA was performed using SAS, which treated block as an independent variable with two levels, particle as an independent variable with nine levels, and age group as an independent variable with two levels. Age groups were determined depending on whether a participant was born before or after 1980. Acceptability was left as a dependent variable, and participant was used as a random variable. It was revealed that there were no significant effects of age group when measuring interactions of all singleton particles, yielding an F statistic of 2.05, and a p-value of 1.526. Particle and block were found to have significant effects on subjects' acceptability ratings, with

particle having an F-statistic of 24.03 and p-value of <.0001 and block having an F-statistic of 5.42 and p-value of .0202. A graph representing the interaction between block and rating plot, drawn in Excel, is provided in Figure 6.

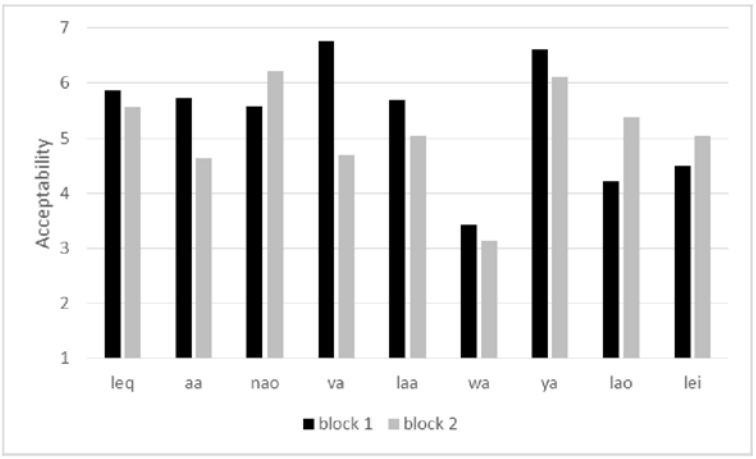

Figure 6: Comparison of Block Effects

Figure 6 demonstrates similar patterns to those observed in Figure 5, with severely diminished ratings for *wa*, and generally high ratings for the remaining seven particles. In addition to this, Figure 6 demonstrates some discrepancies that occurred between blocks. Stimuli in block 2 were lower on average than those in the block 1, which may hint at effects of participant fatigue over the course of the experiment. More inexplicable is the large difference between blocks observed for particle *va* (represented in the x-axis as item 4). This could be an influence of either participant fatigue or the context which was used as a base for all stimuli which occurred within that block.

4.3.2 Particle Permutations

As a general summary of acceptability ratings for the 36 stimuli containing two-particle clusters, Figure 7 represents a comparison of the sample means for all 18 permutations. For each set of columns in Figure 7, the leftmost (black) column represents

the combination that was predicted to be acceptable by the model described in Chapter 2. The rightmost column (grey) in each set is the mirror of the predicted cluster (e.g. for the first pair, *leq aa* is on the left, and *aa leq* is on the right).

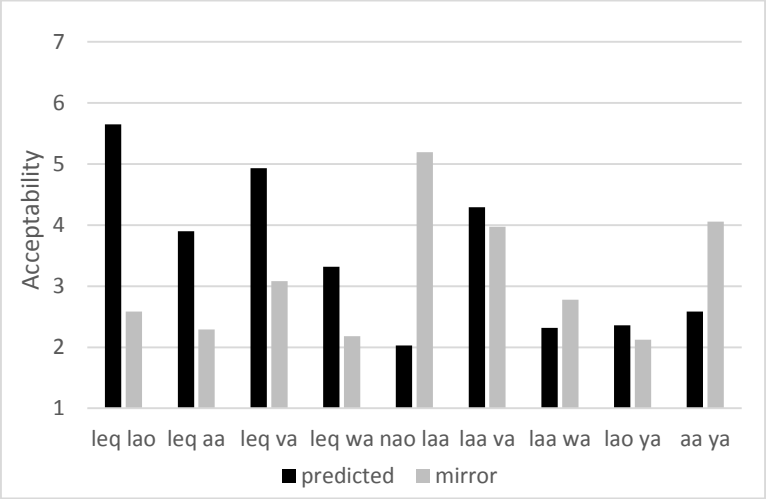

Figure 7: Mean Acceptability of Permutations

Figure 7 shows that, for six of the nine pairs of permutations, the predicted pattern was preferred on average. The predictions for ordering of the pairs *laa/wa, nao/laa, aa/yaa* did not follow the predicted pattern. Similar to findings from the singleton group, all combinations including particle *waa* experienced much lower ratings than the rest of the pairs. Also suffering were the combinations of *ya* and *lao*. A three-way ANOVA tested for the effects of order, block, and age group on acceptability ratings. This test revealed significant effects of order, block, and age group. The effect of order was showed an F-statistic of 10.23, with a p-value < .001, with the predicted order being more acceptable than the mirror order. The effect of age group showed an F-value of 6.77, with a p-value of .009. Older speakers showed higher acceptability ratings than younger

speakers. The block effect showed an F-value of 13.14 and a p-value of < .001. Block 1 showed higher acceptability than block 2.

Immediately visible is the overall decrease in acceptability when more than one particle is in use for a given sentence. The mean of all two-particle samples came out to 3.310, while the mean of all predicted pairs was set at 3.486 with the mirrors at 3.135. These summary statistics were obtained using Microsoft Excel.

More precise comparisons of each combination, including tests for significance of ordering and age group effects are given in subsequent subsections below in Sections 3.2-3.10. The ordering of these sections is moving from the bottom up, as predicted from the predicted ordering. The data used in comparing the mirroring permutations were obtained by running a two-way ANOVA in SAS. The script used for the ANOVA treated order and age group as independent variables and rating as the dependent variable. Just as in the singleton analysis, participant was treated as a random variable. Although as reported above, block 1 was rated as significantly higher than block 2 overall, block was not included as a factor in analyzing the individual combinations of particles. Although the unexpected block effect can be seen as resulting from a limitation of the design, it was decided to only focus on the linguistically important factors (order and age group) for the analysis of individual particle pairs.

4.3.2.1 Leq + Aa

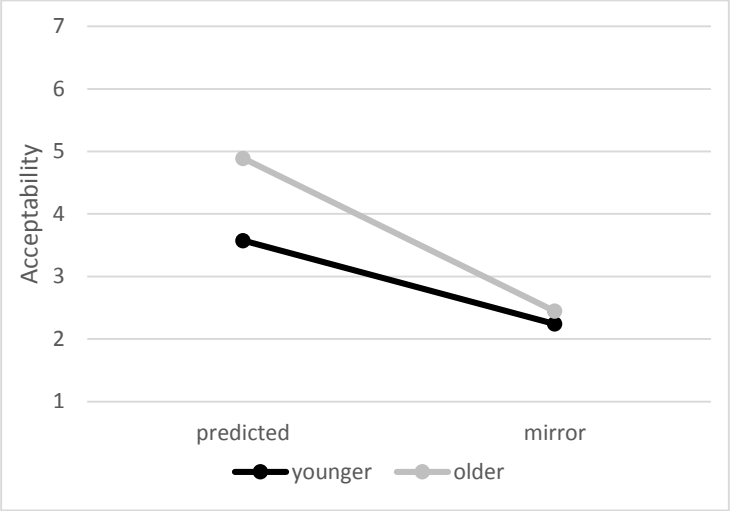

Figure 8: Mean ratings of Leq/Aa Pairs by Age

The sample means of permutations *leq aa* and *aa leq* were 3.871 and 2.271, respectively. A main effect of order was observed for the rating, witnessing an F statistic of 36.92 and a p-value of <.0001. Speakers born before 1980 showed more pronounced acceptance of the predicted ordering with a mean acceptance of *leq* aa of 4.888 compared to the younger group's mean of 3.574. A main effect of age group was ruled to be significant for these items, with an F statistic of 5.97 and a p-value of .0158.

4.3.2.2 Leq + Va

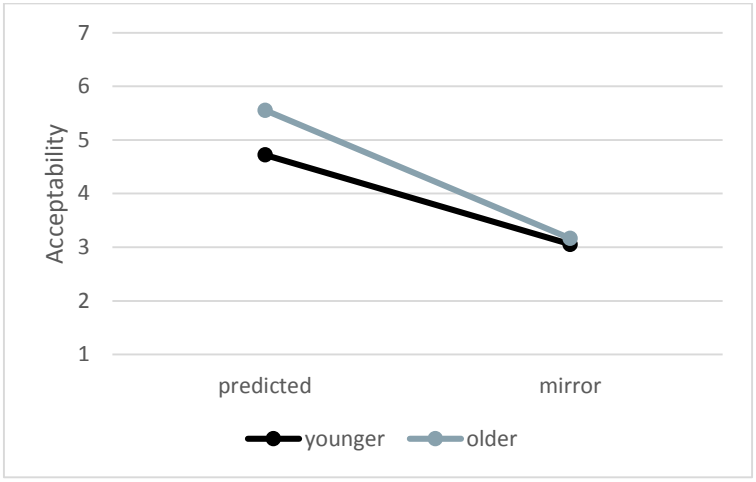

Figure 9: Mean Ratings of Leq/Va Pairs by Age

The sample means of permutations *leq va* and *va leq* were 5.138 and 3.111, respectively. A main effect of order was observed, witnessing an F statistic of 33.03and a p-value of <.0001. Speakers born before 1980 showed more pronounced acceptance of the predicted ordering with a mean acceptance of *leq* va of 5.555 compared to the younger group's mean of 4.722. This effect of age group was ruled to be insignificant for these items, with an F statistic of 1.67 and a p-value of .1830.

4.3.2.3 Leq + Wa

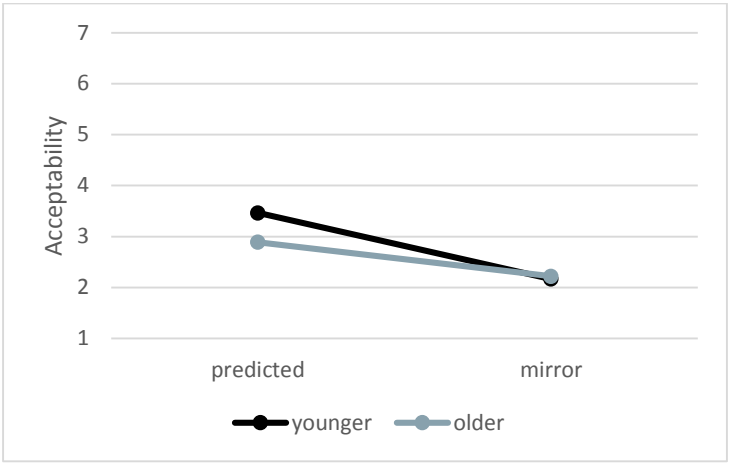

Figure 10: Mean Ratings of Leq/Wa Pairs by Age

The sample means of permutations *leq wa* and *wa leq* were 3.175 and 2.194, respectively. A main effect of order was observed, witnessing an F statistic of 20.70 and a p-value of <.0001. Speakers born before 1980 showed a less confident acceptance of the predicted ordering with a mean acceptance of *leq* wa of 2.888 compared to the younger group's mean of 3,462. A main effect of age was ruled to be insignificant for these differences, with an F statistic of .80 and a p-value of .3713.

4.3.2.4 Leq + Lao

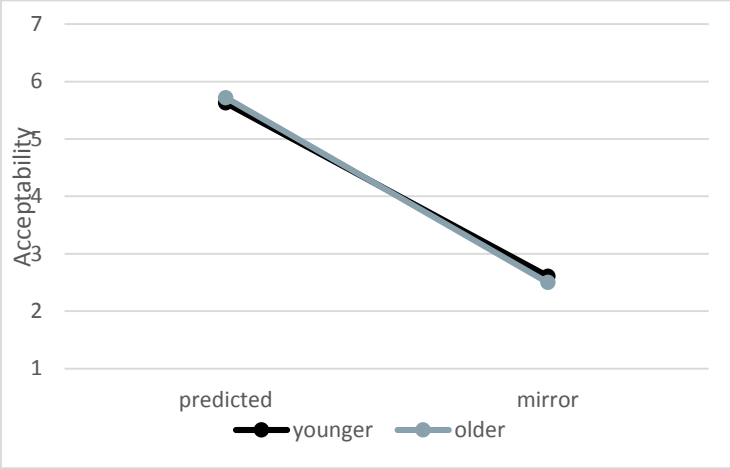

Figure 11: Mean Rating of Leq/Lao Pairs by Age

The sample means of permutations *leq lao* and *lao leq* were 5.675and 2.555, respectively. A main effect of order was observed, witnessing an F statistic of 123.70 and a p-value of <.0001. Speakers born before 1980 showed only slightly more pronounced acceptance of the predicted ordering with a mean acceptance of *leq lao* of 5.722 compared to the younger group's mean of 5.629. A main effect of age group was ruled to be insignificant for these items, with an F statistic of .000 and a p-value of .9737.

4.3.2.5 Lao + Ya

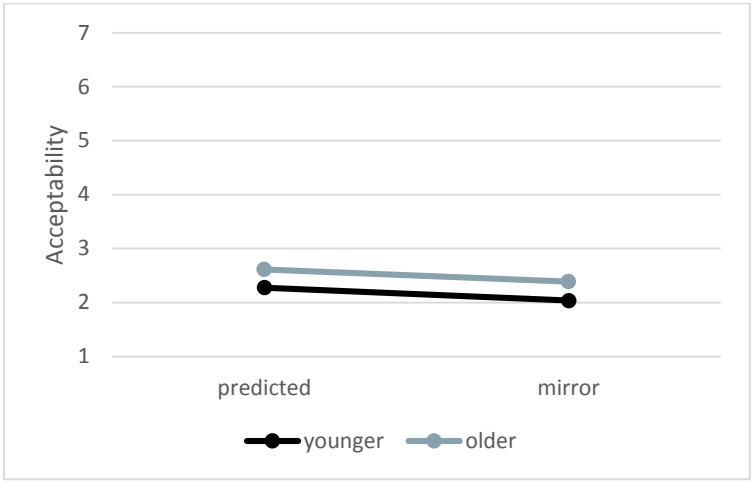

Figure 12: Mean Ratings of Lao/Ya Pairs by Age

The sample means of permutations *lao ya* and *ya lao* were quite low, siting at 2.444 and 2.212, respectively. A main effect of order was observed, though non-significant, witnessing an F statistic of .65 and a p-value of .3437. Speakers born before 1980 showed slightly more pronounced acceptance of the predicted ordering with a mean acceptance of *lao ya* of 2.611 compared to the younger group's mean of 2.277. However, this effect of age group was ruled to be insignificant for these items, with an F statistic of 1.43 and a p-value of .2346.

4.3.2.6 Aa + Ya

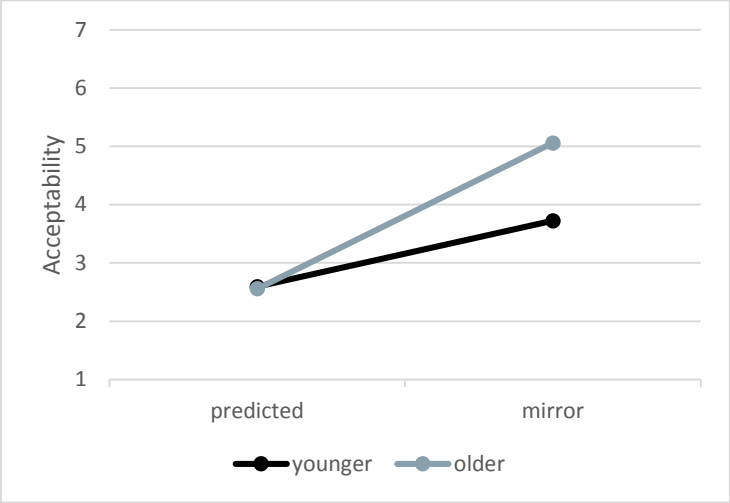

Figure 13: Mean Ratings of Aa/Ya Pairs by Age

The sample means of permutations *aa ya* and *ya aa* came out to be 2.574 and 4388, respectively. This is notable because this patterned in the opposite manner than expected. A significant main effect of order was observed, witnessing an F statistic of 31.66 and a p-value of <.0001. Speakers born before 1980 showed slightly more pronounced acceptance of the unpredicted ordering with a mean acceptance of *yaa aa* of 5.055 compared to the younger group's mean of 3.722. The predicted ordering *aa ya*, though dispreferred, had a mean rating by the older group of 2.555, and the younger group gave this ordering 2.592. Age was shown to be was ruled to be significant for these differences, with an F statistic of 4.04 and a p-value of .0464.

4.3.2.7 Laa + Va

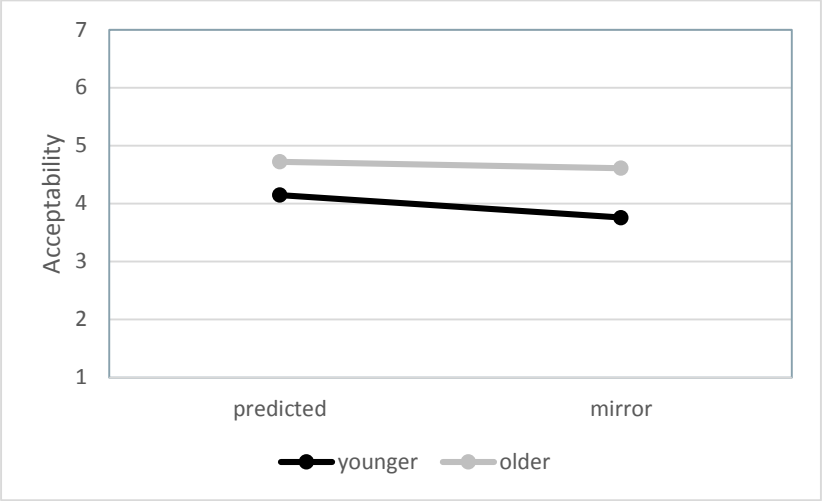

Figure 14: Mean ratings of Laa/Va pairs by Age

The sample means of permutations *laa va* and *va laa* were relatively high, sitting at 4.435 and 4.185, respectively. A main effect of order was observed, though non-significant, witnessing an F statistic of .42 and a p-value of .5173. Speakers born before 1980 showed slightly more pronounced acceptance of the predicted ordering with a mean acceptance of *laa va* of 4.722 compared to the younger group's mean of 4.148. However, this effect of age group was ruled to be insignificant for these items, with an F statistic of 3.43 and a p-value of .0662.

4.3.2.8 Nao + Laa

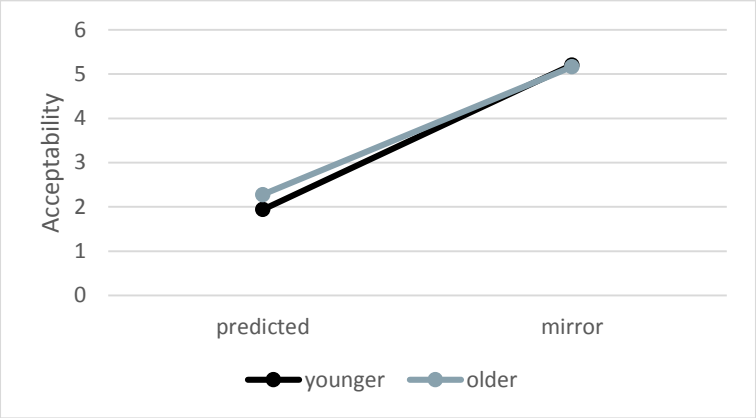

Figure 15: Mean Ratings of Nao/Laa Pairs by Age

The sample means of permutations *laa nao* and *nao laa* came out to be 2.361 and 5.185, respectively. These findings show that the predicted ordering is not favored over its mirror. A significant main effect of order was observed, witnessing an F statistic of 175.67 and a p-value of <.0001. With respect to age group, speakers born before 1980 showed slightly more pronounced acceptance of the unpredicted ordering with a mean acceptance of *laa nao* of 5.166 compared to the younger group's mean of 5.203. In judgments of the predicted ordering, the younger group edged slightly over, though to a non-significant degree, with the younger group granting *laa nao* a rating of 2.277, compared to the older group's 1.944. However, this effect of age group was ruled to be insignificant for these items, with an F statistic of .29 and a p-value of .5921.

4.3.2.9 Laa + Wa

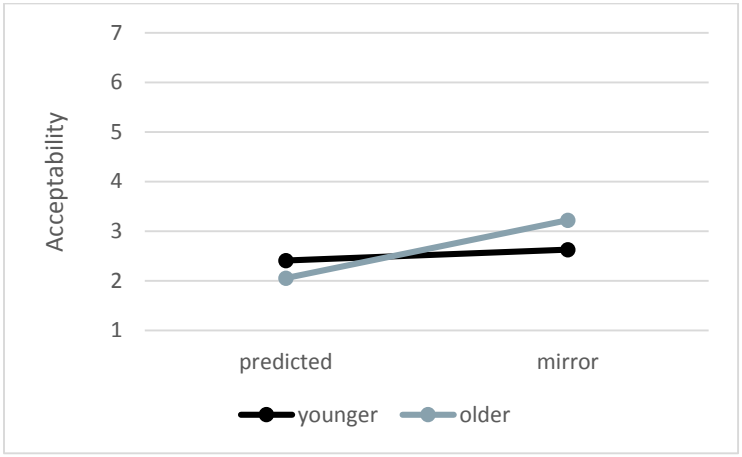

Figure 16: Mean Ratings of Laa/Wa Pairs by Age

The sample means of permutations *laa wa* and *wa laa* came out to be 2.231 and 2.932, respectively. This low mean is believed to suffere from the same degradation in acceptability that affected all *wa* stimuli These findings also show that the predicted ordering is not favored over its mirror. A significant main effect of order was observed, witnessing an F statistic of 3.12 and a p-value of .0207. With respect to age group, speakers born before 1980 showed slightly more pronounced acceptance of the unpredicted ordering with a mean acceptance of *laa wa* of 3.235 compared to the younger group's mean of 2.629. In judgments of the predicted ordering, the younger group edged slightly under, though to a non-significant degree, with the younger group granting *laa wa* a rating of 2.407, compared to the older group's 2.055. However, this effect of age group was ruled to be insignificant for these items, with an F statistic of .18 and a p-value of .6725.

The trend related to age group and rating seems to be that for all ‘acceptable’ examples, or above some threshold for what speakers determined to be acceptable, older speakers gave higher ratings. For other examples, the judgments given by both groups were often very close, if not even.

Table 4: Comparison of Means by Permutation

Predicted	**Mean$_{Pred}$**	**Mirror**	**Mean $_{Mir}$**	**Order F-val**	**Order p-val**	**Age F-val**	**Age p-val**
leq aa	4.231	**aa leq**	2.342	36.92	<.0001	5.97	.0158
leq lao	5.765	**lao leq**	2.555	123.70	<.0001	0.0	.9737
leq va	5.138	**va leq**	3.111	33.03	<.0001	1.79	.1830
leq wa	3.175	**wa leq**	2.194	11.53	.0009	.80	.3713
aa ya	2.574	**ya aa**	4.388	31.66	<.0001	4.04	.0464
nao laa	2.361	**laa nao**	5.185	124.16	<.0001	.29	.5921
laa va	4.435	**va laa**	4.185	.42	.5173	3.43	.0662
laa wa	2.231	**wa laa**	2.932	5.47	.0207	.18	.6725
lao ya	2.444	**ya lao**	2.212	.65	.4213	1.43	.2346

4.4 Discussion of Predicted Permutations

To briefly review the hypothesis presented in Chapter 3, a potential structure of Shanghainese discourse particles was proposed based on other structures which were used to explain patterns of discourse particles observed in other languages, such as Cantonese (Sybesma and Li 2007). In this hypothesis, the particles occupy head positions in a string of functional projections, as depicted in here as Figure 17 and (29) (c.f. Figure 4 and example (22a) in Chapter 2).

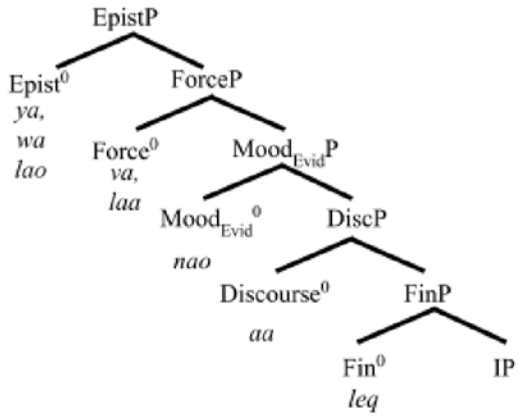

Figure 17: Preliminary Hypothetical Structure

(29) Predicted ordering (First to Last / Low to High)
leq > *aa* >*nao*> *va, laa*> *lao, wa, ya*
(Repeated from (22a) above)

Interpreting Table 4 and Figure 17, this predicts that a particle lower in the tree should always have a higher mean rating when preceding (in surface word order) a particle which occurs in a higher position. The reasoning for this is that a roll-up movement is assumed wherein upon merger of a functional head, the then-complement of the head is copied to the Specifier position of the phrase. This process is assumed to happen iteratively, with the result that particles closer to the end of the sentence are actually the last to be merged in the structure.

Another explanation, with the same predictions, is that particles merge as heads of head-final structures. This alternative removes the necessity of copying large chunks, though at the cost of introducing mixed-headedness within a single category. However, this also assumes that all particles compose a homogenous category which admittedly seems to be a large assumption, though it is not known what computational costs mass-copying may invoke compared to the computational costs of maintaining category specific linearizations. To put this issue aside for now, it will be assumed that the

categories are treated in an equal fashion, and that roll-up movement is the default operation.

Starting at the bottom, it was demonstrated in the survey experiment that all combinations in which *leq* occurred first in a pair (c.f. *leq aa, leq lao, leq va, leq wa*) were preferred over the *leq*-second pairs by the majority of participants to a statistically significant degree. As *leq* was predicted to be the lowest in the hierarchy, this bodes well for the foundations of the hypothesis. While *leq*'s homophony with the perfective aspect morpheme may confound initial glances at these examples, the fact that a sentence may carry both the perfective aspect and the particle demonstrates the existence of a distinct morpheme *leq* which acts as a SFP. One interesting finding in the survey experiment was the mean ratings of mirror pair *va leq*, which showed uncharacteristically high ratings, given that *va* is believed to reside in ForceP. This finding is one that will direct future investigations, though one speculative proposal for this finding could be in the lexical frequency of both particles, which may cause a bias in its judgment. Regarding the syntactic position of *leq*, it may be that the position in which *leq* occurs remains the same as in its origin language. *Leq* is considered to be a loanword from standard Mandarin, so occurring in the head of a low Finite P or DeicticP (as Sybesma and Li 2007 would have) is plausible.

The mirrored pairs *aa ya* and *ya aa* elicited surprising results in the acceptability judgment task, which affords a number of options in attempting to explain the discrepancy between the predictions of the hypothesis and what was revealed empirically. Since the gap in acceptability between the preferred *ya aa* and the dispreferred *aa ya* was so great, there can be little ambiguity in its interpretation. As *ya* was believed to occupy

the head of the highest proposed node Epist^0, with *aa* in the head of Sybesma and Li's (2002) Disc^0 it is at first unclear, given this distance, why the order *ya aa* should be preferred.

The first possibility is that there are many possible morphemes which provide the varied interpretations of the particle *aa*. Since interpretations of the particle were so varied among speakers who were interviewed in the fieldwork task and authors who have published on this topic (interrogative, non-interrogative, emphatic, assumptive, etc.) then there may be a few related morphemes with the same phonetic form. Differences in tone may also be a clue into this issue, as it was mentioned by a speaker in the Fieldwork exercise that changing the tone to a rising one allows the interrogative interpretation of *aa*.

A second possibility for why *ya* precedes *aa* is the dilemma of headedness, which is currently a matter of theoretical debate among researchers of discourse particles, especially SFPs. If headedness were to be the cause for the pair *ya aa*, one might reason that the projection headed by *aa* is head-final. With no other changes, however, the ordering *aa ya* would still be predicted. Following this line of thought, in order for the preferred surface order to remain intact, the amount of content assumed to raise through iterative roll-up operations would need to be restricted, or at least simplified to only move a lower constituent, such as the IP. Since this solution introduces a number of additional complications involving which content to copy (since it can no longer uniformly be assumed to be the complement of each SFP-head) or where to raise it (since this also bears an effect on the linearization), then headedness of the particles seems to not introduce an ameliorative effect on the ultimate word order, still necessitating movement

to maintain it. Thus adjusting linearization of the functional projection containing *aa* is a sub-optimal solution to predicting the ordering *ya aa*.

Yet a third possibility for explaining the preference for the unpredicted ordering over its alternative is movement of the Disc^0 head *aa* moving to a position above Epist^0 *ya*, although this solution seems to invoke a changing in features of the head *aa* which would be needed to motivate such a movement. This seems to hint back at the first solution to the issue, which is to allow for either a single *aa* with many 'denotations', or to postulate multiple homophonous lexical items of the form *aa*. Maintaining the earlier assumption of "one-particle one head", the most reasonable solution is to postulate a position above Epist^0.

In attempting to maintain the assumptions of linear transitivity, the invention of a novel projection may be the most reasonable analysis available for *aa*'s dominance over *ya*, though the diversity of meanings afforded by *aa* make it difficult to pinpoint the specific feature which fills this head above EpistP. In the interest of syntactic simplicity and nominal elegance, and perhaps at the cost of lexical or semantic elegance, the proposal involving parametric variation in the ordering of merger of certain features, and the clustering of these features into lexical items will used as a bookmark solution to the dilemma and the issue of multiple *aa* morphemes will be left to further investigation. For now, it will be proposed that the features contained by *aa* are merged into a projection, *aa*P above EpistP. Though there still remains a question about the final position of *ya*, which tested poorly in judgments of the *lao* + ya pairs

Moving on to the particle *nao* which was predicted to rest in the head of a MoodP, and thus occur before laa, its pair member in the experiment. The survey, however, revealed another rather unexpected preference, that the ordering *laa nao* was significantly preferred over *nao laa*. This is surprising because *laa* was believed to head ForceP, which dominates MoodP in the sketch provided in the preliminary hypothesis.

Part of this preference may be explained by a slight tendency to interpret *laa* as *leq* in some structures, which was documented in the fieldwork interviews. While written feedback from participants, such as a transcription, may have helped reveal the prevalence of this phenomenon, this was not implemented into the methodology of the survey experiment for a number of reasons. The lack of a standard Shanghainese orthography notwithstanding, this would have dramatically increased the time needed to complete the task, and it is speculated that this may have resulted in much greater levels of attrition.

Relying on transitivity, namely the fact that *laa* tested above both *wa* and *nao*, as evidenced by the not-quite-significantly higher mean ratings of permutations *wa laa* and *nao laa* over their mirrors, as well as demonstrations of *laa*'s higher ratings in positions before *va*, speaker judgments seem to prefer *laa* in a position between *wa* in MoodP and *va in* ForceP. Despite what was suggested in the initial hypothesis, both *laa* and *va* cannot occupy the Force head simultaneously. This makes it reasonable to put *laa* in a position directly under ForceP, known as $\text{Mood}_{\text{Informative}}$P, as taken from Sybesma and Li (2007)

Another issue is the particle *nao*, which had been predicted to reside in MoodP, although subjects preferred it in a position above *laa*, which is now believed to reside in

$Mood_{Informative.}$ Aside from mishearing the particle, a certain feature may be the cause of this unexpected pattern. When comparisons were originally drawn, a focus was on the adhortative, reproachful meanings which were evoked upon using this particle, such as those used in Cantonese particles *wo3* and *gaa3*, which all contain minimal meaningful units in the Mood field (Sybesma and Li 2007). It may be the case that Shanghainese speakers prioritize another feature of this particle, which is its focus on prior knowledge and given information. If this epistemic feature Given is more prominent in the interpretation of *nao* than other features, then *nao* may rest in the head of EpistP, above both ForceP and $Mood_{Informative}$. No testing has yet been done to look at the ordering relations between *va* and *nao*.

Related to the discussion above, *laa va* failed to obtain significant effects of ordering when compared to its mirror *va laa*. One potential reason for this is that these particles were hypothesized to occupy the same position, $Force^{0}$. While forcing speakers to choose between two structures in which a singular head position is filled by multiple objects would seemingly predict an ungrammatical result for both orders, results obtained from respondents for all items using both *laa* and *va* were higher than the sample mean of 3.31 for all two-particle stimuli, sitting at 4.351 for *laa va* and 4.185 for *va laa*. This may indicate that there is a finer distinction to make than the simply putting all force-denoting features into the ForceP. There is a possibility, allowed for by the general acceptability of both structures that *laa* may base-generate in a lower position and move up to take on its force denoting characteristic. As discussed above, it may reside in one of the higher Mood projections used by Sybesma and Li (2007), below ForceP. Another reason for this

blurring, as mentioned above, could be that some subset of speakers are misinterpreting the *laa* as *leq*, which has been shown to occur very low in the structure.

The predictions involving permutations *laa wa* and *wa laa* were another group which failed to live up to speakers' expectations. The predicted ordering of *laa wa* was significantly dispreferred to the mirror, though ratings for both pairs were among the lowest tested, receiving scores of 2.231 and 2.93 out of 7. Part of this effect could be the low prevalence of the particle *wa* among certain groups of speakers. While no subject effects other than age were measured against ratings of singleton *wa* to test whether this is an effect of age or district, it can certainly be said that *wa* was the least acceptable among the 9 singletons tested. As far as positions for *wa*, the ratings indicate that it must occur in a position under *laa*, as a head to $\text{Mood}_{\text{InformativeP}}$, which may indicate that the dominant feature of *wa*, rather than Qian's (1997) described adhortativity, its exclamatory or mirative expression of surprise. While Sybesma and Li (2007) place their mirative-denoting minimal meaningful unit in EpistP, and this was followed in the formulation of the original hypothesis, it can be observed here that there is reason to be uneasy about placing miratives as high as Epist^{0}. Since *wa* was indicated by subjects in the fieldwork experiment to be fairly unilateral in its interpretation, only denoting 'surprise' to the exclusion of other effects, perhaps a more suited location for *wa* is in the lower $\text{Mood}_{\text{mirative}}\text{P}$, below ForceP and $\text{Mood}_{\text{InformativeP}}$.

Finally, looking at the interaction of *lao* and *ya*, no significant conclusions were reached as far as main effects of order. *Lao ya* was preferred by a non-significant margin over *ya lao*, but this pair shares the same dilemma that is observed in the pair *laa* and *vaa* in which these two particles were expected to occupy the same head Epist^{0}. In generating

the hypothetical structure seen in Figure 17, the finer distinctions between similar projections were omitted to minimize the number of proposed functional heads before an analysis had been undertaken, but it seems that finer distinctions must be made. In this case, Sybesma and Li (2007) made use of two separate Epistemic heads in order to accommodate the number of minimal meaningful units that they discovered. While combinations of *lao* and *ya* received fairly low marks across the board compared to more pairs separated by more structural distance, allowing the two particles to remain in the same head would be a violation of the "one particle one head" assumption, and thus, we can entertain the idea of two Epist^0 heads as a temporary solution in light of a more detailed comparison of the two particles. There may yet be semantic distinctions and pragmatic factors that remain untouched by this brief foray into the world of Shanghainese discourse particles. The ultimate position of *ya* in the structure is still an open question, due to the poor reception of pairs containing it, and the inability to test all possible permutations in this study. Due to the slight amelioration caused by putting it above *lao,* it is a reasonable proposal to keep *ya* in an EpistP projection until it can be better tested. A full diagram of the amended structure of Shanghainese is available below in Figure 18

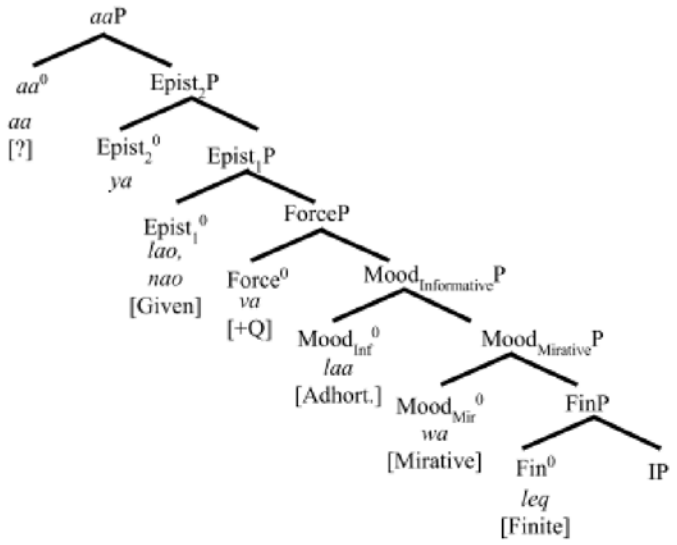

Figure 18: Revised Structure of Shanghainese Left Periphery

Figure 18 displays the revised version of the preliminary structure of Shanghainese that was offered in Figure 17. Based on the acceptability judgment data, several changes have been made to accommodate the patterns observed. Firstly, one can see that the novel projection aaP now dominates EpistP. This helps to show that *aa* dominates all other particles that it was paired with (here *leq* and *ya).* Although Epistemic information has canonically been observed to occupy the highest positions in trees, there is the possibility of parametric variation with respect to the ordering in which functional projections are merged.

Additionally, EpistP has been split into two projections *Epist*$_1$ and *Epist*$_2$, which was motivated by the fact that particles *lao* and *ya* were shown to co-occur, though unfavorably, in contradiction to the prediction that they occupied the same position. In order to maintain the "one feature one head" assumption, then two particles must occupy distinct projections while occurring in a grammatical utterance, barring the possibility of complex heads.

Another adaptation of the original hypothetical structure is the shift of *laa* to a Mood projection. Similarly to the reasoning behind splitting *lao* and *ya*, *laa* and *va* should not occupy the same head. Additionally, since it is unlikely for a sentence to possess two interrogative operators in a single utterance, and since *laa* was observed to not represent interrogative interpretations exclusively, then it may be that *laa* is a manifestation of either the Adhortative or Given feature. After failing to find any epistemic or presumptive meanings in the use of the particle *wa*, and considering the preference of ordering *wa laa* over *laa wa*, it was decided to place *wa* in the head of a distinct Mood phrase,possessing the Mirative feature. This seems appropriate due to speakers' very specific interpretations of *wa*, and their lack of variance in its use. If *wa* is only a phonological manifestation of the Mirative feature, to the exclusion of others, then this position below $Mood_{Inf}P$ and above FinP is most reasonable.

4.5 Summary

This chapter began as a presentation of all relevant results obtained from two small experiments described in Chapter 3; a series of field interviews with Shanghainese students on a college campus, and an online acceptability judgment survey. Section 2 detailed some of the prominent semantic distinctions that native speakers attribute to these SFPs. Section 3 described all of the results from the acceptability survey experiment, which was taken online by 36 native speakers of Shanghainese in China and the United States. Section 4 explained many of the results encountered in Section 3 by referring back to the original hypothetical structure and proposing amendments to some of the assumptions that were made. Section 5 laid out a few directions for future work which include expanding this model to accommodate more particles, speaker variation,

and to provide more empirical bases for the existence of these functional heads in Shanghainese's Left Periphery.

CHAPTER 5. CONCLUSION

In conclusion, it was deemed necessary to begin an initiative investigating Shanghainese as an understudied language. This may extend interest in the language from within the community, and enforce the significant role that the language plays in connecting an old, but rapidly changing city and its people. This study also set out to test current syntactic analyses against data from traditionally understudied languages and ensure that the models withstand empirical analysis.

In an effort to limit the scope of the project to something manageable, only one phenomenon of Shanghainese was probed; its inventory of SFPs. Looking into interactions at this level of the syntax should provide much needed information about the interaction between an internally active syntax, and a mostly externally regulated discourse system. Discourse features which are encoded in the syntax are believed to have a naturally enforced hierarchy or ordering, though the universality of these subtle preferences is currently difficult to test for.

Therefore, languages with much richer discourse morphology serve as an ample testing ground for finding the positions of these discourse features. Much like has been done for Cantonese and Mandarin Chinese, it was believed that the relative orderings of certain SFPs in Shanghainese would either replicate current findings, or that they would reveal new preferential orderings. In order to test these hypotheses, two tasks were conducted

The first task, the field interviews, compared current usage of the particles to their listed descriptions in the literature. This ensured that all relevant features could be accounted for when integrating these articles into the preliminary model. It was revealed that, for the majority of particles, their attested meanings were well represented among the majority of speakers, though the sample size for the interview portion was quite small, at five college-aged participants. With this information, it was a more reasonable exercise in extrapolation, assigning these particles to their corresponding theoretical heads, using studies of Cantonese and Mandarin as a guide.

Secondly, an acceptability judgment experiment was designed and hosted on Qualtrics, an online survey software, which enabled data to be collected from participants living in the densest native speaking population of Shanghainese - Shanghai. This experiment selected a subset of eight SFPs and used their hypothetical syntactic positions to determine what orders were likely to be preferred by native speakers and what orders were likely to be dispreferred. From these 8 particles, a subset of 18 two-particle combinations was composed (9 expected and 9 mirrors) to coordinate their positions in a hierarchical structure. In this experiment, 54 aural stimuli were presented to 36 native speaking participants with a 7-point acceptability rating scale. These stimuli were also

divided into two experimental blocks. The independent variables were the ordering of the particles, such that there were two levels. The predicted order represented one level of this condition, while the mirrored ordering represented the second level. Sentences as with singleton particles were also provided as a control for the speakers' general acceptance of these particles. and to ensure that it was not a single particle which resulted in a lower overall acceptability for a pair, as was demonstrated by pairs containing the particle *wa*.

Results from this experiment demonstrated that four of the nine predicted orderings (*leq waa, leq va, leq lao, leq aa*) were preferred to their mirrors to a statistically significant degree. Three of the predicted orderings (*laa wa, laa va, lao ya*) showed a preference, but failed to exhibit statistical significance. The two remaining orderings *(*aa ya* and **nao laa*) were shown to be dispreferred with respect to their mirrored counterparts, signaling a need to adjust the model. It is believed that these results positively signify the relevance of the Split-CP model of structure at the level of the discourse syntax interface, and that the discourse features responsible for the ordering behaviors observed by the Cartography movement can be observed in Shanghainese as well.

While this study yielded positive results, there are a number of limitations that cannot be ignored, and provide some directions for future research. Firstly, this experiment failed to account for the effects of experimental block and patient fatigue over the course of the experiment, potentially compromising the reliability of the judgments provided in the second block. Secondly, no participant variables were compared against the results to account for sociological effects, such as age or district. This may be

remedied in the future, because it has been noted by some speakers that dialectical differences exist among certain districts, though their effect on the use of SFPs are currently unknown. A third major limitation of the study is that it was not feasible to test all possible permutations of these 8 particles, due to the limited time participants would be willing to sit in for a survey. As mentioned above, this was curbed by selecting a small subset of the permutations that were deemed most likely to either confirm or deny the major hypothesis, though at the expense of total empirical thoroughness. Future investigations will include additional permutations and ensure that all are measured in a controlled environment.

Aside from adjusting for current limitations, other future interests in this topic worthy of investigation include a more thorough semantic and pragmatic analysis of these particles. While this will likely require a corpus of modern Shanghainese, this can also be realized by running a larger scale fieldwork project to catalog more uses of these particles, and potentially discover new particles that do not yet possess descriptions in written literature. Another future project that the author is interested in is testing the final model against more speakers to further validate the existence of an order-determining structure, and investigating the possibility of parametric variation which could account for differences observed among the Sinitic languages.

BIBLIOGRAPHY

BIBLIOGRAPHY

Bayer, J., & Obenauer, H. G. (2011). *Discourse particles, clause structure, and question types*. The linguistic review, 28(4), 449-491.

Cardinaletti, A. (2007). *Für eine syntaktische Analyse von Modalpartikeln.Gesprochene Sprache und Partikeln*, 89-101.

Chafe, W., & Nicholas, J. (1986). *Evidentiality: The Linguistic Coding of Epistemology*. Norwood. NJ: Ablex.

Chen, Z. (2007). *Studies on dialects in the Shanghai area: Their phonological systems and historical developments* (Doctoral dissertation, UMI Ann Arbor).

China Daily. *Shanghai dialect locked in tug of war with Mandarin*. http://usa.chinadaily.com.cn/epaper/2014-02/28/content_17313346.htm. Accessed 04-23-15

Chomsky, N. (2000). *Minimalist inquiries: The framework. Step by step: Essays on minimalist syntax in honor of Howard Lasnik*, ed. by Roger Martin, David Michaels, and Juan Uriagereka, 89–155.

Chomsky, N. (2001). *Derivation by phase. Ken Hale: A life in language*, ed. by Michael Kenstowicz, 1-52.

Chomsky, N. (2008). *On phases*. Current Studies in Linguistics Series, 45, 133.

Chou, C. T. (2013). *Phi-Agree, A-movement, and Complementizer-Tense Relations in Chinese* (Doctoral dissertation, Eastern Michigan University).

Cinque, G. (1998). *Adverbs and functional heads: A cross linguistic perspective*. Oxford University Press, New York USA

Cinque, G., & Rizzi, L. (2008). *The cartography of syntactic structures.* Studies in linguistics, 2, 42-58

.

Cinque, G., & Rizzi, L. (2010).*Mapping Spatial PPs: The Cartography of Syntactic Structures*,(Vol. 6). Oxford University Press.

Cinque, G. (2004). *Issues in adverbial syntax*. Lingua, 114(6), 683710.

Fang, X. (2003). *Guangzhou fangyan jumo yuqi zhuci* [Sentence-final particles in Guangzhou Cantonese]. Jinan University Press, Guangzhou

Fung, R. S. Y. (2000*). Final particles in standard Cantonese: semantic extension and pragmatic inference* (Doctoral dissertation, Ohio State University).

Gilliland, J. (2006). *Language attitudes and ideologies in Shanghai, China* (Doctoral dissertation, The Ohio State University).

Haegeman, L., & Hill, V. (2013). *The syntactization of discourse*. Syntax and its limits, 370390.

Kwok, H. (1984). *Sentence particles in Cantonese* (Vol. 56). Centre of Asian Studies, University of Hong Kong.

Law, A. (2002*). Cantonese sentence final particles and the CP domain*. UCL Working Papers in Linguistics, 14, 375398.

Leung, W. M. (2011). *A Study of Evidential Particles in Cantonese: the case of wo3 & wo5*. The Buckingham Journal of Language and Linguistics, 4, 29-52.

Lewis, D. M. (2006). *Discourse markers in English: a discourse-pragmatic view*. Approaches to discourse particles, 43-60.

Liu, G. (2012). *Social Identity and Sound Change: The Case of Wo in Shanghainese*. Australia Review of Applied Linguistics, 35 (2), 203-214.

Luke, K. K. (1990). *Utterance particles in Cantonese conversation* (Vol. 9). John Benjamins Publishing.

Matthews, S., & Yip, V. (2011). *Cantonese: A comprehensive grammar*. Routledge.

Nationalencyklopedin (in Swedish). 2010. Retrieved 23 February 2015. http://www.ne.se/spr%C3%A5k/v%C3%A4rldens-100-st%C3%B6rsta-spr%C3%A5k-2010. Accessed 04-23-15

Paul, W. (2010). *Why particles are not particular: sentence-final particles in Chinese as heads of a split CP*. Studia Linguistica.

Poletto, C. (2000). *The higher functional field*. Oxford University Press.

Pollock, J. Y. (1989). *Verb movement, Universal Grammar, and the structure of IP*. Linguistic inquiry, 365-424.

Qian, N. (1997). *Shanghai hua yu fa* (上海話語法) [Grammar of ShanghaiWu]. Shanghai: Shanghai ren min chu ban she.

Reuters. (2012) http://www.reuters.com/article/2012/09/04/us-china-dialect-shanghai-idUSBRE88302920120904. Accessed 04-23-15

Rizzi, L. (1997). *The fine structure of the left periphery*. In Elements of grammar (pp. 281337). Springer Netherlands.

Rizzi, L. (2013). *Notes on cartography and further explanation*. International Journal of Latin and Romance Linguistics, 25(1), 197-226.

Shanghai Daily. http://www.shanghaidaily.com/Metro/entertainment-and-culture/Fears-for-Shanghai-dialect-as-youngsters-rank-below-average/shdaily.shtml . Accessed 04-23-15

Speas, P. (2008). *On the syntax and semantics of evidentials*. Language and Linguistics Compass, 2(5), 940-965.

Sybesma, R., & Li, B. (2007). *The dissection and structural mapping of Cantonese sentence final particles*. Lingua, 117(10), 17391783.

VanNess Simmons, R., (2011) *Shanghainese-English/English-Shanghainese Dictionary and Phrasebook*. New York: Hippocrene Books.

Wang, H. Cheng, Y. (2014) *Shanghai kindergartens to promote local dialect*. China daily. Accessed online 12132014: http://www.chinadaily.com.cn/china/2014-02/07/content_17269669.htm

Zhu, X. (2006). *A grammar of Shanghai Wu* (Vol. 66). Lincom Europa.

APPENDICES

Appendix A: Stimuli Sentences*, **

Q #	Audio# - Block 1	Transcription of file	Particles	Translation
Q11	B1-01	Gong ma leq yaq zi seu-biao va laa	va laa	Gong bought a watch
Q12	B1-02	Lei teu leq yiq diao ang'li ya lao	ya lao	Lei stole a necklace
Q13	B1-03	Yue ma leq yiq ngae hao koe eq weijin leq va	leq va	Yue bought a good-looking scarf
Q14	B1-04	yang loq-teq leq yi eq b'i -ga-dzi leq aa	leq aa	Yang lost his wallet
Q15	B1-05	fang san-h'ueq siidou leq wa	leq wa	Fang came to work late
Q16	B1-06	XiaoLi yong leq yiq vae kuae aa ya	aa ya	XiaoLi spent 10,000 RMB
Q17	B1-07	Wuxi be leq ngu yiq b'in shiangsi laa nao	laa nao	WuXi gave me a bottle of perfume
Q18	B1-08	di-ti wa teq leq laa wa	laa wa	The elevator broke
Q19	B1-09	Shiao teu dao teq leq lao	leq lao	The thief escaped
Q20	B1-10	Gong ma leq yaq zi seu-biao aa leq	aa leq	Gong bought a watch
Q21	B1-11	Lei teu leq yiq-diao ang'li wa laa	wa laa	Lei stole a necklace
Q22	B1-12	Yue ma leq yiq ngae hao koe eq weijin nao laa	nao laa	Yue bought a good-looking scarf
Q23	B1-13	yang loq-teq leq yi eq b'i -ga-dzi lao ya	lao ya	Yang lost his wallet
Q24	B1-14	fang san-h'ueq siidou leq ya aa	ya aa	Fang came to work late
Q25	B1-15	XiaoLi yong leq yiq vae kuae laa va	laa va	XiaoLi spent 10,000 RMB
Q26	B1-16	Wuxi be leq ngu yiq b'in shiangsi lao leq	lao leq	WuXi gave me a bottle of perfume
Q27	B1-17	di-ti wa teq leq va leq	va leq	The elevator broke
Q28	B1-18	Shiao teu dao teq leq wa leq	wa leq	The thief escaped
Q29	B1-19	Gong ma leq yaq zi seu-biao wa	wa	Gong bought a watch
Q30	B1-20	Lei teu leq yiq-diao ang'li nao	nao	Lei stole a necklace
Q31	B1-21	Yue ma leq yiq ngae hao koe eq weijin lao	lao	Yue bought a good-looking scarf
Q32	B1-22	yang loq-teq leq yi eq b'i -ga-dzi aa	aa	Yang lost his wallet
Q33	B1-23	fang san-h'ueq siidou leq ya	ya	Fang came to work late
Q34	B1-24	XiaoLi yong leq yiq vae kuae leq	leq	XiaoLi spent 10,000 RMB
Q35	B1-25	Wuxi be leq ngu yiq b'in shiangsi laa	laa	WuXi gave me a bottle of perfume
Q36	B1-26	di-ti wa teq leq lei	lei	The elevator broke
Q37	B1-27	Shiao teu dao teq leq va	va	The thief escaped

Q #	Audio# - Block 1	Transcription of file	Particles	Translation
Q38	B2-01	Jing sanleq yaq ze lubiao leq aa	leq aa	Jing bought a watch
Q39	B2-02	Ming zang tsuo leq gong-jia tsuo leq lao	leq lao	Ming boarded the wrong bus
Q40	B2-03	Jin-tsaq nao Yang zuo chi lae leq laa wa	laa wa	The police arrested Yang
Q41	B2-04	Sao nao yi eq jia-d'aq-tsuo sang wa te leq aa ya	aa ya	Xiao crashed her bike
Q42	B2-05	Xia mangji te yi eq ngatou leq va laa	va laa	Xia forgot his coat
Q43	B2-06	Qiang lagei mulu gaodeu siijiang laa nao	laa nao	Qiang is painting on the street
Q44	B2-07	Yu dae leq siidang li leq ya lao	ya lao	Yue stepped in a puddle
Q45	B2-08	Lei nao leq jiao gwei tsaopiao leq va	leq va	Lei withdrew a lot of money
Q46	B2-09	Wei nao yi eq kafi dang fei te leq wa	leq wa	Wei spilled his coffee
Q47	B2-10	Jing sanleq yaq ze lubiao lao leq	lao leq	Jing bought a watch
Q48	B2-11	Ming zang tsuo leq gong-jia tsuo va leq	va leq	Ming boarded the wrong bus
Q49	B2-12	Jin-tsaq nao Yang zuo chi lae leq lao ya	lao ya	The police arrested Yang
Q50	B2-13	Sao nao yi eq jia-d'aq-tsuo sang wa te leq laa va	laa va	Xiao crashed her bike
Q51	B2-14	Xia mangji te yi eq ngatou leq nao laa	nao laa	Xia forgot his coat
Q52	B2-15	Qiang lagei mulu gaodeu siijiang wa leq	wa leq	Qiang is painting on the street
Q53	B2-16	Yu dae leq siidang li leq ya aa	ya aa	Yuan stepped in a puddle
Q54	B2-17	Lei nao leq jiao gwei tsaopiao wa laa	wa laa	Lei withdrew a lot of money
Q55	B2-18	Wei nao yi eq kafi dang fei te aa leq	aa leq	Wei spilled his coffee
Q56	B2-19	Jing sanleq yaq ze lubiao va	va	Jing bought a watch
Q57	B2-20	Ming zang tsuo leq gong-jia tsuo lao	lao	Ming boarded the wrong bus
Q58	B2-21	Jin-tsaq nao Yang zuo chi lae leq laa	laa	The police arrested Yang
Q59	B2-22	Sao nao yi eq jia-d'aq-tsuo sang wa te leq	lei	Xiao crashed her bike
Q60	B2-23	Xia mangji te yi eq ngatou leq aa	aa	Xia forgot his coat
Q61	B2-24	Qiang lagei mulu gaodeu siijiang	leq	Qiang is painting on the street
Q62	B2-25	Yu dae leq siidang li leq nao	nao	Yue stepped in a puddle
Q63	B2-26	Lei nao leq jiao gwei tsaopiao ya	ya	Lei withdrew a lot of money
Q64	B2-27	Wei nao yi eq kafi dang fei te leq wa	wa	Wei spilled his coffee

* As a note about the table in Appendix A, The orthography used in depicting Shanghainese is an adapted PinYin script, which also uses some elements of the writing system used in Richard VanNess' Shanghainese-English Dictionary (2011). Adaptations resulted from variations in the speech of younger participants.

** This chart is also sorted by Block. The second column in each chart denotes the block in which each item occurred.

Appendix B: Demographic Questions*

Linguistic Background Questionnaire
Shanghainese Research Study

Age (in years) _____

Gender M F

Speak Shanghainese in home Y N

Other languages spoken _____

Time (in years) Lived in Shanghai ☐ (0-5) ☐ (6-10) ☐ (11-15)
☐ (16-20) ☐ (21-25) ☐ (26-30)
☐ (31-35) ☐ (36-40) ☐ (40+)

How often do you use Shanghainese?
☐ Every day
☐ Often (A few times a week)
☐ Occasionally (A few times a month)
☐ Seldom (A few times a year)

If you live or have lived in Shanghai, what district did you live in?

☐ Yangpu	☐ Hongkou	☐ Zhabei	☐ Putuo	☐ Changning	☐ Chongming
☐ Xuhui	☐ Jingan	☐ HuangPu	☐ Baoshan	☐ Luwan	☐ Minhang
☐ Jiading	☐ Jinshan	☐ Songjiang	☐ Nanhui	☐ Qingpu	☐ Fengxian

What is an estimate of your comfort in using Shanghainese? (1 is least: 7 is most)

1 2 3 4 5 6 7

Further comments?

_____
_____
_____
_____
_____

*The questions in this survey were used in both the fieldwork and online acceptability judgment exercises.