The two sides of Wh-indeterminates in Mandarin : a prosodic and processing account

Yang, Y.

THE TWO SIDES OF *WH*-INDETERMINATES IN MANDARIN: A PROSODIC AND PROCESSING ACCOUNT

Published by
LOT
Trans 10
3512 JK Utrecht
The Netherlands

phone: +31 30 253 6111
e-mail: lot@uu.nl
http://www.lotschool.nl

Cover illustration: Two winking owls. This is a Chinese ink painting made and photographed by Linjun Zhang, for this dissertation.

ISBN: 978-94-6093-288-5
NUR 616

THE TWO SIDES OF *WH*-INDETERMINATES IN MANDARIN: A PROSODIC AND PROCESSING ACCOUNT

PROEFSCHRIFT

ter verkrijging van
de graad van Doctor aan de Universiteit Leiden,
op gezag van Rector Magnificus prof. mr. C.J.J.M. Stolker,
volgens besluit van het College voor Promoties
te verdedigen op woensdag 30 mei 2018
klokke 10:00 uur

door

YANG YANG

geboren te Shandong, China
in 1987

Promotores:	Prof. dr. Lisa L.S. Cheng
	Prof. dr. Niels O. Schiller
Co-promotores:	Dr. Leticia Pablos
	Dr. Stella Gryllia
Promotiecommissie:	Prof. dr. Shengli Feng (The Chinese University of Hong Kong)
	Prof. dr. Aoju Chen (Utrecht University)
	Prof. dr. Jenny Doetjes
	Dr. Yiya Chen

Contents

Acknowledgements

Working as a PhD and writing a dissertation are challenging. I wouldn't have come so far without a great deal of help and support. My greatest gratitude goes to my promotors, Lisa Cheng and Niels Schiller, two of the most intelligent people in the world I have known. Back in 2013, your prompt reply in taking me as your PhD student made all this happen. And the scope of your knowledge and critical thinking not only advance my current work but also set a model for how to be a researcher.

My PhD life has changed remarkably since I joined the NWO project Understanding Questions and started to work with Lisa Cheng, Jenny Doetjes, Leticia Pablos, Stella Gryllia and Aliza Glasbergen-Plas. I sincerely thank Lisa for admitting me into the project, providing me with constant support and always thinking ahead for me. Jenny, you are always so patient and considerate in helping me and I am very grateful for that. Leticia, I am so happy to have you as not only a daily supervisor who helps me enormously but also a close friend; the time we spent together in Spain and China (with your husband Jose) is very precious for me. Stella, my other daily supervisor, thank you so much for spending so much time in working out the details and because of your help, I like working on prosody now. Aliza, as the only two PhDs in the project, we have done so many things together and witnessed all of each others' improvements and it is so good to have you marching together in this bumpy journey.

I would like to thank my fellow colleagues, who have been very supportive and helpful, making my work and life in LUCL so enjoyable: Rint Sybesma, Jos Pacilly, Yiya Chen, Arie Verhagen, Anikó Lipták, Hang Cheng, Menghui Shi, Saskia Lensink, Han Hu, Qiong Peng, Kate Bellamy, Güliz Günes, Jing Lin, Min Liu, Yifei Bi, Amanda Delgado, Elena Karvovskaya, Daan van de Velde, Andreea Geambasu, Bobby Ruijgrok, Elly Dutton, Qian Li, Man Wang, Ting Zou, Junru Wu, Zhaole Yang, and Mulugeta Tarekegne Tsegaye. Specially, Yiya, thank you so much for sharing with me your knowledge and for your constructive comments on my work. My neighbor Hang and my officemate Han, thanks a lot for our support for each other. Saskia and Menghui, my supportive paranymphs and close friends, thank you so much for assisting me in the final preparations.

Special thanks go to professors Xiaolu Yang, Anastasia Giannakidou and Dylan Tsai. Xiaolu, thank you very much for all your help during my data collection time in Tsinghua University; my study would not have progressed so smoothly without your help. Anastasia, I am grateful for your timely and insightful discussion with me, which is of great help for my research. Dylan, thank you so much for your constant encouragement and support as an "intellectual elder"; it is such an interesting and rewarding experience to collaborate with you.

Thanks also go to my close friends outside the field of linguistics: Jingli Xin and Zhongwu Sun, who have made sure that I have lived a diversified life in addition to academia.

Finally, I would like to give my sincere thanks to my families: my dearest grandpa Quanhai Zhang, my (adoptive) parents Ruihua Zhang and Hongmin Yang, my (biological) parents Xiuling Han and Canmin Zhang, my parents in-law Guizhi

Miao and Xueting Liu and my three sisters. I feel so grateful that I have so many close families. I would never have achieved so much without their unconditional love and constant support. Home is where my feet may leave, but never my heart.

Speaking about family, my dear husband and truthful companion Yingwei Liu, your love and devotion has been my external power source and because of you, I understand how to love and be loved.

My study at Leiden University was financially supported by the China Scholarship Council, as well as support from LUCL, and I would like to express my sincere gratitude for their support.

Chapter 1 General Introduction

Mandarin is a *wh*-in-situ language, in which *wh*-words remain at their base position just as their declarative counterparts do. Mandarin is also known to be a *wh*-indeterminate language, in which *wh*-words [1] such as *shénme*, can have an interrogative interpretation 'what' or a non-interrogative existential interpretation 'something'. Due to the *wh*-in-situ and *wh*-indeterminates nature of Mandarin, clausal typing (e.g. classify the clause as a *wh*-question) in Mandarin and the licensing of Mandarin *wh*-indeterminates have long been two intriguing topics. This dissertation investigates the clausal typing in Mandarin and the licensing of Mandarin *wh*-indeterminates, from the perspective of prosody and processing. In section 1.1, I briefly introduce what *wh*-in-situ questions, *wh*-indeterminates and clausal typing are, compare three *wh*-in-situ and *wh*-indeterminate languages and motivate why Mandarin *wh*-indeterminates require further investigations. In section 1.2, I present my research questions in this dissertation. In the final section, I conclude this chapter by providing an overview of the dissertation.

1.1 Background

1.1.1 An introduction to *wh*-in-situ, *wh*-indeterminates and clausal typing

Human languages are diversified on the surface. For instance, when asking a *wh*-question, most Indo-European languages (e.g. English) front their *wh*-words to the clause initial position while most East Asian languages (e.g. Mandarin) tend to keep the *wh*-words in their base position. This is a common typological distinction among languages based on the formation of *wh*-questions, with the former known to be *wh*-movement languages and the latter *wh*-in-situ languages (see Cheng, 1991, among others for a detailed discussion of the typology of *wh*-questions). Examples (1) and (2) illustrate the two types of *wh*-questions respectively.

(1) What$_i$ did John buy t$_i$? [*wh*-movement]

(2) 张三 买了 什么 (呢)? [*wh*-in-situ]
Zhāng Sān mǎi-le shénme (ne[2])?
Zhang San buy-PERF what (SFP)
'What did Zhang San buy?'

The distinction of *wh*-movement languages and *wh*-in-situ languages are based on the way *wh*-questions are formed. Looking at the interpretations of *wh*-words, there are languages that are known to be *wh*-indeterminate languages such as Japanese,

[1] Unless specially mentioned, all the *wh*-words refer to arguments such as *what* and *who*, instead of adjuncts such as *why* and *how*. It should be noted that not all *wh*-words are *wh*-indeterminates, e.g., adjuncts such as *why*.
[2] The sentence final particle (SFP) *ne* is optional in Mandarin.

Korean, Mandarin and Vietnamese, in which *wh*-words can have both interrogative and non-interrogative interpretations (Kuroda, 1965; Huang, 1982; Nishigauchi, 1990; Cheng, 1991, 1994; Lin, 1998, 2014, among others). In these languages, *wh*-words have no inherent quantificational force and behave like variables (Cheng, 1991), whose interpretations are subject to licensing by (sentential) operators or licensors. Under different operators or licensors, *wh*-words can have interrogative, universal, existential or free choice interpretations[3]. Take Mandarin for example. As illustrated in (3a-b), a maximality operator (e.g. *dōu* 'all') can license the universal interpretation of the *wh*-word *shénme* 'everything' (see Giannakidou & Cheng, 2006; Xiang, 2008; Cheng, 2009 for the analysis of *dōu* as a maximality operator) and a nonveridical context[4] (like negation) can license the existential interpretation of the same *wh*-word *shénme* 'something/anything'. Henceforth, I refer to the *wh*-word as a *wh*-existential when it is interpreted as an existential as in (3b). For declaratives containing *wh*-existentials[5] like (3b), I refer to them as *wh*-declaratives, in order to distinguish them from regular declaratives.

(3) a. 张三 什么 都 买。 [universal]
Zhāng Sān shénme dōu mǎi.
Zhang San SHENME[6] all buy
'Zhang San buys everything.'

b. 张三 不想 买 什么。 [*wh*-existential]
Zhāng Sān bùxiǎng mǎi shénme.
Zhang San not want buy SHENME
'Zhang San doesn't want to buy anything.'

As illustrated in (4), when there is no overt licensor, the sentence containing the same *wh*-word *shénme* is interpreted as a *wh*-question, with the *wh*-word licensed by the null interrogative operator (Q) at Spec-CP or C^0. Henceforth I refer to the *wh*-word as a *wh*-interrogative when it is interpreted as a question word as in (4). According to Tsai (1994, 1999), all the licensors/operators of Mandarin *wh*-words are merged at the CP level and hence a *wh*-interrogative can only be interpreted at the sentence level.

(4) (+Q) 张三 买了 什么？ [*wh*-interrogative]
Zhāng Sān mǎi-le shénme?
Zhang San buy-PERF SHENME
'What did Zhang San buy?'

[3] The free choices conditions are more complicated to be generalized, see Cheng and Giannakidou (2013) for the licensing details of free choice interpretations in Mandarin.
[4] We will briefly introduce nonveridical contexts in section 1.1.2. The definition and all the examples of "nonveridicality" will be given at length in chapter 4.
[5] *Wh*-existentials can also be licensed in non-declaratives contexts which will be introduced in example (7) but in this dissertation, we mainly focus on *wh*-existentials in declaratives for the ease of discussing clausal typing of questions and declaratives.
[6] From here on, I gloss *shénme* simply as SHENME, as it can have different interpretations.

One typical question concerning *wh*-in-situ languages is the question of how clausal typing is realized in these languages. According to the Clausal Typing Hypothesis by Cheng (1991: 29), "Every clause needs to be typed. In the case of typing a *wh*-question, either a *wh*-particle in C^0 is used or else fronting of a *wh*-word to the Spec of C^0 is used, thereby typing a clause through C^0 by spec-head agreement." Hence *wh*-movement languages such as English type their *wh*-questions by fronting their *wh*-words, while *wh*-in-situ languages realize that by utilizing a *wh*-particle. For instance, in Japanese the presence of *ka* at clause final position (e.g., C^0) can type a *wh*-question. However, this hypothesis requires further considerations. Taking Mandarin as an example, it is problematic to take the particle in *wh*-questions (*ne*, a type of SFP) to be a *wh*-particle: First, *ne* is optionally used in *wh*-questions; second, *ne* appears in matrix *wh*-questions but it does not appear in embedded *wh*-questions; third, in addition to *wh*-questions, *ne* is also used in A-not-A questions[7]; moreover, *ne* appears not only in questions, with a different prosody it can also appear in declaratives (Li, 2006; Constant, 2014). If *wh*-particles and the movement of *wh*-words are not present to type a *wh*-question, then how is clausal typing realized in a language like Mandarin?

Recent studies also show that sentence final particles and some specific intonation are in complementary distribution (Zhang, 2014; Tang, 2015; Wakefield, 2016), and a strong version of the relations between particles and intonation claims that intonation and sentence final particles are the same thing, just in different forms (Wakefield, 2016). Based on an analysis of the historical changes of Mandarin particles, tones and intonations, Feng (2015) proposes that Mandarin particles can be analyzed as a variant of intonation with the former being a segmental realization of the latter. If there is indeed such a correlation between sentence final particles and intonation, it indicates that intonation or prosody could potentially do the clausal typing for *wh*-questions. Although this line of analysis requires more empirical and theoretical support, it may shed light on an alternative analysis of clausal typing, namely, from the perspective of intonation/prosody. This is one of the aims of this dissertation, i.e., to investigate whether and how prosody types sentences containing *wh*-words (i.e. to differentiate questions from declaratives) when there is no *wh*-movement or particles.

Japanese, Korean and Mandarin are the three commonly discussed *wh*-in-situ and *wh*-indeterminate languages. Although in this dissertation I investigate the licensing and clausal typing of *wh*-indeterminates in Mandarin, with a focus on *wh*-interrogatives in questions and *wh*-existentials in declaratives, I briefly compare *wh*-interrogatives and *wh*-existentials in Japanese and Korean with that in Mandarin, through which I show that the case of Mandarin merits further investigations.

1.1.2 *Wh*-indeterminates in Japanese, Korean and Mandarin

Different from Mandarin, which bears a Subject-Verb-Object (SVO) word order and does not allow scrambling, both Japanese and Korean bear an SOV word order and

[7] A-not-A question is a type of yes-no question in Mandarin (and other Chinese languages) offering a choice between an affirmative sentence and its negative counterpart.

allow the scrambling of *wh*-words. For the ease of comparisons with Mandarin, I focus on the unmarked word order of Japanese and Korean (SOV) instead of the scrambled version. I discuss the licensing of *wh*-interrogatives and *wh*-existentials in these two languages and the clausal typing of questions and declaratives containing them, as compared with Mandarin.

Japanese. As mentioned above, Japanese is also a *wh*-indeterminate language where *wh*-words can have interrogative or non-interrogative interpretations, depending on the particles (Kuroda, 1965; Nishigauchi, 1990). According to Tsai (1994, 1999), different from Mandarin in which *wh*-words can only be interpreted at the CP level, Japanese *wh*-words can be licensed at a lower level by particles and thus interpreted at the DP level. For instance, the *wh*-word *dare* has the existential interpretation of 'someone' when combined with particle *ka* as in *dare-ka*, as shown in (5a). Particles in Japanese can license the interpretations of *wh*-words as operators license variables (Cheng, 1991; Watanabe, 1992). When no other particles are used, a sentence-final *wh*-particle *no/ka* at Spec-CP or C^o can license the interrogative reading yielding a *wh*-question, as illustrated in (5b).

(5) a. Dare-ka-kara henna tegami-ga todoi-ta. [*wh*-existential]
DARE-some-from strange letter-Nom arrived
'A strange letter came from someone.'

b. Dare-ga ki-masu-ka? [*wh*-interrogative]
DARE-Nom come-Q
'Who's coming?'

As for clausal typing, the use of *wh*-particles *ka/no*[8] sentence finally can already type *wh*-questions, according to Cheng (1991). Recent prosodic studies also demonstrate that *wh*-questions are marked with a post-*wh*-word lexical accents compression (Deguchi & Kitagawa, 2002; Ishihara, 2002, 2003; Kitagawa & Fodor, 2003, among others). A strong version based on the prosodic markings of *wh*-questions even claims that the post-*wh*-word compression not only marks the clause type but also allows Japanese *wh*-questions to have *wh*-in-situ. According to Richards (2010: 145), "languages try to create a prosodic structure for *wh*-questions in which the *wh*-phrase and the corresponding C^o (clausal-final particle *no*/*ka* in Japanese) are separated by as few prosodic boundaries as possible." Japanese creates this prosodic structure through the post-*wh*-word compression (so that the *wh*-phrase and the corresponding C^o have few prosodic boundaries in between) and hence Japanese *wh*-questions can have *wh*-in-situ (Richards, 2010; 2016).

Korean. Similar to Japanese, Korean is also a *wh*-indeterminate language where the interpretations of *wh*-words largely depend on particles. For example, *wh*-word

[8] The usage of *no/ka* in Japanese *wh*-questions is optional under some conditions (Yoshida & Yoshida, 1997). But the optionality of particles in Japanese appears to be more restricted than in Mandarin. Recently, it has been claimed that *no* has some pragmatic functions; for instance, its presence is associated with certain presuppositions (see Sudo, 2013). As the pragmatic usage of particles is not directly relevant to our study, we still describe the presence of particles *ka/no* at the clause-final as a way to identify questions in Japanese.

nwukwu has the existential interpretation of 'someone' when combined with particle *(i)nka* or *(i)nci* as in *nwukwu-(i)nka/(i)nci*, as illustrated in (6a). When no other particles are used, a sentence-final *wh*-particle *ni* at Spec-CP or C^o can license the interrogative reading, as illustrated in (6b).

(6) a. Nwukwu-inka-ka wass-ta.
nwukwu-some-Nom came-Dec
'Someone came.'

b. Nwukwu-ka wass-ni?
nwukwu-Nom came-Q
'Who came?'

As for clausal typing, particles can do the job but they are optionally used in Korean. Korean also utilizes prosody, especially pitch accent and prosodic phrasing to differentiate *wh*-questions and declaratives containing *wh*-words. Normally a *wh*-interrogative bears a high-pitch accent and a *wh*-existential bears a low-pitch accent; *wh*-questions are characterized by a post-*wh*-word de-phrasing, namely, a deletion of accentual phrasings following the *wh*-word (Jun & Oh, 1996; Shin, 2005; Yun, 2012).

Although Japanese, Korean and Mandarin are all *wh*-indeterminate and *wh*-in-situ languages, the licensing of *wh*-indeterminates in Japanese and Korean in general depends on particles and the clausal typing of *wh*-questions relies on *wh*-particles or perhaps both *wh*-particles and prosody. As opposed to Japanese and Korean, the licensing of *wh*-indeterminates (*wh*-interrogatives and *wh*-existentials) and clausal typing in Mandarin are less straightforward. I will elaborate on it in two points.

First, with respect to the licensing of *wh*-indeterminates, Mandarin has no particles like Japanese and Korean to license *wh*-indeterminates and the particle *ne* is optionally used in *wh*-questions. *Wh*-existentials have been proposed to be licensed by nonveridical contexts only, in which the truth of a proposition cannot be entailed in the sentence. Examples (7a-d) illustrate typical nonveridical contexts, containing nonveridical operators like negation, questions, conditionals and epistemic modalities (Li, 1992; Lin, 1998; Xie, 2007; Lin, Weerman & Zeijlstra, 2014; Huang, 2017).

(7) a. 张三 不想 买 什么。 [negation]
Zhāng Sān bùxiǎng mǎi shénme.
Zhang San not want buy SHENME
'Zhang San doesn't want to buy anything.'

b. 张三 买了 什么 吗? [yes-no question]
Zhāng Sān mǎi-le shénme ma?
Zhang San buy-PERF SHENME yes-no particle
'Did Zhang San buy anything?'

c. 如果 张三 买了 什么, 别 生气。 [conditionals]
Rúguǒ Zhāng Sān mǎi-le shénme, bié shēngqì.
If Zhang San buy- PERF SHENME don't angry
'If Zhang San buys something, please don't be angry at him.'

d. 可能 张三 要去 买 什么。 [epistemic modality]
Kěnéng Zhāng Sān yàoqù mǎi shénme.
Possibly Zhang San go buy SHENME
'Possibly, Zhang San goes to buy something.'

Nevertheless, we observe that counter-examples also exist in which *wh*-existentials are licensed in veridical contexts. For instance, sentences in (8), which contain an adverb *zúotiān* ('yesterday') and a perfective maker *le* are typical veridical contexts, whose truth value is already ascertained. Yet the *wh*-word *shénme* can still have an existential interpretation. This observation challenges the general assumption in previous studies that *wh*-existentials are licensed in nonveridical contexts only (Lin, 1998; Xie, 2007; Lin et al., 2014; Huang, 2017), showing that the existing licensing conditions of Mandarin *wh*-existentials need to be revised.

(8) a. 张三 昨天 买了 点儿 什么 [veridical context]
Zhāng Sān zúotiān mǎi-le diǎnr shénme
Zhang San yesterday buy-PERF a.little SHENME
'Zhang San bought a little of something yesterday.' or
'What did Zhang San buy (a little of) yesterday? '

b. 李四 昨天 买了 个 什么 [veridical context]
Lǐ Sì zúotiān mǎi-le gè shénme
Li Si yesterday buy-PERF CL SHENME
'Li Si bought something yesterday. ' or
'What did Li Si buy yesterday? '

Second, with respect to clausal typing, Mandarin has neither *wh*-movement nor reliable *wh*-particles to type the *wh*-question. Furthermore, even if Mandarin can potentially utilize prosody to type clauses (as discussed in section 1.1.1), it cannot use the same post-*wh*-word lexical accent/accent phrase compression as in Japanese and Korean to disambiguate *wh*-questions from *wh*-declaratives, as Mandarin is a tone language with no lexical accent or accent phrase. It hence raises an additional point concerning clausal typing based on prosody. In particular, if prosody can do clausal typing in Mandarin, is there a specific clausal typing region in the sentence (pre-*wh*-word region, *wh*-word itself, post-*wh*-word region) or is the whole sentence marked with distinctive prosody? Both Korean and Japanese report that *wh*-questions are marked with a lexical accent/accent phrase compression in the post-*wh*-word region. But it remains unknown whether it is only the prosodic marking in the post-*wh*-word region or *wh*-word itself that can mark the clause type. Furthermore, different from Korean and Japanese which are SOV languages, Mandarin is SVO and when the question word is an object, the post-object or post-

wh-region is in fact very limited (see Huang, 1984 for the details). The longer pre-*wh*-word region in Mandarin as opposed to Korean and Japanese raises the question of whether the pre-*wh*-word region also plays a role in clausal typing, especially from the perspective of perception. In other words, whether listeners can identify the clause types in the pre-*wh*-word region based on prosody would be interesting to investigate. Taken together, different from Korean and Japanese in which the post-*wh*-word regions are often the focus for investigation, in Mandarin, it will be insightful to also investigate the prosodic markings of the pre-*wh*-word regions in *wh*-questions and *wh*-declaratives, for a better understanding and a complete picture of clausal typing in Mandarin.

In short, the clausal typing in Mandarin based on prosody needs to be thoroughly investigated, especially in the pre-*wh*-word region. Further investigations are also required to explore the detailed constraints and evidence in the licensing of *wh*-indeterminates in Mandarin, through for instance, investigating the cases where *wh*-existentials are licensed in *wh*-declaratives of veridical contexts.

1.2 Research questions

This dissertation addresses two general research questions each consisting of several subresearch questions. I will first present the general research questions and then elaborate on them at length.

(9) General research questions:

Q1: What kind of clausal typing mechanism(s) can we find in Mandarin?
Q2: How are *wh*-interrogatives and *wh*-existentials licensed in Mandarin?

Research question Q1 originates from the rethinking of traditional clausal typing hypothesis (Cheng, 1991): if Mandarin cannot use *wh*-movement or particles to type clauses, can prosody function as a clausal typing mechanism? I will address Q1 by investigating the prosodic markings of *wh*-questions and *wh*-declaratives, testing listeners' identification of the two clause types and seeking possible neurocognitive evidence for clausal typing based on prosody (i.e. electrophysiological evidence). Hence Q1 can be divided into three subresearch questions.

(10) Subresearch questions of Q1:

q1: Can *wh*-questions be differentiated from *wh*-declaratives through prosody? And if so, how?
q2: Can listeners make use of prosody to recognize the clause type or even anticipate the clause type?
q3: What neural correlates can we find for clausal typing based on prosody?

Subresearch question q1 in (10) aims to understand the role prosody plays in clausal typing by offering a detailed prosodic analysis on the first-hand audio recordings of *wh*-questions and *wh*-declaratives. Subresearch question q2 goes further to investigate the role of prosody in clausal typing from the perspective of perception,

namely, whether listeners can detect and anticipate clause types using prosody. Subresearch question q3 aims to build on neurocognitive evidence during real-time language processing for an accurate understanding of the prosodic clausal typing in Mandarin.

Research question Q2 concerns the licensing of Mandarin *wh*-existentials and *wh*-interrogatives, which are the two most frequently used interpretations of *wh*-indeterminates in Mandarin. I can address question Q2 by addressing two subquestions.

(11) Subresearch questions of Q2:

q1: What detailed licensing constraints and evidence can we find in the licensing of *wh*-existentials?
q2: What processing evidence can we find in the licensing of *wh*-interrogatives?

Subresearch question q1 in (11) originates from the fact that Mandarin *wh*-existentials are argued to be licensed only by nonveridical operators (Lin, 1998; Xie, 2007; Lin et al., 2014; Huang, 2017) though counter-examples also exist. We aim to discuss the licensing environments of *wh*-existentials by investigating the counter-examples. Subresearch question q2 can be addressed based on the evidence of online reading, through which we aim to find evidence for the covert dependency and licensing between the *wh*-interrogative and the interrogative operator (Q) at Spec-CP or C^{o}.

1.3 Overview of the chapters

This dissertation is organized as follows. Chapters 2, 3 and 6 focus on the clausal typing mechanism of *wh*-questions and *wh*-declaratives based on prosody. Chapters 4 and 5 focus on the licensing of *wh*-interrogatives and *wh*-existentials respectively. Below I introduce these chapters in this sequence.

Chapter 2 reports a production study on *wh*-questions and *wh*-declaratives, which are string identical. By analyzing the acoustic data, we come to understand how prosody is utilized to mark each clause type throughout the sentence, in particular in the pre-*wh*-word region. This chapter directly answers the subquestion q1 under the general question Q1.

Based on the results of the production experiment in Chapter 2, Chapter 3 reports an audio-perception and an audio-gating study (Grosjean, 1980). The perception study specially investigates whether listeners can differentiate the two clause types, *wh*-questions and *wh*-declaratives respectively, by completing a dialogue based on their identification of the *wh*-questions or *wh*-declaratives. Instead of a whole sentence as in the perception study, the audio-gating study segments the audios of *wh*-questions or *wh*-declaratives into several fragments, i.e., gates, and it specially tests whether listeners can anticipate clause types before hearing the *wh*-word and if yes, at which part of the sentence or at which gate can the anticipation happen. This chapter directly answers subquestion q2 under the general question Q1.

The perception and gating studies in Chapter 3 are offline studies that help to demonstrate the role of prosody on clause type identification and anticipation. With

respect to the role of prosody in clausal typing during online sentence processing, I seek the direct evidence and neural correlates by conducting two auditory ERP (Event-Related Potentials) studies reported in Chapter 6. These two auditory ERP studies examine *wh*-questions and *wh*-declaratives preceded by contexts biasing each clause type and their manipulated conditions. Experiment 1 tests *wh*-questions and *wh*-declaratives and their cross-spliced conditions. Audios of *wh*-questions and *wh*-declaratives are cross-spliced from the onset of *wh*-words onwards, in other words, the *wh*-word and the following constituents from a *wh*-question audio are spliced and combined to the pre-*wh*-word constituents of a *wh*-declarative audio, and vice-versa. It investigates the electrophysiological evidence of clausal typing incongruity. To be specific, when listeners expect a question word, as predicted by the context and the prosody of the pre-*wh*-word region (Q), they hear a *wh*-existential from the *wh*-declarative (D) instead (with a different prosodic marking as compared with the expected one), leading to an incongruent clause type (Q-D). By the same token, when they expect a *wh*-existential from the *wh*-declarative, they hear a *wh*-interrogative from the *wh*-question instead (D-Q). Experiment 2 manipulates the congruity between the contexts participants hear biasing *wh*-questions/*wh*-declaratives and the critical sentences of *wh*-questions/*wh*-declaratives participants hear. It investigates whether we can find neural correlates for detecting the early clausal typing incongruity based on prosody already at the subject position of *wh*-questions/*wh*-declaratives. The two ERP results directly address subquestion q3 under the general question Q1.

Chapter 4 offers a theoretical discussion on the licensing environments of *wh*-existentials based on empirical evidence, by focusing on *wh*-declaratives containing *diǎnr* 'a little' and its licensing on *wh*-existentials in veridical contexts. This chapter challenges the assumption that the licensing of *wh*-existentials is restricted in nonveridical contexts (Lin, 1998; Xie, 2007; Huang, 2017) and discusses the role of *diǎnr* in licensing *wh*-existentials as well as its detailed constraints in the licensing. Chapter 4 directly answers subquestion q1 under general question Q2.

Chapter 5 investigates the covert licensing of *wh*-interrogatives by conducting online-reading studies through word-by-word self-paced reading paradigm (Just, Carpenter & Wooley, 1982), which closely resembles natural reading. Different from the other chapters that utilize prosody, this chapter collects pure processing evidence in *wh*-questions as compared with declaratives with indefinite noun phrase (e.g. 'a classmate'). The working hypothesis is as follows. If processing a *wh*-question requires the construction of a covert dependency between the *wh*-phrase and the interrogative operator (Q) at Spec-CP or C^{o}, the processing cost in *wh*-questions is thus higher than that in declaratives. Chapter 5 addresses subquestion q2 under general question Q2.

Finally, Chapter 7 concludes the dissertation and summarizes the answers to the research questions posed in this chapter. Since the clausal typing mechanism of Mandarin *wh*-questions and *wh*-declaratives is based on prosody, I extend the traditional clausal typing hypothesis of *wh*-questions (Cheng, 1991) by integrating prosody into it. Furthermore, I summarize the discussions of how *wh*-interrogatives and *wh*-existentials are licensed in Mandarin based on all the investigations, as well as its implications for the licensing mechanism of *wh*-indeterminates cross-linguistically.

Chapter 2 *Wh*-question or *Wh*-declarative? Prosody Makes the Difference

2.1 Introduction

As introduced in Chapter 1, Mandarin Chinese is a *wh*-in-situ language in which *wh*-words remain at their base position just as their declarative counterparts do, as illustrated in (1a-b). Mandarin is also known to be a *wh*-indeterminate language (like Japanese and Korean) in which *wh*-words like *shénme* can have non-interrogative interpretations, see for instance (1c) (Huang, 1982; Cheng, 1991; Li, 1992; Lin, 1998). (1c) is a declarative sentence (*wh*-declarative) and the *wh*-word *shénme* is interpreted as an existential/indefinite, meaning "something". (1d) is the interrogative counterpart of (1c). As we can see, (1c) and (1d) are string identical. They both contain the word *diǎnr* 'a little', which is considered to be a determiner with existential quantificational force (Tsai, 2010), which licenses the indefinite reading of *shénme* in (1c).

(1) a. 张三 买了 书。 [declarative]
Zhāng Sān mǎi-le shū.
Zhang San buy-PERF book
'Zhang San bought a book.'

b. 张三 买了 什么？ [*wh*-question]
Zhāng Sān mǎi-le shénme?
Zhang San buy-PERF what
'What did Zhang San buy?'

c. 张三 买了 点儿 什么。 [*wh*-declarative]
Zhāng Sān mǎi-le diǎnr shénme.
Zhang San buy-PERF a.little SHENME
'Zhang San bought a little of something.'

d. 张三 买了 点儿 什么？ [*wh*-question]
Zhāng Sān mǎi-le diǎnr shénme?
Zhang San buy-PERF a.little SHENME
'What did Zhang San buy (a little of)?'

When (1c) / (1d) are presented in written form without a punctuation mark and out of context, they are in principle ambiguous between a *wh*-declarative and a *wh*-question interpretation.[9] Hence the clause types cannot be marked as in a *wh*-

[9] In the absence of a punctuation mark and out of context, the ambiguity between 1(c) and 1(d) is confirmed by an empirical study (a reading study), which is reported in Chapter 4.

movement language (e.g. English) or a language (like Japanese) that uses particles a lot, as introduced in Chapter 1.

Previous studies have shown that when declaratives are string identical with questions, the clause type is prosodically marked (Bolinger, 1978; Ohala, 1983, 1984; Jun & Oh, 1996; Frota, 2002; Face, 2004; Vion & Colas, 2006; Baltazani, 2007, among others). For instance, in Greek, yes-no questions are string identical with declaratives and interrogativity is prosodically encoded. In particular, as shown in Baltazani (2007) the two clause types differ with respect to the nucleus pitch accent (NPA) and the boundary tone (BT); in yes-no questions the NPA is L* and the BT is H-L%, while in declaratives the NPA can vary among a H*, H*+L or L+H* and the BT is L-L%. In Portuguese, yes-no questions are also string identical with declaratives and interrogativity is also prosodically encoded. As reported in Frota (2002), a Portuguese declarative often bears an NPA H+L* and the BT is L%; a yes-no question often bears an NPA H+L* but the BT is L-H% or H%.

In Mandarin, yes-no questions can be syntactically marked or unmarked[10]; when syntactically unmarked, they are also string identical with their declarative counterparts and the clause types are also prosodically encoded. Prosodic markings in Mandarin[11] are often investigated from either a global perspective like the sentence F0 curve/contour or local prosodic features like duration, F0 (range) or intensity (range) on the syllable or word level or a combination of both (Shi, 1980; Shen, 1994; Yuan, 2004; Liu, 2009; Jiang & Chen, 2011). As reported in previous studies, Mandarin yes-no questions are marked with a higher sentence F0 curve as compared with their declarative counterparts (Shi, 1980; Shen, 1990; Shen, 1994; Yuan, 2004, 2006; Jiang & Chen, 2011, among others) and the biggest F0 difference between the two clause type often lies in the final syllable (Yuan, 2004, 2006).

The prosodic marking of clause types has mainly been examined for yes-no questions in comparison with their declarative counterparts, as introduced above. For the identical strings of *wh*-declaratives and *wh*-questions as in (1c-d), so far, only one study investigates the prosodic markings of them and finds that *wh*-words in *wh*-questions bear higher pitch and expanded pitch range than in *wh*-declaratives (Liu, Li & Jia, 2016). The other two relevant studies also investigate the prosodic marking of sentences containing *wh*-words (as question words or as indefinites), although not on string identical cases and they also find that *wh*-words have higher pitch and expanded pitch range when used as question words than as indefinites (Hu, 2002; Dong, 2009). Although we know based on existing studies that there is a different F0 marking on *wh*-words, the prosodic properties of other parts of the *wh*-declaratives/*wh*-questions remain to be investigated. In terms of prosodic features, the existing studies have mainly investigated pitch, and it is not clear whether other prosodic features can mark the two clause types (like the word duration or syllable intensity).

---

[10] In Mandarin, the most frequently used yes-no question is string identical to its declarative counterpart but can have a sentence-final yes-no question particle *ma* used optionally; when *ma* is not used, the yes-no question is not marked. Here I am only discussing the unmarked yes-no questions.

[11] As a tone language, debate exists on whether Mandarin has NPA and BT (see Xu & Mok, 2011 for discussions about BT).

In this chapter, we scrutinize the prosody of *wh*-declaratives and *wh*-questions investigating the following research questions: (1) Do *wh*-declaratives and *wh*-questions as in (1c) and (1d) differ in terms of prosodic marking? 2) If they do, which is the first point in the utterance that the two start to differ, and in which prosodic features do they differ?

The current chapter is organized as follows. In section 2.2, we discuss relevant studies on the prosody of *wh*-declaratives and *wh*-questions. Section 2.3 presents the results of a production experiment. Section 2.4 concludes and examines the implications of the results on the focus of *wh*-questions and *wh*-declaratives.

2.2 Relevant studies

The prosodic markings of questions and declaratives in Mandarin have been a topic of research for some time, from the early descriptive and introspective studies (Chao, 1932, De Francis, 1963, among others) to the more recent laboratory-based studies (Shi, 1980; Wu, 1982; Gårding, 1987; Shen, 1990; Shen, 1994; Hu, 2002; Yuan, 2004, 2006; Dong, 2009; Liu, 2009; Liu, Li & Jia, 2016, among others). Most of the above studies discuss the prosody of yes-no questions in comparison to their string identical declaratives while few studies focused on the prosodic marking of sentences containing *wh*-words (as question words or as indefinites). Below we discuss these limited studies containing *wh*-words in more detail (Hu, 2002; Dong, 2009; Liu, Li & Jia, 2016).

Hu (2002) reports that the *wh*-word in a *wh*-question bears the focus prominence with an expanded pitch range, while in yes-no questions (containing *wh*-words as indefinites meaning 'something'), the *wh*-word has a reduced pitch range. She conducted a production experiment, comparing yes-no questions containing *wh*-words with *wh*-questions containing *wh*-words as shown in example in (2).[12] Here the indefinite reading of the *wh*-word is triggered by the yes-no question particle *ma* (for detailed discussions see Cheng 1991, 1994; Li, 1992; Lin, 1998, among others).

(2) a. 张三 买了 什么 呢？ [*wh*-question]
Zhāng Sān mǎi-le shénme ne?
Zhang San buy-PERF SHENME *wh*-particle
'What did Zhang San buy?'

b. 张三 买了 什么 吗？ [yes-no question]
Zhāng Sān mǎi-le shénme ma?
Zhang San buy-PERF SHENME yes-no-particle
'Did Zhang San buy something?'

In (2b) where *wh*-words are interpreted as an indefinite, Hu finds that it is the verb that bears the prosodic prominence with greater expanded pitch range. Hu doesn't find any consistent duration or intensity differences between the two conditions.

[12] There are more experimental conditions in Hu (2002). Here we only list the ones that are related to the current study.

However, Hu's results should be taken cautiously, as they are based on a very small sample of speakers (4 speakers), who show a lot of variations across themselves.

Dong (2009) compares *wh*-questions and *wh*-declaratives in terms of F0. As illustrated in (3), the *wh*-word is interpreted as a question word in (3a) and as an indefinite in (3b) meaning 'someone'. Dong's results show that *wh*-word is the most prosodically prominent item in *wh*-questions as represented by its expanded pitch range, while in contrast it has a compressed pitch contour in *wh*-declaratives; instead, the verb bears the prosodic prominence with expanded pitch range in *wh*-declaratives.

(3) a. 梅 揶揄了 谁？ [*wh*-question]
Méi yéyú-le shéi?
Mei ridicule-PERF who
'Who did Mei ridicule?'

b. 好像 梅 揶揄了 谁。 [*wh*-declarative]
Hǎoxiàng Méi yéyú -le shéi.
Seem Mei ridicule-PERF someone
'It seems that Mei ridiculed someone.'

Although Dong examines only F0, duration (and intensity) can also be informative prosodic cues to check (Shen, 1993; Chuang & Fon, 2010).

Liu, Li and Jia (2016) investigate the prosodic marking of *wh*-questions with the string identical *wh*-declaratives. They compare the two clause types from the perspective of both local prosodic features (prosodic properties of each word) and global prosodic features (sentence F0 curve). The examples (4a-b) illustrate the comparisons between the two clause types. Note that Liu et al. use the *wh*-word *shénme* as a modifier of the noun, different from the cases we discussed above.

(4) a. 张三 打算 吃 点儿 什么 糕？ [*wh*-question]
Zhāng Sān dǎsuàn chī diǎnr shénme gāo?
Zhang San intend to eat a.little SHENME cake
'What kind of cake does Zhang San intend to eat?'

b. 张三 打算 吃 点儿 什么 糕。 [*wh*-declarative]
Zhāng Sān dǎsuàn chī diǎnr shénme gāo.
Zhang San intend to eat a.little SHENME cake
'Zhang San intends to eat whatever cake.'

The results of Liu et al. show that local prosodic features contribute more in differentiating the two clause types than global features. Contrary to the results by Dong and Hu, Liu et al. do not find any prosodic differences at the verb between the two clause types, which may be due to the fact that Liu et al. use the *wh*-word *shénme* as a modifier of the object noun while the other two studies use the *wh*-word as the object.

From these limited studies, we can only conclude that *wh*-words have expanded pitch range when used as question words as compared with their indefinite counterparts. Our research questions with respect to when *wh*-declaratives and *wh*-questions start to be different in prosody, and in what prosodic properties they differ remain to be investigated.

2.3 Production experiment

2.3.1 Participants

Forty native speakers of Beijing Mandarin (23 females and 17 males, $\bar{x}$ age = 21 years old) were paid to participate in the production experiment. All of them were born and raised in Beijing and at the time of recording they were students at Tsinghua University. None of them reported any speech disorder or vision impairment (after correction). Prior to recording informed written consent was obtained from each participant.

2.3.2 Experimental materials

We created a total of 56 stimuli; half of the stimuli were *wh*-declaratives (see example in 5a), while the other half were *wh*-questions (see example in 5b). *Wh*-declaratives were string identical to their corresponding *wh*-questions except for the punctuation at the end of the sentence. As shown in (5), for construcing the stimuli, we used the following word order: Subject (proper name, e.g. *Táo Wēi* "Tao Wei"), Adverb (e.g. *zúotiān* "yesterday"), Verb (e.g. *ná* "bring") + Perfective marker (*le*), *diǎnr*, Direct Object (*shénme* "what/something"), preposition phrase (e.g. *gěi* Líu Gāng "to/for Liu Gang"). We chose this word order as it is a basic word order in Mandarin (Li, 1990). Each stimulus consisted of 12 syllables and the stimulus length was constant across clause types and items. As Mandarin is a Tone (T) language with four full lexical tones (T1 a high level tone, T2 a rising tone, T3 a low tone and T4 a falling tone) and a neutral tone (T0), we kept the combination of tones constant across items and clause types for all constituents but the verb. For the verb, we included all four possible tones, to obtain more natural stimuli. An example of a stimulus set is given in (5).

(5) a. 陶薇 昨天 拿了 点儿 什么 给 刘刚。 [*wh*-declarative]
Táo Wēi zúotiān ná-le diǎnr shénme gěi Líu Gāng.
T2 T1 T2 T1 T2-T0 T3 T2 T0 T3 T2 T1
Tao Wei yesterday bring-PERF a.little something to Liu Gang
'Tao Wei brought a little something to Liu Gang yesterday.'

b. 陶薇 昨天 拿了 点儿 什么 给 刘刚？ [*wh*-question]
Táo Wēi zúotiān ná-le diǎnr shénme gěi Líu Gāng?
T2 T1 T2 T1 T2-T0 T3 T2 T0 T3 T2 T1
Tao Wei yesterday bring-PERF a.little what to Liu Gang
'What did Tao Wei bring (a little) to Liu Gang yesterday?'

2.3.3 Procedure

The recordings took place in a sound-proof booth in a lab of the Department of Foreign Languages and Literatures at Tsinghua University in Beijing. For recording we used a head-worn unidirectional dynamic microphone (Shure SM10A) which was connected to an external sound card (UA-1G), and Audacity software (sampling rate 44.1 kHz, 16 bit, mono). The stimuli were presented on screen without any preceding context using Praat (Boersma & Weenink, 2016) and the presentation pace of each stimulus was controlled by the experimenter. Participants were instructed to read silently the stimulus on screen to understand its meaning, and then to utter it as if they were talking with someone. Once they had uttered the sentence, the new stimulus appeared on screen. A pseudo-randomized list of stimuli was prepared for every participant.

We recorded a total of 2240 utterances (40 participants × 56 stimuli). 338 utterances were excluded from any further analysis due to slips of the tongue, disfluencies and unnatural pausing.

2.3.4 Acoustic analysis

The remaining 1902 stimuli/utterances were manually annotated using Praat (Boersma & Weenink, 2016), as shown in Figure 1. Then, we obtained the following measurements using a number of Praat scripts.

Duration
(i) Utterance duration in ms.
(ii) Word duration in ms; this was calculated based on the syllable duration, see Figure 1.

F0
Mandarin tones are dynamic pitch targets (Xu, 2001; Xu & Wang, 2001), but for the ease of measurements, we used the notation L, H, LH and HL to describe the four full lexical tones of Mandarin, T1 (H), T2 (LH), T3 (L), T4 (HL) (Duanmu, 2004), and hence we measured the following F0's.
(iii) F0-maximum (H) of the syllable that bore T1 (high level tone).
(iv) F0-minimum and then F0-maximum (LH) of the syllable that bore T2 (rising tone).
(v) F0-minimum (L) of the syllable that bore T3 (low tone).
(vi) F0-maximum and then F0-minimum (HL) of the syllable that bore T4 (falling tone).
(vii) For T0 (neutral tone) of the perfective marker *le*, following Li (2002), we measured first the F0-maximum and then the F0-minimum, when the preceding syllable (verb) bore T1, T2 or T4; while we measured first the F0-minimum and then the F0-maximum, when the preceding syllable bore Tone 3, as illustrated in Figure 2. For the second syllable of the *wh*-word *shénme*, namely, *me*, when found in isolation it bears Tone 0. However, in our data, it behaved like a rising tone (T2),

and thus we treated it as such, measuring the F0-minimum and then the F0-maximum.

The obtained F0 values in Hz were converted into semitones (ST) to reduce variation across speakers; following Li and Chen (2012), for female speakers we used formula (i) ST = 12log2 (Hz/100), while for male speakers we used formula (ii) ST = 12log2 (Hz/50).

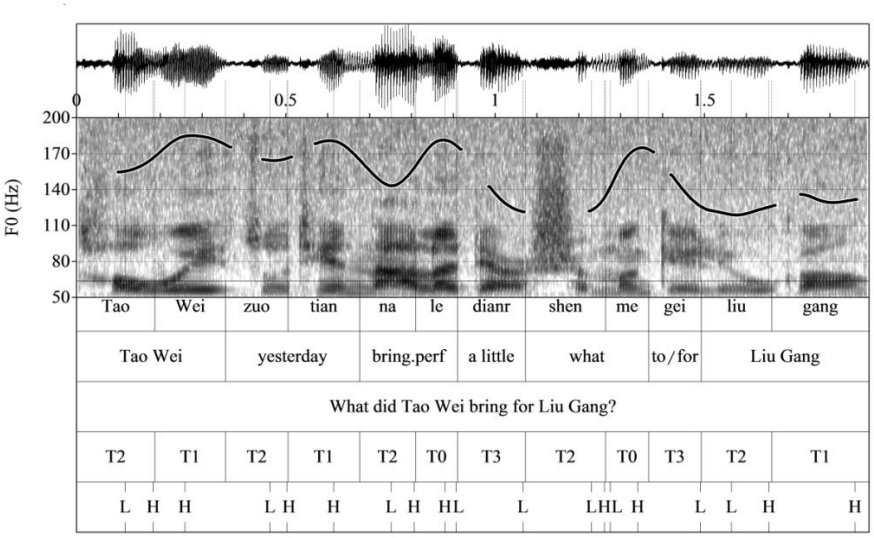


Figure 1. An exemplar waveform and spectrogram from a male participant with superimposed F0-contours, syllables, glosses and F0 measurements obtained based on the specific tones.

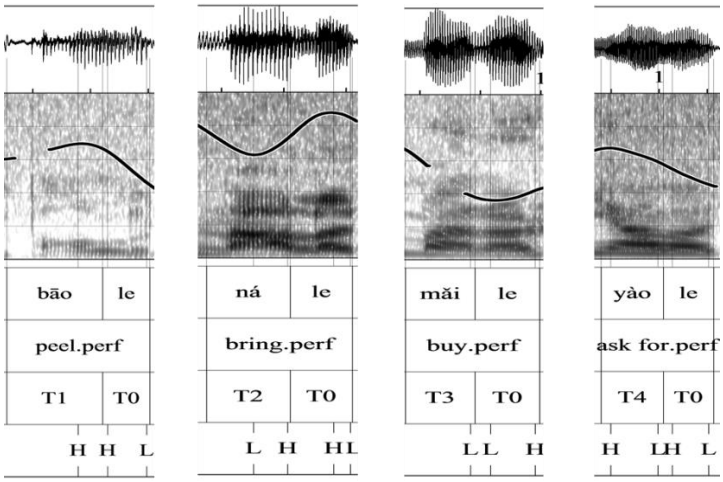


Figure 2. The F0 measurement of le obtained based on the preceding T1, T2, T3 and T4 respectively.

F0 range

We also calculated the F0 range in ST of the *wh*-word (*shénme*) and the post *wh*-word region, following previous studies (Dong, 2009; Liu et al., 2016).

(viii) F0 range of *shén* and *me*. Given that *shén* is a rising tone and that *me* also behaves like a rising tone, we calculated the F0 range of *shén* and *me* respectively, *shén* as F0-maximum of *shén* – F0-minimum of *shén* and *me* as F0-maximum of *me* – F0-minimum of *me*.

(ix) F0 range of post-*wh*-word region, namely, the preposition phrase (e.g. *gěi* plus indirect object *Líu Gāng* "to/for Liu Gang"). Given that *gěi* carries a low tone, and that the first syllable of the indirect object bears a rising tone and the second syllable a high tone, the pitch contour in the whole preposition phrase is in general a rising contour; hence we calculated the F0 range of the preposition phrase as F0-maximum of the second syllable of the indirect object (e.g. *Gāng*) – F0-minimum of *gěi*. See Figure 3.

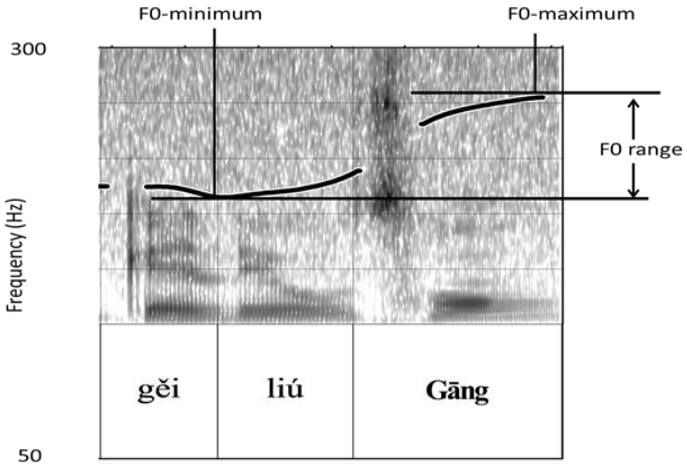


Figure 3. The F0 contours of the preposition phrase *gěi* indirect object and its F0-range obtained.

Intensity range

(x) Intensity range of each syllable defined as Maximum-Intensity – Minimum-Intensity (Chen 2005; Ouyang & Kaiser, 2015), see Figure 4.

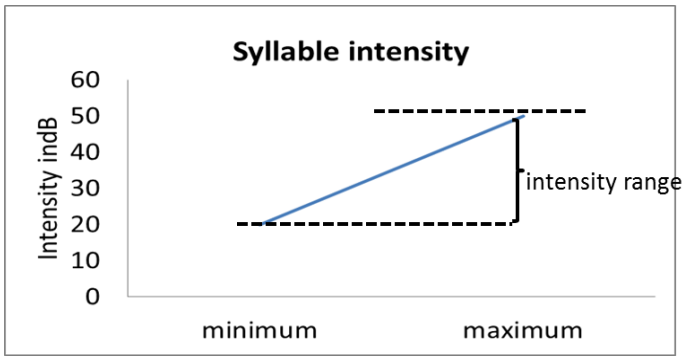


Figure 4. The Maximum intensity and Minimum intensity of a syllable and its intensity range.

2.3.5 Statistical analysis

As mentioned in section 2.1, the aim of the production experiment is to examine the prosodic differences between *wh*-declaratives and *wh*-questions. In other words, our aim is to examine the effect of clause type on the duration, F0 and intensity of *wh*-sentences. Hence, we ran a series of linear mixed-effects models using the lmerTest package (Kuznetsova, Brockhoff & Christensen, 2013) in R. Specifically, for every measurement, we first ran a null model with the relevant measurement as the dependent variable, and participants and items as random factors. A second model included in addition clause type as a fixed effect factor. Finally, we ran a third model that included the relevant measurement as the dependent variable, clause type as fixed factor, and participants and items as random factors, allowing by-participant and by-item random intercepts, and by-participant and by-item random slopes for clause type. Model fit was compared using the likelihood ratio test (Pinheiro & Bates, 2000; Bolker, Brooks, Clark, Geange, Poulsen, Stevens, & White, 2009). See Appendix A for the details of the fitting models in each measurement.

2.3.6 Results

Utterance duration. Figure 5 presents the average utterance duration of *wh*-declaratives and *wh*-questions. We found an effect of clause type on utterance duration. The average duration of *wh*-declaratives ($\bar{x}$ = 2050 ms) is significantly longer than that in *wh*-questions ($\bar{x}$ = 2020 ms), [β = 28.634, S.E. = 9.307, t = 3.077, $p < 0.01$].

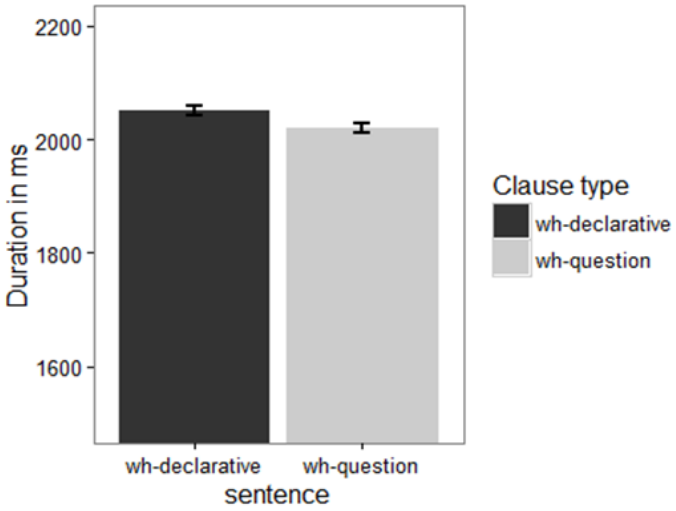


Figure 5. Mean sentence duration in ms with error bars showing standard error across clause types.

Word duration. Figure 6 presents the mean duration of all words in the utterance. We found an effect of clause type on the duration of the Subject, of the Verb plus the perfective marker *le*, of *diănr* and of the *wh*-word. Specifically, the duration of the Subject, Verb plus *le* and *diănr* in *wh*-declaratives ($\bar{x}$ = 347 ms, $\bar{x}$ = 289 ms and $\bar{x}$ = 171 ms respectively) are significantly longer than those in *wh*-questions ($\bar{x}$ = 341 ms, $\bar{x}$ = 261 ms and $\bar{x}$ = 166 ms respectively), [β = 5.626, $p < 0.01$; β = 27.993, $p < 0.001$; β = 5.339, $p < 0.001$]. The pattern changes when examining the duration of the *wh*-word *shénme*. *Shénme* in *wh*-declaratives ($\bar{x}$ = 294 ms) is significantly shorter than in *wh*-questions ($\bar{x}$ = 305 ms), [$\beta = -11.065$, $p < 0.01$]. When looking at the post-*wh*-word region, the preposition phrase (*gěi* plus indirect object) in *wh*-declaratives ($\bar{x}$ = 572 ms) does not differ from *wh*-questions ($\bar{x}$ = 572 ms) [$\beta = -0.397$, $p > 0.1$]. The detailed results of the mixed effects model can be found in Table 1.

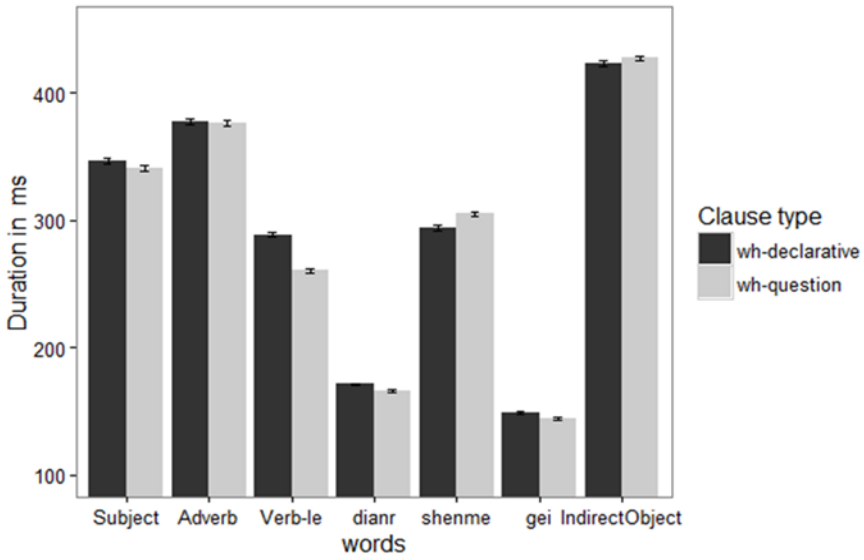


Figure 6. Mean word duration in ms with error bars showing standard error across clause types.

Table 1. Summary of the linear mixed effects models on the duration of each word and the sentence.

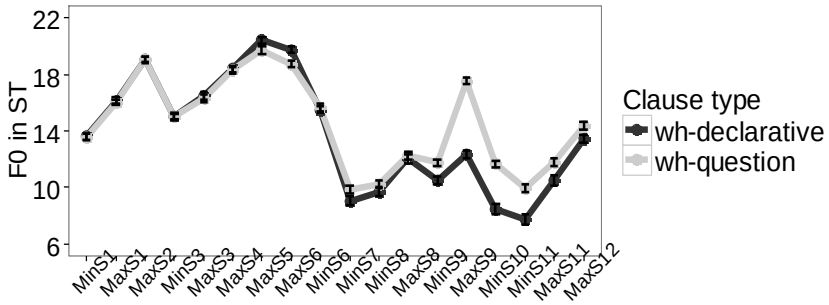
	Estimate β	Std. Error	t-value	p-value
subject	5.626	1.860	3.024	< 0.01
adverb	1.146	3.039	0.377	> 0.1
verb-*le*	27.993	2.730	10.253	< 0.001
diǎnr	5.339	1.239	4.310	< 0.001
shénme	− 11.065	3.588	− 3.084	< 0.01
preposition phrase	− 0.397	3.818	− 0.104	> 0.1

F0. Figure 7 presents the stylized means of F0 curves of the two clause types broken per verb tone. As shown, the most striking F0 difference between the two clause types is at the *wh*-word *shénme*, which shows a steep rise in *wh*-questions but is relatively flat in *wh*-declaratives, and the F0 in *wh*-questions remains higher than that in *wh*-declaratives until the end of sentence. To be specific, *shénme* in *wh*-declaratives has lower F0 at the F0-minimum of *shén* ($\bar{x}$ = 9.964 ST) [$\beta = -0.714$, $p < 0.001$], the F0-minimum of *me* ($\bar{x}$ = 10.462 ST), [$\beta = -1.630$, $p < 0.001$] and the F0-maximum of *me* ($\bar{x}$ = 12.392 ST), [$\beta = -5.298$, $p < 0.001$] than in *wh*-questions ($\bar{x}$ = 10.579 ST for F0-minimum of *shén*, $\bar{x}$ = 12.008 ST for F0-minimum of *me* and $\bar{x}$ = 17.625 ST for F0-maximum of *me*).

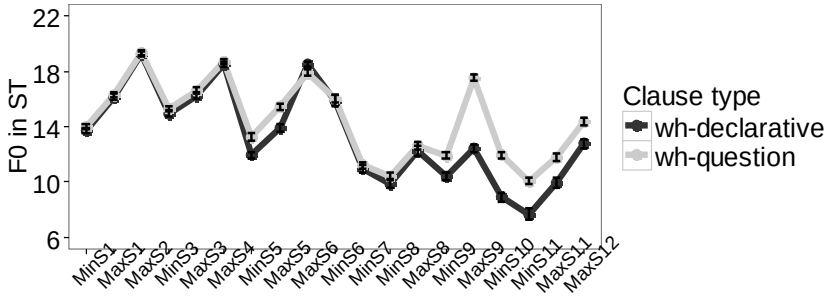
In the pre-*wh*-word region, we also found F0 differences at the verb when it bears T2. The F0-minimum and F0-maximum of the T2 verb in *wh*-declaratives ($\bar{x}$ = 11.947 ST, $\bar{x}$ = 13.929 ST) are lower than that of *wh*-questions ($\bar{x}$ = 13.274 ST, $\bar{x}$ = 15.465 ST) respectively, [$\beta = -1.351$, $p < 0.05$; $\beta = -1.561$, $p < 0.001$]. In the

post-*wh*-word region, *wh*-declaratives are continuously lower in F0 than in *wh*-questions. To be specific, the F0-minimum of *gěi* in *wh*-declaratives ($\bar{x}$ = 8.820 ST) is lower than that in *wh*-questions ($\bar{x}$ = 12.036 ST), [$\beta = -3.355$, $p < 0.001$]; the F0-minimum and F0-maximum of the first syllable of the indirect object in *wh*-declaratives ($\bar{x}$ = 7.601 ST, $\bar{x}$ = 10.262 ST) are also lower than that in *wh*-questions ($\bar{x}$ = 10.328 ST, $\bar{x}$ = 12.017 ST), [$\beta = -2.714$, $p < 0.001$; $\beta = -1.364$, $p < 0.001$]; Finally, the F0-maximum of the second syllable of the indirect object in *wh*-declaratives ($\bar{x}$ = 13.327 ST) is again lower that in *wh*-questions ($\bar{x}$ = 14.560 ST), [$\beta = -1.364$, $p < 0.001$]. The detailed results of the mixed effects model can be found in Table 2.

Verb=T1



Verb=T2



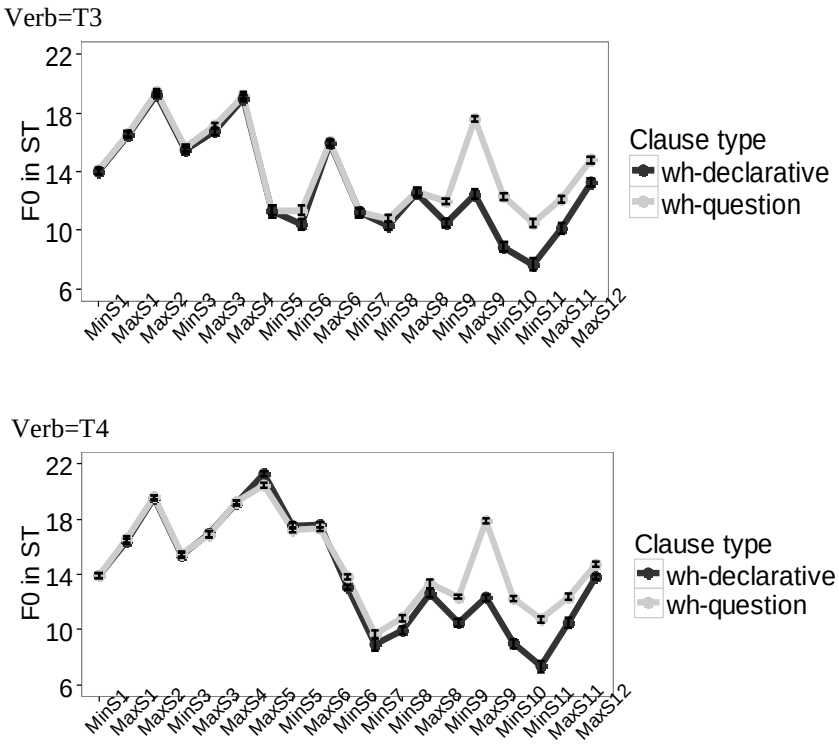


Figure 7. Stylized[13] means of F0 curves in ST across clause types with error bars showing standard errors.

Table 2. Summary of the linear mixed effects models on the F0 measurement with significant differences between clause types.

	Estimate *β*	Std. Error	*t*-value	*p*-value
F0-min *shén*	− 0.714	0.171	− 4.167	< 0.001
F0-min *me*	− 1.630	0.263	− 6.199	< 0.001
F0-max *me*	− 5.298	0.393	− 13.468	< 0.001
F0-min verb (T2)	− 1.351	0.541	− 2.497	< 0.05
F0-max verb (T2)	− 1.561	0.265	− 5.894	< 0.001
F0-min *gěi*	− 3.355	0.438	− 7.661	< 0.001
F0-min indirect object(1st syllable)	− 2.714	0.426	− 6.365	< 0.001
F0-max indirect object (1st syllable)	− 1.817	0.237	− 7.659	< 0.001
F0-max indirect object (2nd syllable)	− 1.364	0.310	− 4.401	< 0.001

[13] The stylized mean F0 curve is based on our F0 measurement as illustrated in Figure 1, namely we use a F0-maximum to represent T1, F0-minimum and F0-maximum to represent T2, F0-minimum to represent T3, and for T4 we use F0-maximum and F0-minimum.

F0 range. The F0 range differences between clause types were found at the *wh*-word and the post-*wh*-word region. The F0 range of *me* in *wh*-declaratives ($\bar{x}$ = 1.930 ST) is smaller than that in *wh*-questions ($\bar{x}$ = 5.617 ST), [$\beta = -3.653$, $p < 0.001$]. The F0 range of the preposition phrase after the *wh*-word is bigger in *wh*-declaratives ($\bar{x}$ = 4.507 ST) as compared with that in *wh*-questions ($\bar{x}$ = 2.524 ST), [$\beta = 1.975$, $p < 0.001$], indicating a compressed F0 range in the post-*wh*-word region of *wh*-questions. The detailed results of the mixed effects model can be found in Table 3.

Table 3. Summary of the linear mixed effects models on F0 range with significant differences between clause types.

	Estimate β	Std. Error	t-value	p-value
F0 range *me*	− 3.653	0.274	− 13.346	< 0.001
F0 range preposition phrase	1.975	0.487	4.054	< 0.001

Intensity range. Figure 8 shows the mean intensity range of each syllable (S) across the two clause types. As illustrated in Figure 8, the main differences between *wh*-declaratives and *wh*-questions are at the verb-*le* (S5 and S6), the *wh*-word *shénme* (S8 and S9) and *gěi* (S10). *Wh*-declaratives have a bigger intensity range at the verb ($\bar{x}$ = 20.210 dB), [$\beta = 1.119$, $p < 0.001$] and at *le* ($\bar{x}$ = 7.701 dB), [$\beta = 0.833$, $p < 0.001$] than *wh*-questions ($\bar{x}$ = 19.079 dB, $\bar{x}$ = 6.865 dB). The direction changes when looking at the *wh*-word. *Shén* and *me* in *wh*-declaratives are significantly smaller in intensity range ($\bar{x}$ = 9.477 dB, $\bar{x}$ = 7.745 dB) than those in *wh*-questions ($\bar{x}$ = 10.095 dB, $\bar{x}$ = 9.470 dB), [$\beta = -0.620$, $p < 0.001$; $\beta = -1.729$, $p < 0.001$]. Finally, for the preposition phrase after the *wh*-word, we only find intensity range differences at *gěi*, which is smaller in *wh*-declaratives ($\bar{x}$ = 15.315 dB) than that in *wh*-questions ($\bar{x}$ = 17.966 dB), [$\beta = -2.662$, $p < 0.001$). The detailed results of the linear mixed effects model are summarized in Table 4.

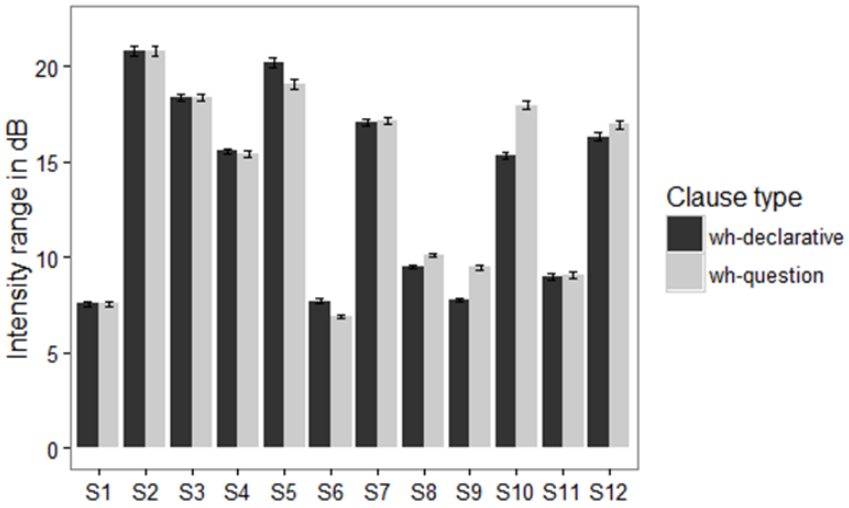


Figure 8. Mean intensity range of each syllable (S) across clause types with error bars showing standard errors.

Table 4. Summary of the linear mixed effects models on the intensity range with significant differences between clause types.

	Estimate *β*	Std. Error	*t*-value	*p*-value
verb	1.119	0.179	6.243	< 0.001
le	0.833	0.224	3.713	< 0.001
shén	− 0.620	0.157	− 3.963	< 0.001
me	− 1.729	0.351	− 4.927	< 0.001
gěi	− 2.662	0.461	− 5.767	< 0.001

In summary, our production study shows that given a string identical *wh*-declarative and *wh*-question, prosody indeed marks them differently. *Wh*-declaratives differ from *wh*-questions in word and sentence durations, F0 and F0 range and intensity range. When the sentence is uttered by the speaker as a question, it has a steep F0 rise and expanded F0 range at the *wh*-word, and also a longer word duration and a bigger intensity range. It is worth noting that before the *wh*-word and after the *wh*-word, there are also prosodic differences. Our study reveals that it is duration that provides an early cue to differentiate the two interpretations from the onset of the utterance (sentence subject), with *wh*-declaratives always longer than *wh*-questions and at the verb-*le* the difference reaches the peak. In terms of intensity range, the verb-*le* in *wh*-declaratives is bigger than that in *wh*-questions. It is of interest that in the post-*wh*-word region, *wh*-questions show a long-lasting higher pitch and their F0 range is smaller than that in *wh*-declaratives, indicating a F0 compression in *wh*-questions, consistent with previous studies (Xu, 1999; Kuo, Xu & Yip, 2007).

2.4 Discussion and conclusion

In the current study, we aim to investigate the role prosody plays in clausal typing by conducting a production study on the two clause types, the string identical cases of *wh*-questions and *wh*-declaratives. The detailed prosodic analysis has shown that the *wh*-word in *wh*-questions is manifested with a steep F0 rise and expanded F0 range, consistent with the results of previous studies on *wh*-sentences containing *wh*-words (Hu, 2002; Lee, 2005; Dong, 2009; Liu, 2009; Liu, Li & Jia, 2016), and also in general consistent with previous studies saying that the presence of a high pitch is often a property of question intonation cross-linguistically (Hermann, 1942). In addition, our study reveals that prosody marks *wh*-questions and *wh*-declaratives differently with different prosodic properties, ranging from utterance and word duration, F0 and F0 range to intensity range, not limited to F0.

Furthermore, we find that prosody (mainly duration) marks the clause type "early". From the onset of the utterance, *wh*-questions are always shorter than *wh*-declaratives in terms of word duration in the pre-*wh*-word region, which provides an early cue to differentiate the two clause types from the point of the sentence subject. This early durational property (a shorter duration in *wh*-questions as opposed to declaratives) can be perceived as another defining feature of Mandarin *wh*-questions, in addition to the commonly known F0 property; this finding of the duration property is in general consistent with studies on other languages reporting that duration plays a role in marking questions (Lindsey, 1985; Van Heuven & Van Zanten, 2005; Cangemi & D'Imperio, 2013).

In addition to the above findings on the different prosodic markings of the two clause types, our production results also shed light on the discussions of the focal property of *wh*-words and their prosodic realizations. *Wh*-words are claimed to be the focus of the sentence (Cho, 1990; Lambrecht & Michaelis 1998; Deguchi & Kitagawa, 2002; Ishihara, 2002, among others) when used in *wh*-questions, but few studies discuss the focus status of *wh*-words when used in different contexts, for instance, when used as question words ('what') as compared with non-question words ('something'). Cross-linguistically speaking, there is a general correspondence between focus and the prosodic marking: focused constituents are characterized with an expanded pitch range and the post-focal regions typically show a compressed F0 range (for Germanic languages, see Cruttenden, 2006, Fery & Kugler, 2008, among others; for Mandarin studies, see Xu, 1999; Yuan, 2004; Li, 2009; Chen, 2010; Xu, Chen & Wang, 2012, among others). In addition, a focused element typically has a longer duration and greater intensity (Xu, 1999; Chen, Wang & Xu, 2009; Li, 2009). In our production study, we also found that *wh*-questions have a raised and expanded F0 range, lengthened duration and greater intensity range at the *wh*-word *shénme* and an F0 range compression after *shénme*. Based on this, we can ascertain that Mandarin *wh*-words are focused in *wh*-questions and bear the prosodic prominence. In contrast, *wh*-words in *wh*-declaratives are largely suppressed with an almost flat pitch, a short duration and a small intensity range, as shown in our production study. This indicates that *wh*-words cannot be the focus in *wh*-declaratives, which is compatible with their interpretation (a narrow scope indefinite 'something').

To conclude, in this chapter we tackled the question of how prosody marks *wh*-questions and *wh*-declaratives respectively. Our production results have demonstrated that the two clause types differ in terms of various prosodic properties, not limited to F0; it is duration that marks the two clause types early as *wh*-declaratives are continuously longer than *wh*-questions in terms of word duration since the onset of the sentence (subject). The current study also implies that *wh*-questions and *wh*-declaratives also differ in terms of focus, namely, a *wh*-word in Mandarin is a focus when it is a question word (e.g., 'what') but it is not a focus when it is an indefinite ('something') used in a *wh*-declarative.

Chapter 3 Clause Type Anticipation Based on Prosody — An Audio-perception and Gating study

3.1 Introduction

In the previous chapter we investigate how prosody differentiates clause types from the perspective of the speakers. In the current chapter, we explore a related topic from the perspective of the listeners, namely, clause type anticipation. Before scrutinizing clause type anticipation, we first briefly introduce the anticipation in spoken sentences. Our daily communication involves anticipation in the oral speech processing and here anticipation often refers to predicting how the discourse of a speaker will evolve (Seeber, 2001). To be precise, spoken sentences proceed in an incremental manner and when a sentence unfolds over time, listeners interpret the information and formulate a prognosis of how the sentence may further unfold. Generally speaking, listeners can process incoming information rapidly to dissolve ambiguities and predict the most plausible interpretation from the sentence (Braun & Chen, 2012). For example, listeners can predict upcoming referents by utilizing intonation cues such as accent (see Braun & Chen, 2012 for details). To make predictions/anticipations, listeners use all kinds of linguistic knowledge and rely on every cue they may get, in combination of information from syntax, morphology, discourse-semantics and prosody/intonation (Kohn & Kalina, 1996; Braun & Chen, 2012).

With respect to anticipation and prosody, a few studies so far attempted to establish a link between the two. These studies mainly focus on the role of pitch accent/prominence in reference resolution or in information structure interpretation (Terken & Hirschberg, 1994; Dahan, Tanenhaus & Chambers, 2002; Weber, Braun & Crocker, 2006; Chen, den Os & de Ruiter, 2007, among others), how the stress pattern of the sentence can help to anticipate upcoming stressed syllables and the end of the speech sequence (Shields, McHugh & Martin, 1974; Buxton, 1983), or how listeners can use prosodic information to predict upcoming syntactic structure or to resolve syntactic ambiguity (Beach, Katz & Skowronski, 1996; Kjelgarrd & Speer, 1999; Carlson, Clifton & Frazier, 2001; Snedeker & Trueswell, 2003, among others). Fewer studies have investigated the question of whether clause type (question or declarative) can be anticipated by utilizing prosodic cues only, that is, whether the sentence unfolding to a question or a declarative can be predicted based on pure prosodic information. Among the limited studies on clause type anticipation, the authors examined the role of pitch accent in predicting the correct clause type (yes-no question and declarative) in Castilian Spanish (Face, 2004), the role of downstepped pitch (in the middle of the sentence) in predicting French yes-no questions (Vion & Colas, 2006), the contribution of pre-nuclear pitch accent properties in predicting yes-no questions and declaratives in Northern standard German (Petrone & Niebuhr, 2014), the intonation contours in anticipating clause types in English (e.g. question 'Want a candy?' and declarative 'Want a candy.') (Heeren, Bibyk, Gunlogson & Tanenhaus, 2015), and the prosodic cues (i.e. high F0

onset) in the pre-*wh*-word region in predicting in-situ *wh*-questions and declaratives in Persian.

For a tonal language like Mandarin Chinese, however, little is known about whether and when clause type anticipation takes place based on prosodic cues only. The consensus so far is that in an identical string of Mandarin yes-no questions[14] and declaratives, the sentence final syllable plays the most important role in identifying clause types as "question intonation has the highest prosodic strengths (high F0 curve) in the sentence final syllables" (Yuan, 2004, 2006). It is worth noting that Mandarin Chinese offers an ideal case of investigating clause type anticipation based on prosody, as we can easily find identical strings, not only yes-no questions and their string identical declarative counterparts, but also *wh*-questions and their declarative counterparts. We will spell out below why Mandarin *wh*-questions can be string identical to their declarative counterparts.

As introduced in previous chapters and repeated here, Mandarin is a *wh*-in-situ language in which *wh*-words remain at their base position just as their declarative counterparts do, as illustrated in (1a-b).

(1) a. 罗薇 昨天 买了 什么？ [*wh*-question]
Lúo Wēi zúotiān mǎi-le shénme?
Luo Wei yesterday buy-PERF what
'What did Luo Wei buy yesterday?'

b. 罗薇 昨天 买了 提子。 [declarative]
Lúo Wēi zúotiān mǎi-le tízi.
Luo Wei yesterday buy-PERF grapes
'Luo Wei yesterday bought grapes.'

Moreover, Mandarin *wh*-words like *shénme* are not only in-situ but also known as *wh*-indeterminates, which can have various interpretations (including a number of non-interrogative interpretations) depending on the context and licensors as in Japanese and Korean (Huang, 1982; Cheng, 1991; Li, 1992; Lin, 1998). When *diǎnr* 'a little' appears in front of a *wh*-word, it can have both an interrogative and a declarative interpretation, as illustrated in (2). (2a) is a declarative sentence and the *wh*-word *shénme* is interpreted as an indefinite, meaning 'something'. (2b) is its corresponding *wh*-question. (2a) and (2b) are string identical, but differ in their interpretations.

(2) a. 张三 买了 点儿 什么。 [*wh*-declarative]
Zhāng Sān mǎi-le diǎnr shénme.
Zhang San buy-PERF a.little SHENME
'Zhang San bought a little of something.'

[14] Mandarin yes-no question can have a sentence-final question particle *ma* but *ma* is optionally used, for example *Zhāngsān chī mǐfàn (ma)?* ('Does Zhangsan eat rice?'); when *ma* is not used, the yes-no question is a string identical to its declarative counterpart *Zhāngsān chī mǐfàn.* ('Zhangsan eats rice.').

b. 张三 买了 点儿 什么？ [*wh*-question]
Zhāng Sān mǎi-le diǎnr shénme?
Zhang San buy-PERF a.little SHENME
'What did Zhang San buy (a little of)?'

As reported in the production experiment in Chapter 2, the identical strings of *wh*-questions and *wh*-declaratives have different prosodic markings as early as in the clause initial (i.e. subject position). A question thus arises: can listeners differentiate and anticipate *wh*-questions and *wh*-declaratives by utilizing prosodic cues?

Though no studies have investigated the clause type anticipation of *wh*-declaratives and *wh*-questions so far, we take note of an audio-gating study on Mandarin *wh*-questions and their declarative counterparts with a non-interrogative noun phrase (Gryllia, Yang, Doetjes & Cheng, 2016). As shown in (1a-b), though *wh*-questions and their declarative counterparts are not entirely string identical, they are nonetheless identical up to the *wh*-word / noun phrase. To investigate whether listeners can anticipate clause types in the pre-*wh*-word region, Gryllia et al. (2016) conducted an audio-gating experiment on *wh*-questions and declaratives. They reported that in the beginning of the utterance such as the subject position, listeners can already have a preference towards the correct clause type that was intended by the speaker based on the prosodic cues. The results are insightful in that they indicate that clause type anticipation in Mandarin does take place early. Nevertheless, the question of whether the same anticipation can be found in the string identical case of *wh*-questions and *wh*-declaratives remains to be investigated.

In the current study we address the following questions: 1) Can listeners perceive the prosodic differences between *wh*-questions and *wh*-declaratives as in (2) and make use of them to differentiate the two clause types? 2) If yes, at which point can they perceive them and predict the clause types? In other words, how early does the clause type anticipation take place?

The chapter is organized as follows. In section 3.2, I present the results of a perception experiment that directly address question 1). Section 3.3 presents a series of audio-gating experiments based on *wh*-questions and *wh*-declaratives, the results of which directly address question 2). Section 3.4 is a summary and conclusion.

3.2 Perception experiment

In this section, we tried to tackle our first research question, namely, "Can listeners perceive the prosodic differences between *wh*-questions and *wh*-declaratives and make use of them to differentiate the two clause types?", by designing a perception experiment on the two types of *wh*-sentences. Like in a dialogue, after hearing each *wh*-sentence, participants were asked to continue the discourse by choosing one of two discourse continuations. Example (3) illustrates one type of *wh*-sentence (the one with *wh*-question prosody) which is auditorily presented and example (4) illustrates the two responses/discourse continuations: (4a) a noun phrase and (4b) a *wh*-question, which are visually presented on the screen. Given that (3) is a sentence with *wh*-question prosody, we predict that participants would choose (4a) as a response if they interpret it correctly as a *wh*-question. On the other hand, if the

audio stimulus consists of a sentence with *wh*-declarative prosody, we expect that participants would choose a *wh*-question in order to get more information to continue the discourse, if they correctly perceive the audio as a *wh*-declarative.

(3) 陶薇 昨天 拿了 点儿 什么 给 刘刚？ [*wh*-question]
TáoWēi zúotiān ná-le diǎnr shénme gěi LíuGāng?
TaoWei yesterday bring-PERF a.little what to LiuGang
'What did Tao Wei bring (a little) to Liu Gang yesterday?'

Discourse continuations on the screen

(4) a. 提子。
Tízi.
Grapes
Grapes.

b. 陶薇 拿了 什么？
TáoWēi ná-le shénme?
TaoWei bring-PERF what
'What did TaoWei bring?'

3.2.1 Participants

Thirty-six native speakers of Beijing Mandarin (16 female, 20 male, $\bar{x}$ age = 19 years old) participated in our experiment and were reimbursed for it. They were students at Tsinghua University coming from the northern part of China. None of them have participated in the production experiment reported in Chapter 2. They did not report any hearing or vision disorders (after correction). Prior to recording, informed written consent was obtained from each participant.

3.2.2 Acoustic stimuli

40 stimuli (20 items × 2 clause types) were selected from the recordings of a female native speaker of Beijing Mandarin (age = 20 years old), see the production study in Chapter 2 for details. These stimuli were chosen for their clearness and the speaker's moderate speech rate. Example (5) illustrates a sample stimulus which can be interpreted either as a *wh*-question or as a *wh*-declarative. Each stimulus consisted of 12 syllables and the stimulus length was constant across clause types and items. Below we reported the acoustic properties of the 40 audio stimuli in detail.

(5) a. 陶薇 昨天 拿了 点儿 什么 给 刘刚？ [*wh*-question]
TáoWēi zúotiān ná-le diǎnr shénme gěi LíuGāng?
T2 T1 T2 T1 T2-T0 T3 T2 T0 T3 T2 T1
TaoWei yesterday bring-PERF a.little what to LiuGang
'What did TaoWei bring (a little) to LiuGang yesterday?'

b. 陶薇 昨天 拿了 点儿 什么 给 刘刚. [*wh*-declarative]
TáoWēi zúotiān ná-le diǎnr shénme gěi LíuGāng.
T2 T1 T2 T1 T2-T0 T3 T2 T0 T3 T2 T1
TaoWei yesterday bring-PERF a.little something to LiuGang
'TaoWei brought a little something to LiuGang yesterday.'

Duration. Figure 1 presents the mean sentence duration in milliseconds (ms) of *wh*-questions and *wh*-declaratives. In general, *wh*-questions ($\bar{x}$ = 2049 ms) are shorter than *wh*-declaratives ($\bar{x}$ = 2197 ms). Figure 2 depicts the mean word duration for *wh*-questions and *wh*-declaratives. As shown in the same Figure, already from the subject, there is a consistent duration difference between the two clause types. *Wh*-questions are shorter than *wh*-declaratives at the subject, verb-*le*, and the preposition phrase (e.g., *gei Liu Gang* 'to/for Liu Gang') as well. A series of linear mixed effects models were run in R using the lmerTest package (Kuznetsova, Brockhoff & Christensen, 2013) with sentence/word duration as a dependent variable, clause type as a fixed-effect factor and item as a random factor. The detailed linear mixed effects results are summarized in Table 1.

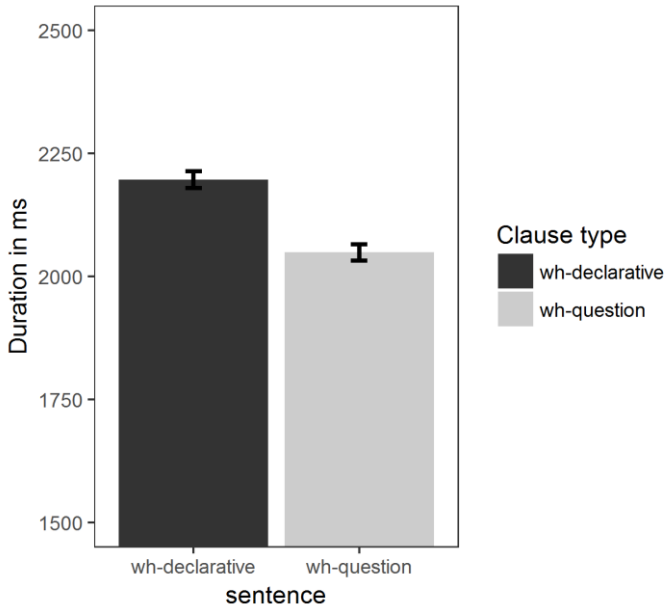


Figure 1. Mean sentence duration across clause types with error bar showing standard error.

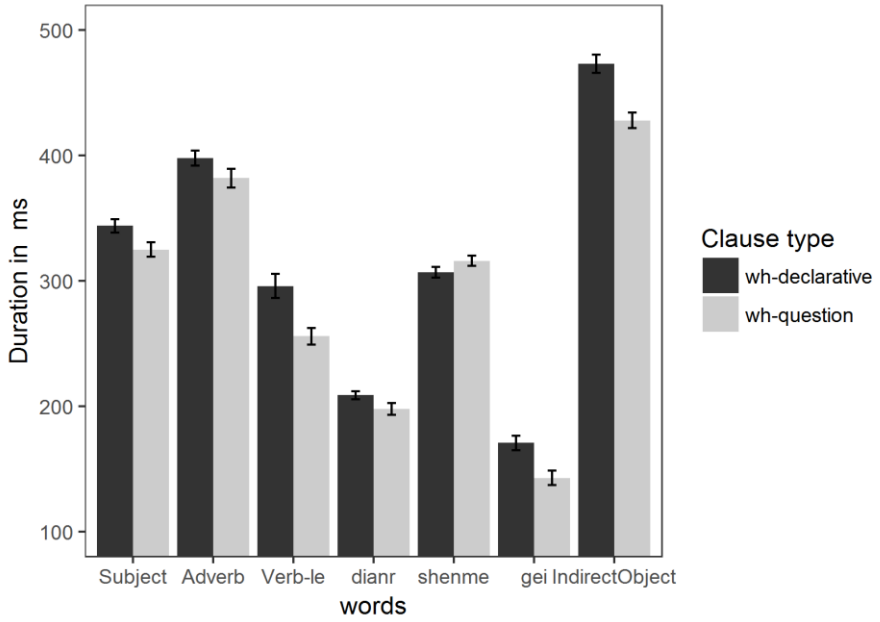


Figure 2. Mean word duration across clause types with error bar showing standard error.

Table 1. Summary of the linear mixed effects models on the duration of each word and the sentence.

	Estimate *β*	Std. Error	*t*-value	*p*-value
subject	18.45	5.36	3.44	< 0.01
verb-*le*	39.45	8.78	4.49	< 0.001
shénme	-9.30	5.97	-1.56	> 0.1
preposition phrase	73.00	9.13	8.00	< 0.001
the whole sentence	148.00	20.00	7.40	< 0.001

F0. As reported in Chapter 2 and repeated here, for the F0 measurement of syllables bearing Tone 1, we measured the F0-maximum (H), while for Tone 3 we measured the F0-minimum (L). For Tone 2 we measured first the F0-minimum and then the F0-maximum (LH), while for Tone 4 we measured first the F0-maximum and then the F0-minimum (HL). For Tone 0 (neutral tone) of the perfective marker *le*, following Li (2002), we measured first the F0-maximum and then the F0-minimum (HL), when the preceding syllable (the verb) bore Tone 1, Tone 2 or Tone 4, while we measured first the F0-minimum and then the F0-maximum (LH), when the preceding syllable bore Tone 3.

Based on the F0 measurement of each syllable (S), the stylized means of F0 curves for *wh*-questions and *wh*-declaratives split per verb tone was given in Figure 3. As we can see, the most striking F0 difference between the two clause types is at

the *wh*-word *shénme*, which shows a steep rise in *wh*-questions but is relatively flat in *wh*-declaratives, and the F0 in *wh*-questions remains higher than that in *wh*-declaratives until the end of sentence. Linear mixed effects model was also run in R with F0 in Hz as a dependent variable and the fixed-effect factor and random factor are the same as in duration models. The detailed linear mixed effects results are summarized in Table 2.

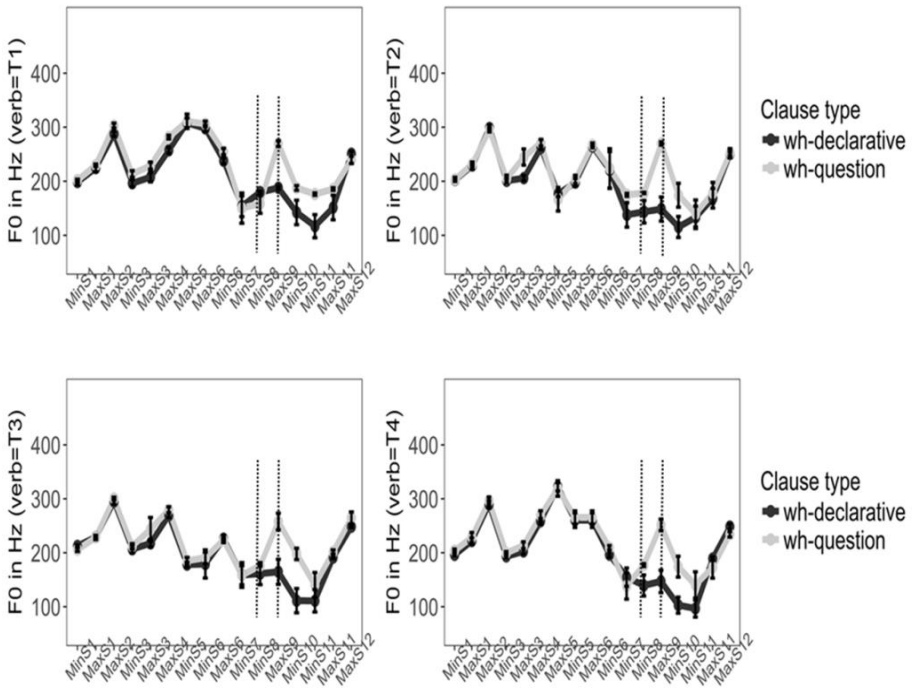


Figure 3. Stylized means of F0 curves across clause types and verb tones with error bar showing standard error; F0 of *shénme* highlighted with dashed lines.

Table 2. Summary of the linear mixed effects models on the F0 measurement between clause types (significant differences were reported).

	Estimate *β*	Std. Error	*t*-value	*p*-value
F0-min adverb (1st syllable)	8.66	2.67	3.24	< 0.01
F0-max adverb (1st syllalbe)	24.65	7.95	3.10	< 0.01
F0-max adverb (2st syllalbe)	14.16	3.88	3.65	< 0.01
F0-min *le*	8.46	3.21	2.64	< 0.05
F0-max *me*	100.60	10.15	9.91	< 0.001
F0-min *gěi*	65.27	11.24	5.81	< 0.001
F0-min indirect object (1st syllable)	34.61	14.23	2.43	< 0.05

3.2.3 Procedure

The perception experiment was conducted in a dim and sound-proof booth in a lab of Tsinghua University in Beijing. Participants were seated in front of a computer where the experiment was running with MFC Praat (Boersma & Weenik, 2016). The procedure was as follows. First, participants read the instructions that appeared on the computer screen and, once they were ready, they pressed OK on the screen to continue. After 1.0 second, the audio stimulus was played (either a *wh*-question or a *wh*-declarative as in (6a-b)). While the audio was played, the screen was blank. 0.3 seconds after the offset of the audio stimulus, two discourse continuations appeared on screen, namely, a noun phrase and a *wh*-question. Participants were instructed to listen to the audio stimulus and then complete the discourse selecting one of the two discourse continuations. After their selection, participants clicked the OK button and 1.0 second after clicking OK, the next audio stimulus was played. The audio stimuli were randomized for each participant to avoid a sequence effect; the two choices on the left or right of the screen were also counterbalanced to avoid any left/right preference among participants. Participants were not forced to make a choice under time pressure. The perception experiment lasted about 10 minutes.

3.2.4 Results

We obtained a total of 1440 responses (40 stimuli × 36 participants). The results showed that participants successfully perceived the clause type that was intended by the speaker and completed the dialogue with the corresponding response/dialogue continuation. On average, when the audio stimulus was a *wh*-question, they correctly perceived it as a *wh*-question and thus chose a noun phrase as response to complete the dialogue at 93.9% of the time. When the audio stimulus was a *wh*-declarative, they correctly perceived it and chose a *wh*-question as a continuation at 95.0% of the time (see Figure 4). The highest accuracy rate per participant was 100%, and the lowest accuracy rate was 70%. The chi-square analysis showed that there was a significant association between clause types intended by the speaker and listeners' responses, $x^2 = 1137.92$ (1), $p < 0.001$.

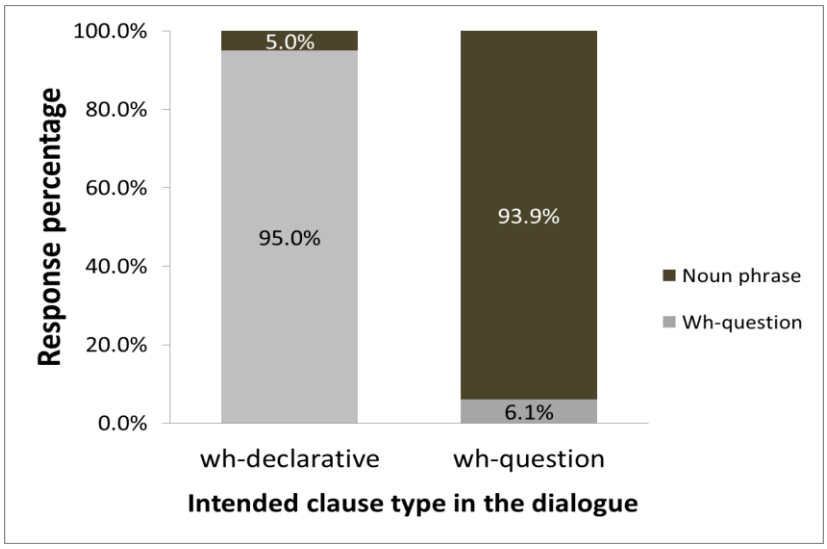


Figure 4. Listeners' responses in percentage (%) in the perception experiment.

3.2.5 Discussion

We conducted the perception experiment to investigate whether the clause types (*wh*-questions and *wh*-declaratives) can be differentiated by listeners. The results demonstrated that participants correctly made use of the prosodic differences between *wh*-questions and *wh*-declaratives to interpret the audio stimuli. When the audio stimulus is a *wh*-question, the participants' accuracy of interpretation is as high as 93.9% and when the audio stimulus is a *wh*-declarative, their accuracy is as high as 95.0%. In short, our experiment provides evidence that the prosodic differences in the string identical case of *wh*-questions and *wh*-declaratives can be perceived and the two clause types can be differentiated by listeners utilizing prosody.

3.3 Audio-gating experiment

As discussed in section 3.1, our second research question is twofold: "At which point can listeners perceive the differences and anticipate the clause types? In other words, how early can listeners perceive the differences?" To answer this research question, we used an audio-gating paradigm where participants listen to audio fragments and choose the corresponding sentence continuations presented on the screen. As we used the same group of participants for the audio-perception experiment and audio-gating experiment, the audio-gating experiment was actually run first to make sure that the information participants heard is incremental and therefore the audio-perception experiment would not affect the results of the audio-gating experiment.

3.3.1 Participants

The same participants from the audio-perception experiment participated in the audio-gating experiment.

3.3.2 Stimuli

The 40 audio stimuli used in the perception experiment (section 3.2.2) were used as a basis for generating audio fragments for the audio-gating experiment. We generated three types of audio fragments that were presented in three gates[15]. Type *a* audio fragment consisted of the subject and was presented in Gate *a*, Type *b* consisted of subject and adverb and was presented in Gate *b*, and Type *d* consisted of subject, adverb, verb plus *le* and *diănr* 'a little', and was presented in Gate *d*. Notice that we included *diănr* in Type *d* together with the verb because *diănr* is described as a clitic to the preceding verb (Shih, 1997; Chen, 2000). For each gate, we also created two kinds of sentence continuations, *wh*-questions and *wh*-declaratives. Examples (6-8) illustrate the fragments used in the audio fragments of three gates and their corresponding sentence continuations visually presented on the screen. As shown, the two kinds of continuations differ only at the sentence final punctuations.

(6) audio Gate *a* — sentence continuations (visually presented)

TáoWēi	-- zúotiān	ná-le	diănr	shénme	gěi	LíuGāng?
TaoWei	yesterday	bring-PERF	a.little	what	to	LiuGang
	-- zúotiān	ná-le	diănr	shénme	gěi	LíuGāng.
	yesterday	bring-PERF	a.little	something	to	LiuGang

(7) audio Gate *b*

TáoWēi zúotiān	-- ná-le	diănr	shénme	gěi	LíuGāng?
TaoWei yesterday	bring-PERF	a.little	what	to	LiuGang
	-- ná-le	diănr	shénme	gěi	LíuGāng.
	bring-PERF	a.little	something	to	LiuGang

(8) audio Gate *d*

TáoWēi zúotiān	ná-le	diănr	-- shénme	gěi	LíuGāng?
TaoWei yesterday	bring- PERF	a.little	what	to	LiuGang
			-- shénme	gěi	LíuGāng.
			Something	to	LiuGang

This resulted in a total of 40 audio fragments of each gate and 120 in total. The detailed acoustic properties can be found in section 3.2.2.

[15] The audio-gating experiment also includes another gate, gate *c*, which is designed for different research questions and is reported in Chapter 4.

3.3.3 Procedure

Similar to the perception experiment, participants were seated in front of a computer where the experiment was running with MFC Praat (Boersma & Weenink, 2016). In each gate, the procedure was similar to that in section 3.2.3. First, participants read the instructions that appeared on the computer screen and, once they were ready, they pressed OK on the screen to continue. After 2.0 seconds the audio fragment was played. While the audio was played, the screen was empty. 0.5 seconds after the offset of the audio stimulus, two sentence continuations appeared on the screen, one on the left, the other on the right. Participants were asked to listen to each audio fragment in that gate and then complete the sentence selecting one of the two continuations. After choosing one option on the screen, participants clicked the OK button and 2.0 seconds after clicking OK, the next audio stimulus was played. The order of the two sentence continuations on the screen was counterbalanced to avoid any left/right preference among participants.

The stimuli in each gate were randomized for each participant to avoid a sequence effect. Participants were not forced to make a choice under time pressure. The audio fragments were presented in three consecutive gates, from gate *a* to gate *d*, to make sure that the information participants heard was incremental.

3.3.4 Results

We obtained a total of 4320 responses (3 gates × 40 stimuli × 36 participants). In general, participants were successful in correctly deciding which of the two clause types were intended by the speaker, as illustrated in Figure 5. In gate *a*, where participants only listened to the subject, participants chose a question continuation 54.6% of the time when it was originated from questions. In other words, their overall response accuracy was 54.6% when the intended clause type was question; when the intended clause type is declarative, their overall response accuracy was 59%. The chi-square analysis showed that there was a significant association between the clause type intended by the speaker and the participants' responses, $x^2 = 26.73$ (1), $p < 0.001$. In gate *b*, where participants listened to both the subject and the adverb, the overall response accuracy was 59.7% when the intended clause type was question, and 64.6% when the intended clause type was declarative. Similarly, the chi-square analysis showed that there was a significant association between the clause type intended by the speaker and the listeners' responses, $x^2 = 85.27$ (1), $p < 0.001$.

In gate *d*, where participants listened to subject, adverb and verb-*le diănr*, the overall response accuracy was 62.1% when the intended clause type was question, and 72.1% when the intended clause type was declarative. The chi-square analysis showed that there was a significant association between the clause type intended by the speaker and participants' responses, $x^2 = 169.80$ (1), $p < 0.001$.

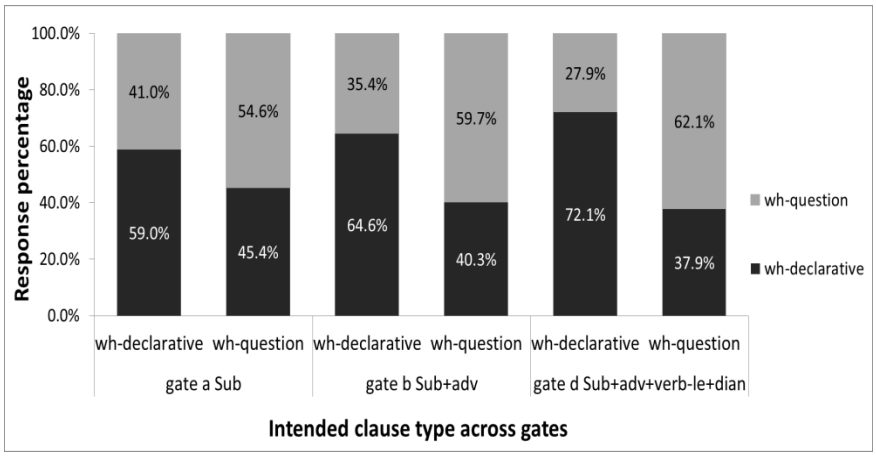


Figure 5. Listener's response in percentage (%) in gate *a*, *b* and *d*

To investigate whether participants' responses can be predicted on the basis of the intended clause type by the speaker in each gate, we also ran a mixed effects logistic regression using the lme4 package in R, with the intended clause type (declarative or a question) as an independent variable, participants' responses (declarative or question) as a dependent variable and items and participants as random factors. Specially, we first ran a null model with participants' responses as a dependent variable, and participants and items as random factors. A second model included in addition the intended clause type as a fixed effect factor to see whether the model was improved. Finally, we ran a third model that included participants' responses as a dependent variable, the intended clause type as a fixed-effect factor, and participants and items as random factors, allowing by-participant and by-item random intercepts, and by-participant and by-item random slopes for the intended clause type. Model fit was compared using the likelihood ratio test (Pinheiro & Bates, 2000; Bolker, Brooks, Clark, Geange, Poulsen, Stevens, & White, 2009). See Appendix B for the details of the fitting model in each gate. The detailed mixed effects results are summarized in Table 3. In general, participants' response on the clause type in each gate can be predicted on the basis of the intended clause type by the speaker.

Table 3. Summary of the results of the mixed effects logistic regression between participants' responses and the intended clause type in each gate.

	Estimate β	Std. Error	z-value	p-value
Gate *a*	0.62	0.30	2.10	< 0.05
Gate *b*	1.51	0.29	4.00	< 0.001
Gate *d*	1.58	0.17	9.35	< 0.001

3.3.5 Discussion

The audio-gating experiment aims at investigating at which point the prosodic differences between *wh*-questions and *wh*-declaratives can be perceived and utilized to predict the sentence clause type. The results from gate *a*, *b* and *d* demonstrate that listeners can make use of prosody to anticipate clause types, distinguishing *wh*-questions and *wh*-declaratives before reaching the *wh*-word *shénme*. In gate *a*, where only the subject of the sentence is heard, there is already a preference for the correct clause type. As the gate goes from *a* to *b* and then to *d*, the more information is perceived by participants, the more accurate their responses are.

With respect to the clause type anticipation, when a sentence unfolds over time and listeners can interpret the information and formulate a prognosis of which clause type the sentence may further unfold, we can hence claim that they can anticipate the clause type. By adopting the audio-gating paradigm, we distinctly collected each prognosis of the clause type in different sentence fragment lengths. The results of the audio-gating study demonstrate that, although *wh*-questions and *wh*-declaratives are string identical, listeners can anticipate their clause types as early as the beginning of the sentences by utilizing the available prosodic cues.

3.4 General discussion and conclusion

The aim of this study was to examine whether and when Mandarin listeners can anticipate the clause type based on pure prosodic cues, by conducting a case study on *wh*-questions and *wh*-declaratives. To achieve this, we first conducted a perception experiment on the identical string of *wh*-questions and *wh*-declaratives to investigate whether listeners can identify the clause type intended by the speaker by perceiving the prosodic differences. The results of the perception experiment served as an experimental baseline for clause type anticipation, since they demonstrated that listeners can indicate the clause type correctly by hearing the prosodic cues of the clause type.

The audio-gating experiment directly addresses the questions of whether listeners can anticipate clause types and when exactly the clause type anticipation takes place based on prosody, by testing different lengths of audio fragments of *wh*-questions and *wh*-declaratives. We obtained listeners' prognosis of the clause type in all the audio-gates and found that listeners can make use of the (limited) prosodic cues to predict the clause type, already from the subject of the sentence. It should be noted that we didn't provide the listeners any prior contexts other than the fragments of *wh*-questions and *wh*-declaratives; in other words, prosodic information is able to serve as the only cue for clause type anticipation in a sentence appearing out of the blue. Further, the more prosodic information is used by listeners, the more accurate their anticipation on clause type is.

With respect to clause type and prosody, our findings also support that, if clause types are marked differently in their specific prosody in a language, these specific prosodic markings (i.e. pitch and duration in our study) are stored in auditory memory and actively used as criteria/filters for listeners' identification of clause

types (Gerard & Clement, 1998). Lastly, concerning anticipation, our findings also imply that listeners seem to make an online assessment of the limited prosodic cues they hear with a reference to the common prosodic marking in each clause type stored in memory and anticipate the most plausible clause type or interpretation accordingly based on the existing prosody.

Although the audio-perception and gating studies can help us to determine whether and at which point prosody is utilized by listeners to identify and anticipate clause types, they are offline studies that cannot provide evidence for the role of prosody on clausal typing during online processing. This is the limitation of the gating paradigm. Different from the behavioral studies like audio-gating, event-related potentials (ERPs) provide a continuous measure of auditory processing with an excellent temporal resolution, and hence serve as an ideal measure for revealing the role of prosody in clausal typing during online processing. We continue this topic of clausal typing by reporting 2 ERP studies in Chapter 6. In the next Chapter, we will first discuss the licensing of *wh*-existentials as the evidence collected in the audio-gating study can illuminate the discussion of the licensing of *wh*-existentials to some extent.

Chapter 4 On the Role of *Diǎnr* in Licensing Mandarin *Wh*-existentials

4.1 Introduction

As we have seen in previous chapters, Mandarin *wh*-words (e.g. *shénme*) can have existential interpretation ('something'), subject to licensing (Huang, 1982; Cheng, 1991; Li, 1992; Lin, 1998, Lin, Weerman & Zeijlstra, 2014). Many of the examples of *wh*-declaratives with *wh*-existentials we introduced also contain the word *diǎnr* 'a little', which is considered to be a determiner with existential quantificational force (Tsai, 2010). In this chapter, we are particularly interested in the licensing of the existential reading of the *wh*-word *shénme* ('something') and the role of *diǎnr* as a potential licensor of *wh*-existentials.

The existential reading of *shénme* has been shown to be licensed by nonveridical contexts. For example, environments containing nonveridical operators like negation, questions, conditionals and epistemic modalities (see Li, 1992; Lin, 1998; Huang, 2017, for details). The definition of (non)veridicality is given in (1).

(1) (Non)veridicality for propositional operators:
A propositional operator F is veridical, iff Fp entails p: Fp ⊨ p; otherwise F is nonveridical.

[Zwarts, 1995]

Essentially, contexts exhibiting nonveridicality are normally contexts where the truth of a proposition cannot be entailed. As illustrated below, none of the examples in (2) entail the truth of 'Zhang San buys something.'

(2) a. 张三 没有 买 什么。 [negation]
Zhāng Sān méiyǒu mǎi shénme.
Zhang San didn't buy SHENME
'Zhang San didn't buy anything.'

b. 张三 买了 什么 吗? [yes-no question]
Zhāng Sān mǎi-le shénme ma?
Zhang San buy-PERF SHENME yes-no particle
'Did Zhang San buy anything?'

c. 如果 张三 买了 什么，别 生气。 [conditionals]
Rúguǒ Zhāng Sān mǎi-le shénme, bié shēngqì.
if Zhang San buy-PERF SHENME don't angry
'If Zhang San buys something, please don't be angry at him.'

d. 在 张三 买 什么 之前， 记得 讲价。 [before-clause]
Zài Zhāng Sān mǎi shénme zhīqián, jìdé jiǎngjià.
at Zhang San buy SHENME before remember bargain
'Remember to bargain about price before Zhang San buys something.'

e. 可能 张三 要去 买 什么。 [epistemic modality]
Kěnéng Zhāng Sān yàoqù mǎi shénme.
possibly Zhang San go buy SHENME
'Possibly, Zhang San goes to buy something. '

All the above contexts exhibit nonveridicality and they can license the existential reading of *wh*-words. Furthermore, it is noticed that some types of nonveridical contexts (i.e. some future environments, see Lin, 1998 for details) cannot license the existential reading of *wh*-words unless the *wh*-word is preceded by *diǎnr* 'a little' or by the classifier *gè*. *Diǎnr* has a meaning of small quantity, according to *Xiàndài hànyǔ cídiǎn* (Contemporary Chinese Dictionary, 2005: p 304). *Gè* is a "general" classifier usually used with nouns without a specific classifier; in some conditions, it can also be used with nouns that typically appear with a specific classifier (see *Xiàndài hànyǔ cídiǎn* (Contemporary Chinese Dictionary), 2005: p 462). As illustrated in (3a), the sentence without *diǎnr/gè* is interpreted as a *wh*-question (*wh*-word licensed by (Q) by default) but the declarative reading is not easily available, whereas adding *diǎnr/gè* in (3b) makes both the interrogative and the existential interpretations available.

(3) a. 张三 打算 买 什么 [future]
Zhāng Sān dǎsuàn mǎi shénme
Zhang San plan buy SHENME
i. 'What does Zhang San plan to buy?' (*wh*-question)
ii. * 'Zhang San plans to buy something.' (*wh*-declarative)

b. 张三 打算 买 点儿/个 什么 [future]
Zhāng Sān dǎsuàn mǎi diǎnr/gè shénme
Zhang San plan buy a.little/CL SHENME
i. 'What does Zhang San plan to buy (a little)?' (*wh*-question)
ii. 'Zhang San plans to buy (a little) something.' (*wh*-declarative)

The contrast between (3a) and (3b) illustrates that some nonveridical contexts (i.e. future environments) still require *diǎnr* or *gè* in order to license existential readings of *wh*-words. On the other hand, is it possible that veridical contexts with the presence of *diǎnr* or *gè* can also allow the licensing of the existential readings of *wh*-words? Previous studies (Xie, 2007; Lin et al., 2014; Huang, 2017, among others) claim that the answer is no as they emphasize that *wh*-existentials are restricted to nonveridical contexts only. Nonetheless, we notice that there are also contexts which are veridical but still can allow the existential interpretation of *wh*-words and their default interrogative interpretation is still available. An example of such context is example (4), which contains veridical sentences with *wh*-words preceded by *diǎnr* or

gè and is ambiguous between interrogative and declarative interpretations. When *shénme* is interpreted as an interrogative, the clauses are information-seeking questions. On the other hand, when *shénme* is interpreted as an existential, the clauses are declarative sentences.

(4) a. 张三 昨天 买了 点儿 什么
Zhāng Sān zuótiān mǎi-le diǎnr shénme
Zhang San yesterday buy-PERF a.little SHENME
i. 'What did Zhang San buy a little of yesterday?' (*wh*-question)
ii. 'Zhang San bought a little of something yesterday.' (*wh*-declarative)

b. 张三 昨天 买了 个 什么
Zhāng Sān zuótiān mǎi-le gè shénme
Zhang San yesterday buy-PERF CL SHENME
i. 'What did Zhang San buy yesterday?' (*wh*-question)
ii. 'Zhang San bought something yesterday.' (*wh*-declarative)

The availability of both interpretations for the sentences in (4) seems to indicate that the existential reading of *wh*-words can also be licensed in veridical contexts, countering previous proposals on the licensing conditions of *wh*-existentials (Lin, 1998; Xie, 2007; Lin et al., 2014; Huang, 2017). Also, the role of *diănr* or *gè* in the licensing of *wh*-existentials in some nonveridical contexts as in (3) and in veridical contexts as in (4) poses the question of how *diănr* or *gè* can function to license the existential reading of *wh*-words. We mainly focus on *wh*-declaratives containing *diănr* in the present study. To answer the above question, we first review the studies which discuss *wh*-declaratives containing *diănr* (section 4.2); then we confirm that *wh*-existential reading is indeed licensed by *diănr* through a reading study (section 4.3) and through an audio-gating experiment (section 4.4); finally we discuss the licensing environments of *diănr* on *wh*-existentials in Mandarin (section 4.5).

4.2 What is *diănr* and discussions about effects of *diănr* on *wh*-existentials

Before we discuss the effect of *diănr* on *wh*-existentials, we first introduce the basic meaning and usage of *diănr*. *Diănr*, also used as *yi-diănr* [16], modifies the noun phrase following it. As illustrated in example (5), (*yi*)-*diănr* denotes the meaning of 'a small quantity' and modifies the noun phrase 'gold'.

(5) 张三 昨天 买了 (一)点儿 金子
Zhāng Sān zuótiān mǎi-le (*yi*)-diănr jīnzi
Zhang San yesterday buy-PERF one-a.little gold
'Zhang San bought a little bit of gold yesterday.'

[16] *Yi* has a literal meaning of 'one'. The presence of *yi* is optional and both *diănr* and *yi-diănr* have the interpretation of 'a little'.

In the above example, the noun modified by (*yi*)-*diănr* is a mass noun. Iljic (1994) and Lü (1999) claim that (*yi*)-*diănr* is usually used with mass or abstract nouns, while Zhang (2013) claims that (*yi*)-*diănr* has little selection restriction on its noun phrase and can also be used to modify a countable noun under some conditions, resulting in the interpretation of 'a few'. We are not particularly interested in under what conditions (*yi*)-*diănr* can also be used to modify a countable noun. Instead, what we are concerned with is the meaning of 'a small quantity' that (*yi*)-*diănr* denotes when it modifies a noun phrase.

From a different perspective, Shih (1997) and Chen (2000) describe *diănr* as a bound morpheme that cliticizes obligatorily leftwards. When *yi* is used in *yi-diănr*, *diănr* cliticizes to *yi* prosodically. On the other hand, when *yi* does not appear as in examples (3b) and (4a), *diănr* cliticizes leftwards onto the verb-(PERF) prosodically, although syntactically it belongs to the DP phrase.

In addition to modifying the ordinary noun phrase and denoting the meaning of 'a small quantity', *diănr* can also be used with a *wh*-word, be it a *wh*-interrogative as in (6a) or a *wh*-existential as in (6b).

(6) 张三 昨天 买了 点儿 什么
Zhāng Sān zuótiān mǎi-le diǎnr shénme
Zhang San yesterday buy-PERF a.little SHENME
a. 'What did Zhang San buy a little of yesterday?' (*wh*-question)
b. 'Zhang San bought a little bit of something yesterday.' (*wh*-declarative)

When *diănr* is used with a *wh*-word as in (6), the mass noun and countable noun distinction does not play a role any more, as no matter it is milk or two books that Zhang San bought, the *wh*-question and *wh*-declarative in (6) are always valid. It is worth noting that the above examples in (6) will be infelicitous if the speaker already knows that the stuff bought by Zhang San is something big in quantity/value/significance. This meaning of 'small in value/significance' is not obtained when *diănr* is modifying a non-*wh* noun phrase, as in *(yi)-diănr jīnzi* ('a little bit of gold', which is valuable). Only in combination with a *wh*-word is the meaning of 'small in value/significance' triggered in addition to 'small in quantity'. It remains to be investigated why this extra meaning of 'small in value/significance' is obtained when *diănr* is used with a *wh*-word. Summarizing briefly, *diănr* is a quantification-denoting expression that denotes the meaning of 'small in quantity' when used with an ordinary noun phrase and a *wh*-word; it adds the qualification of 'small in value/significance' when used with a *wh*-word. Also, *diănr* cliticizes leftwards onto the verb-(PERF) when *yì* is not used.

Consider now the contribution of *diănr* in *wh*-existential interpretation, regardless of whether we are dealing with a veridical or a nonveridical context. It is worth noting that Tsai (2010) analyzes *diănr* as an element with existential quantificational force. Also, Zhang (2013) describes that *diănr* functions as an existential quantifier. In other words, they both consider *diănr* to be the one providing existential force to the *wh*-word, and therefore licensing the *wh*-existential interpretation.

With respect to *diǎnr*'s licensing of *wh*-existential interpretation, as *diǎnr* 'a little' can be used together with the numeral *yì* 'one' optionally, a natural question arises: since *yìdiǎnr* denotes the meaning of 'a little', is it possible that *diǎnr* in Mandarin has the negative interpretation 'little' when *yì* is not present, so that the negative environment licenses the *wh*-existential as in a normal nonveridical context? We think that the answer is no: *diǎnr* has the same interpretation as *yìdiǎnr*, interpreted as 'a little' as *yì* in *yìdiǎnr* can be optionally present. *Diǎnr* is not a negative expression that introduces a nonveridical context such as *little* in English, *weinig* in Dutch, *wenig* in German or *peu* in French. We can illustrate the different interpretations of Mandarin *diǎnr* and French *peu* utilizing a test example proposed by Ducrot (1972: 193).

(7) a. Il semble devenir sobre: il a bu peu de vin hier.
he seems become sober he has drunk little of wine yesterday
'He is sober: he has drunk little wine yesterday.'

b. Il semble devenir moins sobre: *il a bu peu de vin hier.
he seems become less sober he has drunk little of wine yesterday
*'He is not sober: he has drunk little wine yesterday.'

[Ducrot, 1972]

(8) a. 他 似乎 很 清醒： *他 昨天 喝了 点儿 酒。
Tā sìhū hěn qīngxǐng: *tā zuótiān hē-le diǎnr jiǔ.
he seem very sober: he yesterday drink-PERF a.little wine
*'He is sober: he drank a little wine yesterday.'

b. 他 似乎 不太 清醒： 他 昨天 喝了 点儿 酒。
Tā sìhū bútài qīngxǐng: tā zuótiān hē-le diǎnr jiǔ.
he seem not sober: he yesterday drink-PERF a.little wine
'He is not very sober: he drank a little wine yesterday.'

As illustrated in (7), French *peu* is a negative expression and therefore the sequence of statements in (7a) is felicitous, while (7b) is not. In contrast, in the case of Mandarin *diǎnr*, (8b) is falicitous while (8a) is not, since *diǎnr* only has the positive interpretation 'a little'. Based on this, we can ascertain that *diǎnr* is not a negative expression and hence the licensing of *wh*-existentials in (4a) or (6b) with *diǎnr* cannot be accounted for by assuming any type of nonveridicality, as the sentence in (4a) or (6b) is simply a veridical context.

Besides the investigation of *diǎnr* itself, there are a limited number of studies on *wh*-indeterminates that discuss the role of *diǎnr*. Nevertheless, they either only notice the mysterious role of *diǎnr* without giving an account (Lin, 1998), or propose that classifiers provide existential force for the *wh*-words but excluding veridical cases like (4a) or (6b) (Huang, 2017). Below we discuss these studies in detail.

Lin (1998) generalizes three groups of environments/contexts that license the existential interpretation of *wh*-words: group A, consisting of Negation, questions

and *if*-clauses (our example (2a-c) belong to this group); group B, corresponding to epistemic modality contexts (example (2e) belongs to this group); and group C, some sort of "future" environments (example (3) belongs to this group). The three groups of contexts classified by Lin (1998) are all nonveridical. According to Lin, the three groups of contexts in his study decrease in their strength for licensing polarity *wh*-words[17] while the need for a classifier like *diǎnr*[18] (or *gè*) increases. For example, Lin (1998: 222) summarizes that in group B adding *diǎnr* (or *gè*) "may (sometimes) increase naturalness but this is not essential"; but in group C *diǎnr* (or *gè*) seems to be generally required for the licensing of the existential reading, similar to our observation in example (3).

Lin (1998) notices the role of *diǎnr* (or *gè*) and finds that *diǎnr* (or *gè*) increases the naturalness of the *wh*-existential reading when the context strength for licensing the *wh*-existential is decreasing. As the licensing contexts listed in Lin (1998) (negation, question, if-clause, epistemic modality and future environment) can be understood as licensors, Lin's observation can be further generalized as the following: as the licensors' licensing strength decreases, the necessity of *diǎnr* (or *gè*) increases in the licensing of the existential reading of *wh*-words. However, Lin (1998) only briefly introduces the naturalness that *diǎnr* (or *gè*) brings about in context group B and C without discussing how *diǎnr* (or *gè*) helps in the licensing of *wh*-existentials. Furthermore, all the contexts discussed by Lin (1998) are restricted to nonveridical contexts only and the possibility that *wh*-existentials can be licensed in veridical contexts containing *diǎnr* (or *gè*) as in (4) was not considered.

A more recent study is Huang (2017), who systematically investigates the role of classifiers[19] like *gè* or *diǎnr* in licensing the *wh*-existential reading in Mandarin. First, he proposes that Mandarin *wh*-words require two ingredients to be existential, nonveridicality (Zwarts, 1995; Xie, 2007; Lin et al, 2014) and an existential closure (∃) (Kratzer & Shimoyama, 2002) that applies higher than vP. The existential closure provides the existential force to the *wh*-word when the existential closure is in the scope of the nonveridical operator. Furthermore, Huang proposes that existential closure can only apply higher than vP; if the nonveridical operator in the sentence is not higher than vP, the existential closure is hence not in the scope of the nonveridical operator and accordingly the existential closure cannot provide the existential force to the *wh*-word. In this case, classifiers like *diǎnr* or *gè* play the role of providing existential force to the *wh*-word. Take the *wh*-existential reading in (9a,b) as an example. (9a) has an epistemic modal word *hǎoxiàng* (roughly translated as 'seem') while (9b) has a deontic modal[20] word *yīnggāi* 'should'. Both sentences satisfy the nonveridicality condition but the crucial difference is that the two modals/nonveridical operators have different syntactic positions, which affect the application of existential closure. (9a) is grammatical without a classifier

[17] *Wh*-words in Mandarin behave like negative polarity items, as stated in Cheng (1991, 1994), among others.

[18] Both Lin (1998) and Huang (2017) describe *diǎnr* as a classifier, although Tsai (2010) holds a different opinion, as introduced in the beginning of section 2.

[19] See reference 6.

[20] According to Lin (1998) and Huang (2017), sentences with deontic modality can be grouped into future environments.

because the epistemic modal which licenses the *wh*-existential is generated above vP and hence in the scope of the epistemic modal there is applicable existential closure. (9b) is ungrammatical without a classifier like *diǎnr* or *gè* because the deontic modal is generated lower than vP and hence has a scope lower than vP; the existential closure that applies higher than vP therefore falls outside the scope the deontic modal, but the sentence can be rescued by adding *diǎnr* or *gè*, which plays the role of providing existential force to the *wh*-word.

(9) a. 张三 好像 吃了 什么。
Zhāng Sān hǎoxiàng chī-le shénme.
Zhang San seem eat-PERF something
'It seems that Zhang San ate something.'

b. 张三 应该 吃了*(点儿) 什么
Zhāng Sān yīnggāi chī *(diǎnr) shénme.
Zhang San should eat a.little something
'Zhang San should eat something.'

[Huang, 2017]

To our knowledge, Huang's study is the first one that gives an analysis on why licensing *wh*-existentials sometimes needs classifiers like *gè* or *diǎnr* while sometimes not. We agree with Huang that *diǎnr* or *gè* provides existential force for the *wh*-word, but there are a couple of issues in Huang's analysis that require further investigations.

The first issue concerns *diǎnr* as a classifier (in accounts such as Huang, 2017). According to Cheng and Sybesma (1999), which follows Allan (1977), Tai and Wang (1990) and Tai (1992), Mandarin classifiers can be roughly divided into two types, (a) classifiers which combine with count-nouns, denoting some inherent and permanent properties of the object/noun (e.g., *yì-běn shū* [one-CL book] 'a book'), and (b) classifiers that create a unit of measure indicating temporary states of the object/noun, giving a quantifying description of the object (e.g., *yì-píng shuǐ* [one-CL water] 'a bottle of water'). Clearly, *diǎnr* does not denote any inherent and permanent properties of an object like the first type of classifiers. Similar to the latter type of classifiers, *diǎnr* does denote the quantifying description of the object ('a little') and somehow indicates a temporary state of the object. But as opposed to the latter type of classifiers or any type of classifiers, *diǎnr* has a different syntactic distribution. For instance, quantifiers and numerals can co-occur with classifiers, as in *hěnduō píng shuǐ* [many-CL water] 'many bottles of water' and *wǔ-píng shuǐ* [five-CL water] 'five bottles of water', but they cannot co-occur with *diǎnr*, as **hěnduō diǎnr shuǐ* (many DIANR water) and **wǔ-diǎnr shuǐ* (five-DIANR water) are not grammatical. Hence *diǎnr* cannot be considered as a standard type of classifier.

The second issue, which is a key issue, concerns the nonveridicality. According to Huang (2017), nonveridicality is always needed in licensing the *wh*-existential. Only in cases when a nonveridical operator has a syntactic position that is not above vP and hence the existential closure is not available, does *gè* or *diǎnr* start to play a role of existential closure. This line of analysis proposed by Huang (2017) rules out

the non-interrogative readings in a veridical context as in (4) or (6b), contrary to our observation.

In short, it is evident that *diǎnr* is connected to the licensing of *wh*-existentials in addition to 'nonveridicality' by providing existential force for the *wh*-word; nevertheless, previous studies either neglect (Lin, 1998) or deny (Huang, 2017) the possibility that *wh*-existentials can be licensed in veridical contexts containing *diǎnr* as in (4) and (6). As we can see, all the claims so far about whether *wh*-declaratives can be licensed by *diǎnr* in veridical contexts is based on observations. Opting for empirical support, we conduct two empirical studies, a sentence reading study on *wh*-sentences containing *diǎnr* and an audio-gating study on *wh*-sentences containing *diǎnr* respectively.

4.3 A reading study testifying the two interpretations

4.3.1 Participants

Eighty-four native speakers of Beijing Mandarin (53 females and 31 males, $\bar{x}$ age = 20 years old) participated in this study and were reimbursed for their participation. All of them came from the northern part of Mainland China and were students at Tsinghua University. Prior to testing informed written consent was obtained from each participant.

4.3.2 Experimental materials

We created a total of 40 stimuli. Half of the stimuli were the control stimuli that could be interpreted only as *wh*-questions, which involve typical veridical contexts, as in (10b). The other half of the stimuli were target stimuli with the *wh*-word preceded by *diǎnr*, which also involve veridical contexts, see (10a). As shown in (10), target and control stimuli were string identical except for the licensor *diǎnr* 'a little', which was present in the target stimuli but absent in the control.

(10) a. Target stimuli

冯涛	昨天	做了	点儿	什么	给	王英
FéngTāo	zúotiān	zùo-le	diǎnr	shénme	gěi	WángYīng
FengTao	yesterday	make-PERF	a.little	SHENME	for	WangYing

'What did FengTao make for WangYing yesterday?' (*wh*-question)
or 'FengTao yesterday made a little something for WangYing.' (*wh*-declarative)

b. Control stimuli

冯涛	昨天	做了	什么	给	王英
FéngTāo	zúotiān	zùo-le	shénme	gěi	WángYīng
FengTao	yesterday	make-PERF	what	for	WangYing

'What did FengTao make for WangYing yesterday?'

Target and control stimuli were intermingled with 120 fillers. The fillers consisted of 20 yes-no questions with *ma* "yes-no particle", 20 A-not-A questions (a type of

yes-no question in Mandarin) and 80 declaratives. All stimuli lacked punctuation and were randomized for each participant.

4.3.3 Procedure

Participants were tested individually in a quiet room in Tsinghua University in Beijing. The whole experiment lasted about 15 minutes. The participants' task was to read silently the stimulus on screen, and to choose one of the two punctuation marks (i. question mark (?) or ii. full stop (。)) to complete the sentence. The experiment was organized as a semi-self-paced procedure and participants were under time pressure (maximum of 5 seconds for reading silently the stimulus, and maximum of 5 seconds for indicating their response). In particular, the procedure was as follows. First, the stimulus appeared on the computer screen, and participants had a maximum of 5 seconds to read it silently. After participants read the stimulus, they press the space bar and move to the next screen to indicate their response. If participants exceeded the 5 seconds time limit, the next screen appeared automatically. This next screen showed two punctuation marks, namely, a question mark (?) and a full stop (。) and participants had again 5 seconds to indicate their response. If participants exceeded the time limit of 5 seconds, the next stimulus appeared automatically on screen. The question mark and the full stop were counter-balanced on screen to avoid any left or right preference by participants. The experiment was conducted with the experimental software E-prime 2.0 (Psychology Software Tools).

4.3.4 Statistical analysis and results

We first inspected the data for missing responses; we had a total of 10 missing responses (7 for target stimuli and 3 for control stimuli). We obtained a total of 3350 responses (1673 responses for target stimuli and 1677 responses for control stimuli). These responses were analyzed statistically. We also examined participants' stimuli reading time and response time respectively.

Response. As illustrated in Figure 1, participants interpreted the control stimuli (*wh*-sentences without *diănr*) as questions 93.1% of the time. When inspecting participants' responses to the target stimuli (*wh*-sentences with *diănr*), we see a more balanced response distribution; participants interpreted *wh*-sentences with *diănr* as questions 59.7% of the time, while they interpreted them as declaratives 40.3% of the time. The chi-square analysis showed that there was a significant association between the stimulus type (target or control) and the participants' responses, $x^2 = 518.91$ (1), $p < 0.001$.

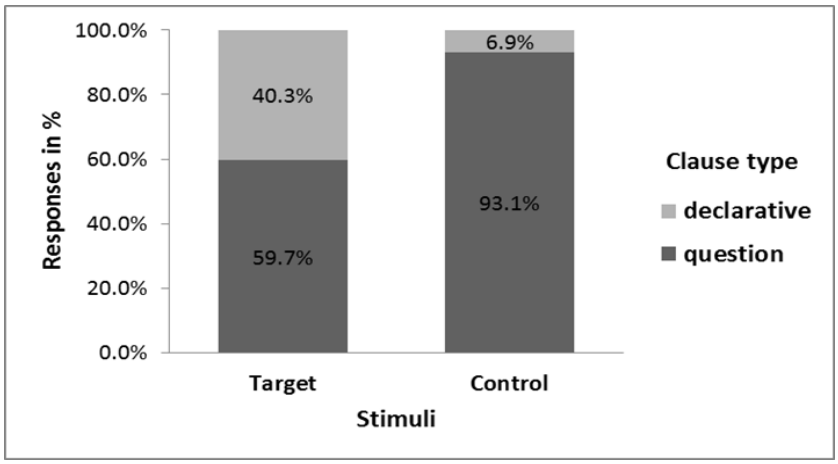


Figure 1. Participants' responses in percentage (%) in the reading study

We also ran a logistic regression analysis in R to investigate whether participants' responses can be predicted on the basis of the stimulus' type by using the stimulus' type (control or target) as an independent variable and participants' responses as a dependent variable. A test of the full model against a constant only model was statistically significant, as shown by the results presented in Table 1, indicating that stimulus' type (control or target) is a reliable predictor for participants' response as in (10a) and (10b).

Table 1. Logistic regression of stimulus' type (Target/Control) on participants' response of the clause type.

		95% CI for Odds Ratio		
	B (*SE*)	Lower	Odds Ratio	Upper
Constant	0.39(0.05)			
Stimulus' Type	2.21*(0.11)	7.36	9.10	11.26

Note: *$p < 0.001$.

Stimuli reading time and response time. We also inspected the time participants needed to read the stimuli and to indicate the clause type. As shown in Figures 2 and 3, participants spent more time reading the target stimuli ($\bar{x}$ = 2364.5ms) and indicating their responses ($\bar{x}$ = 559.7ms) than the control stimuli ($\bar{x}$ = 2024.1ms for the reading time; $\bar{x}$ = 445.7ms for the response time).

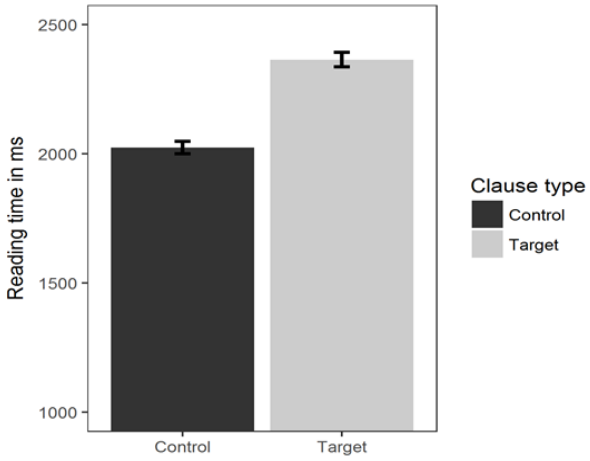


Figure 2. Participants' stimuli reading time with error bars showing standard errors

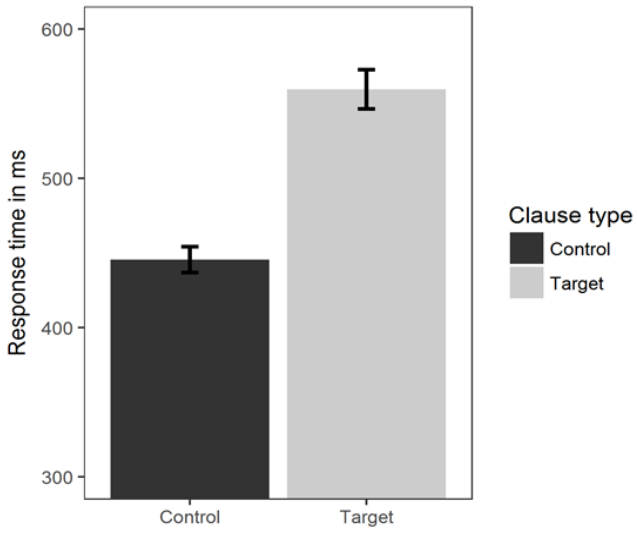


Figure 3. Participants' response time in indicating the clause types with error bars showing standard errors

The paired sample T-test in R showed significant differences between the target and control stimuli in both stimuli reading time and response time. The time participants spent in reading target stimuli is significantly longer than that in control stimuli, t (83) = 8.20, $p < 0.001$. Also, the time participants spent in indicating the clause type is significantly longer in target stimuli than that in control stimuli, t (83) = 6.28, $p < 0.001$. The results revealed that it is less straightforward to interpret a *wh*-sentence with *diǎnr* and more difficult to identify its clause type due to its ambiguity, as

compared with the control stimuli. The longer reading and response time offers another piece of evidence that *wh*-sentences with *diănr* are ambiguous between *wh*-questions and *wh*-declaratives.

Taken together, the results show that in the absence of *diănr*, the control stimuli with a *wh*-word are interpreted dominantly as *wh*-questions; in the presence of *diănr*, participants' response pattern (59.7% as questions and 40.3% as declaratives), as well as the longer stimuli reading time and response time as opposed to *wh*-sentences without *diănr* confirm that a *wh*-sentence containing *diănr* is indeed ambiguous between a declarative and a *wh*-question interpretation. This offers us empirical evidence that even if sentences like (4a) and (10a) do not conform to nonveridicality, *diănr* still can license the *wh*-existential reading, rejecting the analysis in Huang (2017) on empirical grounds.

4.4 An audio-gating study investigating the licensing of *diănr* on *wh*-declaratives

As shown from the above reading study, *wh*-words preceded by *diănr* as in (4a) and (10a) can have *wh*-existential interpretations, although they appear in a veridical context. As we mentioned above, although *diănr* is modifying the noun phrase following it and syntactically it belongs to the DP, *diănr* cliticizes obligatorily leftwards onto the verb-(perf) as in example (4a) and (10a), forming a verb-*diănr* prosodic clitic group (Shih, 1997; Chen, 2000). The question thus arises whether verb-*diănr* is required as a prosodic clitic group in licensing the *wh*-existential. To be specific, we want to investigate whether the licensing of *wh*-existentials is affected if we break this clitic group into different domains.

The answer to this question can be sought from an audio-gating paradigm we have used in Chapter 3, in which the audio recordings of *wh*-questions with *diănr* and *wh*-declaratives with *diănr* like (10a) were generated into different audio fragments for participants to listen to and to anticipate the upcoming clause type (i.e. a *wh*-question containing a question word or a *wh*-declarative containing a *wh*-existential). In this study, we had four types of audio fragments that were presented in four gates[21]. In addition to the three gates reported in Chapter 3, namely Gate *a* (subject), Gate *b* (subject and adverb) and Gate *d* (subject, adverb, verb plus *le* and *diănr*), audio fragments in Gate *c* consisted of subject, adverb, verb plus *le*. As already introduced in Chapter 3, the remaining sentence parts of each gate were shown on the screen as corresponding sentence continuations for listeners to choose from based on their anticipation. As shown in Chapter 3, listeners can anticipate the clause type incrementally in the sense that the more information/the longer audio fragments they hear, the more accurate their anticipation is. Given this, the clause type anticipation accuracy is expected to increase from gate *a* to gate *d* for both clause types, if verb-*diănr* is not required as a clitic group in licensing the *wh*-existential.

On the other hand, if verb-*diănr* is required as a clitic group in licensing the *wh*-existential, the anticipation of clause types might be affected. In gate *a* and *b* where

[21] In chapter 3, we report the results of 3 gates (gate *a*, *b* and *d*) for the investigation of clause type anticipation.

diănr co-occurs with the verb in the sentence continuations on the screen and in gate *d* where *diănr* co-occurs with the verb in the audio-fragment, verb-*diănr* can still be perceived as one unit/clitic group without being separated apart. The critical gate is gate *c*, in which *diănr* and the verb are separated, with the verb showing up in the audios and *diănr* in the sentence continuations on the screen. If verb-*diănr* is indeed required as a clitic group in the licensing of *wh*-existentials, our hypothesis is as follows. Different from the incremental clause type anticipation accuracy (for both *wh*-questions and *wh*-declaratives) from gate *a* to gate *b* and to gate *d*, listeners' anticipation accuracy for *wh*-declaratives in gate *c* would be affected (namely, less accurate than the other three gates), as in this gate *diănr* is separated from the verb. The details will be reported below.

4.4.1 Participants

See 3.3.1 in Chapter 3 for the details.

4.4.2 Acoustic stimuli

As introduced in 3.3.2 in Chapter 3, we have 40 stimuli (20 items × 2 clause types) that were used to generate audio fragments. Example (11) illustrates an item set.

(11) a. 陶薇 昨天 拿了 点儿 什么 给 刘刚? [*wh*-Q]
TáoWēi zúotiān ná-le diănr shénme gěi LíuGāng?
TaoWei yesterday bring-PERF a.little what for LiuGang
'What did TaoWei bring (a little) for LiuGang yesterday?'

b. 陶薇 昨天 拿了 点儿 什么 给 刘刚。 [*wh*-D]
TáoWēi zúotiān ná-le diănr shénme gěi LíuGāng.
TaoWei yesterday bring-PERF a.little something for LiuGang
'TaoWei brought a little something for LiuGang yesterday.'

As already introduced in Chapter 3, the 40 audio stimuli were directly used in our audio-gating experiment for generating gates. Here we only reported the acoustic properties in detail up to *diănr* where the acoustic information of all gates are given.

Duration. As reported in Chapter 3 and briefly mentioned here, the prosodic analysis of the two clause types up to *diănr* revealed that it is mainly duration that differentiates the two clause types before the *wh*-word. Figure 4 depicts the mean word duration in millisecond for *wh*-questions and *wh*-declaratives. As shown in Figure 4, already from the subject, there is a consistent duration difference between the two clause types.

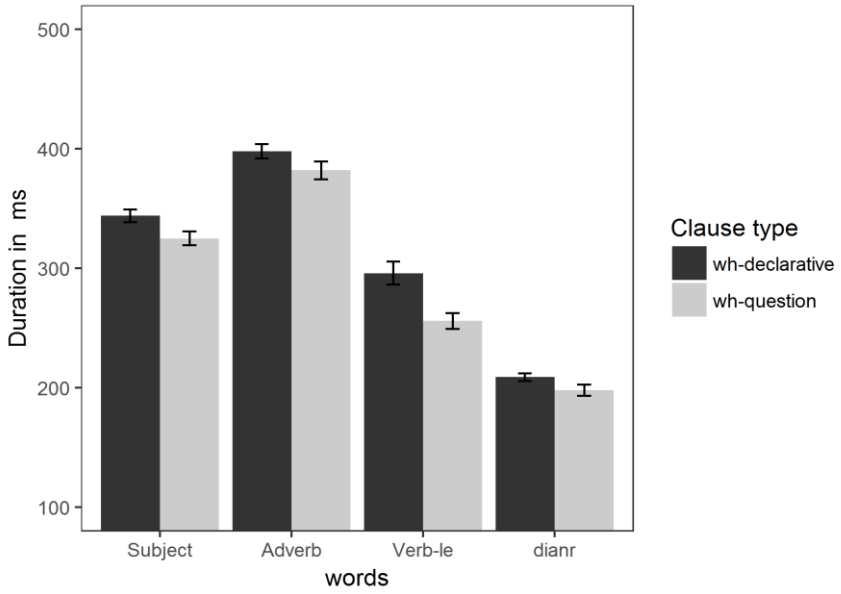


Figure 4. Mean word duration across clause types with error bar showing standard error.

4.4.3 Stimuli for audio-gating experiment

As introduced above, the 40 audio stimuli were used as a basis for generating the four types of audio fragments in the four gates. We also created two kinds of sentence continuations, *wh*-questions and *wh*-declaratives. Examples (12)-(15) showed an example of the audio fragments of the four gates and their corresponding sentence continuations on the screen. As the examples of gate *a*, *b* and *d* were already given in Chapter 3, here we just briefly presented them.

(12) Gate *a* (audio) sentence continuations (visual)
TáoWēi -- zúotiān ná-le diǎnr shénme gěi LíuGāng?
-- zúotiān ná-le diǎnr shénme gěi LíuGāng.

(13) Gate *b* (audio) sentence continuations (visual)
TáoWēi zúotiān -- ná-le diǎnr shénme gěi LíuGāng?
TaoWei yesterday -- ná-le diǎnr shénme gěi LíuGāng.

(14) Gate *c* (audio) sentence continuations (visual)
TáoWēi zúotiān ná-le -- diǎnr shénme gěi LíuGāng?
TaoWei yesterday bring-PERF a.little what for LiuGang
-- diǎnr shénme gěi LíuGāng.
a.little something for LiuGang

(15) Gate *d* (audio) sentence continuations (visual)

TáoWēi zúotiān	ná-le	diǎnr	-- shénme gěi LíuGāng?
TaoWei yesterday	bring-PERF	a little	-- shénme gěi LíuGāng.

This resulted in a total of 40 audio fragments in each gate and 160 in total. The detailed acoustic properties can be found in 4.4.2.

4.4.4 Procedures

See section 3.3.3 in Chapter 3 for the detailed procedure. The audio fragments were presented in four consecutive gates, from gate *a* to gate *d*, to make sure that the information participants heard were incremental. The whole testing time lasted about 25-30 minutes.

4.4.5 Results

The response accuracy for each clause type per gate was presented in Figure 5. As shown in Figure 5 and also reported in Chapter 3, the clause type anticipation accuracy in gate *a*, *b* and *d* demonstrates that listeners can make use of prosody to anticipate clause types and their clause type anticipation accuracy is incremental in the sense that the more information/longer audio fragments listeners hear, the more accurate their anticipation is.

Let us now turn to gate *c*, as this is the gate where *diǎnr* is (prosodically) separated from the verb. As shown in Figure 5, the response pattern of gate *c* differs from all the other three gates. The overall accuracy in *wh*-questions jumped to 73.1%[22] while the accuracy in *wh*-declaratives dropped to chance level 53.5% (lower than than the accuracy of gates *a* and *b*).

[22] As our audio-gating studies are forced choice studies, when participants didn't anticipate the audios to be declaratives (regardless the audios they heard are from *wh*-questions or *wh*-declaratives), they can only choose them to be questions; therefore the accuracy for *wh*-questions increased.

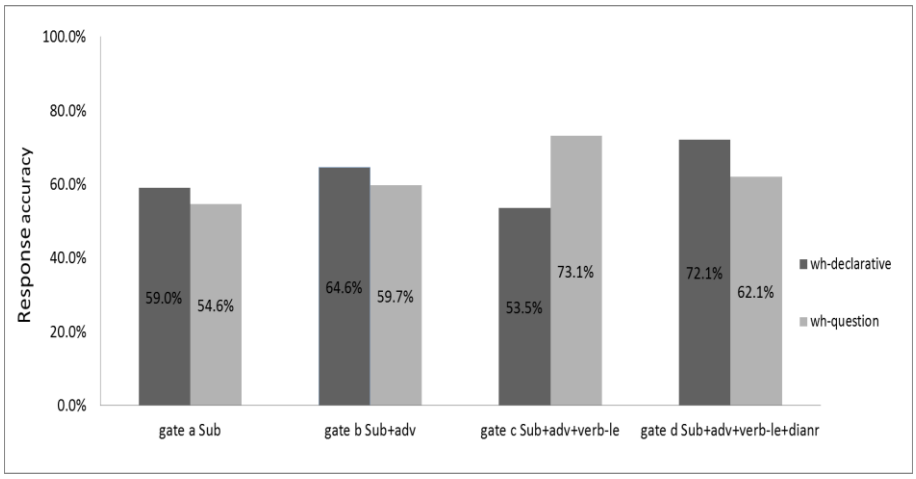


Figure 5. Listener's response in percentage (%) in gate *a*, *b*, *c* and *d*

In summary, the accuracy of the clause type anticipation is incremental as we see from the results of gate *a*, *b* and *d*, where *diănr* co-occurs with the verb in the same domain, either in sentence continuations presented on the screen (gate *a* and *b*) or in audios heard by listeners (gate *d*). When *diănr* was separated from the verb in gate *c*, the anticipation for *wh*-declaratives dropped to the chance level, even lower than the accuracy in gate *a*, in which only the subject is heard by the listeners. The audio-gating results are consistent with our hypothesis that although *diănr* can license the *wh*-existential, verb-*diănr* is required as a clitic group for a legitimate licensing of the *wh*-existential.

4.5 Discussions on the licensing of *diănr* on *wh*-existentials

We started our study from an interesting case, a veridical sentence containing a *wh*-word and *diănr*, which can have both an interrogative reading and a declarative reading. We repeat it here as (16).

(16) 陶薇 昨天 拿了 点儿 什么 给 刘刚
TáoWēi zúotiān ná-le diănr shénme gěi LíuGāng
TaoWei yesterday bring-PERF a.little what for LiuGang
'What did TaoWei bring a little for LiuGang yesterday?' or
'TaoWei brought a little something for LiuGang yesterday.'

Although a sentence such as (16) does not conform to the typical nonveridicality condition (Lin et al., 2014; Huang, 2017), our reading study shows that it has both interrogative (interpreting *shénme* as *wh*-interrogative 'what') and declarative readings (interpreting *shénme* as *wh*-existential 'something'). To our limited knowledge, this study is the first empirical study that investigates the licensing of *wh*-existentials in veridical contexts, confirming that *diănr* can provide the existential force in licensing the *wh*-existential interpretation and at the same time

rejecting the claim that *wh*-existentials can only be licensed in nonveridical contexts. Hence, the licensing environments of *wh*-existentials will be revised and proposed later in (18) based on our findings and the previous analysis of the licensing of *wh*-existentials in nonveridical conditions.

Our audio-gating experiment further demonstrates that although the presence of *diănr* can license the interpretation of a *wh*-existential, the licensing comes with a constraint. In other words, in order to license the *wh*-existential, *diănr* must co-occur with the verb as a prosodic clitic group, while in contrast, the interpretation of a *wh*-interrogative does not have such a requirement.

Taken together, our two empirical studies confirm that *diănr* can license *wh*-existentials, even in veridical contexts, but the licensing has a constraint: *diănr* has to cliticize to the verb prosodically as a clitic group in licensing the *wh*-existential in the VP. Further, these findings not only reject the claim that *wh*-existentials can only be licensed in nonveridical contexts, but also shed light on the puzzle that in the "weak" nonveridical contexts (i.e. future environments as in (3) and (9b)), *diănr* seems to be generally required for the licensing of the *wh*-existential reading. To be specific, we suggest that under these "weak" nonveridical contexts, it is actually *diănr* that is functioning as a licensor for the licensing of the *wh*-existential and that is why the presence of *diănr* becomes necessary.

In the current study, we do not zoom in on the details of the abstract formal semantic discussion on the licensing of *wh*-existentials as Lin (1998) or Huang (2017) did. Based on the above empirical studies and discussions, we aim to shed light on the licensing environments of *wh*-existentials in Mandarin, especially when it is preceded by *diănr*. If a *wh*-word in a sentence can be interpreted as a *wh*-existential, the sentence normally has the following environments:

(18) i. Either the sentence contains a nonveridical operator (negation, questions, conditionals and epistemic modalities) that can license the *wh*-existential.
or ii. When the nonveridical contexts (i.e. containing future environments) are not able to license *wh*-existential or when the sentence is simply veridical, it needs a last resort licensor like *diănr*. *Diănr* provides the existential force in licensing the *wh*-existential in the VP when *diănr* cliticizes to the verb.

Chapter 5 The Processing Mechanism of *Wh*-questions and Declaratives with Indefinites — Evidence from Mandarin

5.1 Introduction

As we have introduced, Mandarin Chinese is a *wh*-in-situ language in which question words remain at their base position, just as their declarative counterparts do, as illustrated in (1a-b). In contrast, English is a *wh*-movement language in which *wh*-words are fronted to sentence initial position (Spec-CP), as illustrated in (1c).

(1) a. 张三 买了 什么？ [*wh*-question]
Zhāng Sān mǎi-le shénme?
Zhang San buy-PERF what
'What did Zhang San buy?'

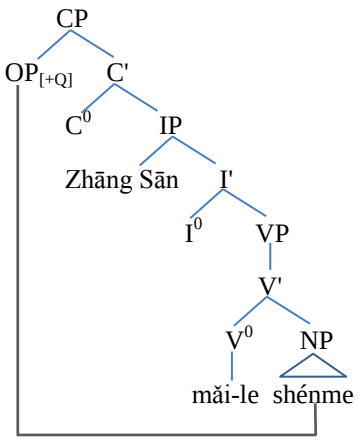
b. 张三 买了 一本 书。 [declarative]
Zhāng Sān mǎi-le yìběn shū.
Zhang San buy-PERF a CL book
'Zhang San bought a book.'

c. What$_i$ did Zhang San buy t_i? [*wh*-question]

In processing a *wh*-movement question like (1c), when the *wh*-word 'what' (filler) is encountered, the parser actively looks for its base position (gap) and establishes an overt dependency, according to existing psycholinguistic literature (Stowe, 1986; Frazier & Clifton, 1989). The filler-gap dependency has been widely investigated in many *wh*-movement languages. Nevertheless, for *wh*-in-situ questions like (1a), where there is no overt (non-local) dependency, little is known about how exactly an in-situ *wh*-word is processed.

As we have known in previous chapters, in addition to being in-situ, *wh*-words in Mandarin are on a par with indefinites in not having inherent quantificational force and their interpretations depend on operators/licensors. (Huang, 1982; Cheng, 1991; Aoun & Li, 1993; Tsai 1994, among others). When there is no overt licensor in the sentence as in example (1), the sentence is unambiguously a *wh*-question. In such cases, the *wh*-word is licensed by the interrogative operator (Q) at Spec-CP or C^o, and obtains the interrogative quantificational force, constructing a covert dependency between the in-situ *wh*-word and the interrogative operator (Q) at Spec-CP or C^o. The licensing/covert dependency in *wh*-questions is illustrated in (2).

(2)



The positing of a covert dependency in *wh*-in-situ questions has raised an interesting question for processing, namely, whether there is any processing evidence for establishing such a dependency in *wh*-in-situ languages like Mandarin. To the best of our knowledge, except for the pioneering work by Xiang, Dillon, Wagers, Liu and Guo (2013), no studies investigated the processing of Mandarin *wh*-in-situ questions, especially comparing *wh*-questions with their declarative counterparts. In the remaining part of this section, we discuss this article in detail.

Xiang et al. (2013) investigated the processing of Mandarin *wh*-questions with different lengths (mono-clausal and multiclausal), using the multiple-response speed-accuracy tradeoff paradigm (SAT)[23]. The *wh*-questions they tested are complex *wh*-questions, also called discourse-linked questions (i.e. *nǎxiē guānyuán* 'which officials'), as shown in (4). Xiang et al. found that in-situ *wh*-questions as in (4a) and (4b) were processed with longer reading times than their declarative counterparts as in (3a) and (3b) at the region of *wh*-phrases, indicating a higher processing cost in *wh*-questions.

(3) a. Declarative; Short

市政府	严惩了	那些官员.
Shìzhèngfǔ	yánchěng-le	nàxiē guānyuán.
city-council	punish	those officials

'The city council punished those officials.'

[23] There are more conditions in Xiang et al. (2013). We only reported the conditions that directly compared questions and declaratives.

b. Declarative; Long

市长	命令	市政府	严惩了	那些官员.
Shìzhǎng	mìnglìng	shìzhèngfǔ	yánchěng-le	nàxiē guānyuán.
mayor	order	city-council	punish	those officials

'The mayor ordered the city council to punish those officials.'

(4) a. *Wh*-question; Short

市政府	严惩了	哪些官员?
Shìzhèngfǔ	yánchěng-le	nǎxiē guānyuán?
city-council	punish	which officials

'Which officials did the city council punish?'

b. *Wh*-question; Long

市长	命令	市政府	严惩了	哪些官员?
Shìzhǎng	mìnglìng	shìzhèngfǔ	yánchěng-le	nǎxiē guānyuán?
mayor	order	city-council	punish	which officials

'Which officials did the mayor order the city council to punish?'

Xiang et al. (2013) interpreted the higher processing cost in *wh*-questions than in declaratives as empirical evidence for the establishment of a covert dependency between the in-situ *wh*-word and Spec-CP or C^o of the clause, as illustrated in (2). Despite of the insightful results bridging between the processing cost and the establishment of a covert dependency in in-situ questions, there are two issues in Xiang et al. (2013) that merit further discussions.

The first issue concerns the fact that the conclusion by Xiang et al. (2013), namely, *wh*-questions are more costly to be processed than declaratives, is based on *wh*-questions containing discourse-linked *wh*-phrase (i.e., a complex *wh*-phrase, *nǎxiē guānyuán* 'which officials') only. The notion "discourse-linking" was first defined by Pesetsky (1987). Similarly, Avrutin (2000) distinguished discourse-linked ('which x') and non-discourse-linked ('who') *wh*-phrase such that the former involves a discourse presupposition but the latter does not. That is, there is a mutual understanding in the discourse between the speaker and addressee that there is a presupposed set of items and the speaker wants to know which item in the presupposed set is the one in question. Rullmann and Beck (1998) also claimed that complex *wh*-phrases ('which x') are presuppositional elements, a particular kind of definites, different from other *wh*-words (i.e. simplex *wh*-phrases), which are indefinites in nature. It is also shown that the discourse-linking in complex *wh*-phrases incurs longer reading times as opposed to simplex *wh*-phrases in many languages (see De Vincenzi, 1996 for Italian; see Donkers, Hoeks & Stowe, 2013 for Dutch, among others). In other words, it remains a question whether the additional processing cost in *wh*-questions with discourse-linked *wh*-phrase can be extended to *wh*-questions with simplex *wh*-words. To address this issue, simplex *wh*-phrases (e.g. 'who') which entail no discourse-linking should be considered as compared with their declarative counterparts. We can then address the question of whether simplex *wh*-questions also result in higher processing cost than their declarative counterparts.

The second issue concerns the position of the critical region, namely, the *wh*-phrase in questions and the corresponding noun phrase in declaratives. In Xiang et al. (2013), all critical regions were posited at the sentence final position where the sentence-final wrap-up effects were often reported. According to the concept of wrap-up effects, readers tend to spend longer time in reading clause-final words than clause-internal words for an integration of information (Aaronson & Scarborough, 1976; Just & Carpenter, 1980; Rayner, Sereno, Morris, Schmauder, & Clifton, 1989). Although both *wh*-questions and declaratives may be affected by clause-final wrap-up effects, it remains unclear whether the two clause types are affected equally, which might weaken Xiang et al.'s empirical evidence of the additional processing cost they found at the *wh*-phrase. Given this, the processing cost incurred in *wh*-questions requires reinvestigation by excluding confounds involving sentence-final wrap-up effects.

To address the above issues and to attest whether the evidence of establishing a covert dependency in processing *wh*-questions can be extended to *wh*-questions without discourse-linking (i.e. containing simplex *wh*-words), we conducted altogether three self-paced reading studies on *wh*-questions and their declarative counterparts, with the *wh*-phrase and the corresponding noun phrase in the non-final position. Experiment 1 compared the processing of *wh*-questions of simplex *wh*-words (i.e. 'who') with their declarative counterparts with indefinites (i.e. 'someone'). The reason we compared *wh*-questions with simplex *wh*-words only with their declarative counterparts containing indefinites has to do with the definiteness of the noun phrase in declaratives. As we aim to examine whether simplex *wh*-questions result in higher processing cost than their declarative counterparts, the definiteness of the noun phrase in declaratives needs to be properly controlled. According to Warren & Gibson (2002), definite noun phrases are known to be more costly to process than their indefinite counterparts due to their different referential nature. Given that *wh*-phrases are indefinites in nature, comparing the processing of *wh*-questions with their declaratives counterparts containing indefinites would be a better comparison.

As the declaratives in Experiment 1 were shown to have ambiguous readings (see section 5.2.2 for details), we conducted Experiment 2 with a different set of stimuli. The stimuli constructed in Experiment 2 exclude the ambiguous readings on indefinites in declaratives and hence we can clearly detect the processing of simplex *wh*-questions and declaratives containing pure indefinites.

After comparing the processing of *wh*-questions containing simplex *wh*-phrases ('who') with their declaratives counterparts containing indefinites ('someone'), we conducted Experiment 3. In this experiment, we compared the processing of *wh*-questions containing complex *wh*-phrases (i.e. 'which x') with their declarative counterparts containing indefinites (i.e. 'an x'), different from Xiang et al. (2013) where complex *wh*-questions were compared with their declarative counterparts containing definites (i.e. 'the/those x'). The motivations for conducting Experiment 3 with a different comparison from Xiang et al. (2013) are two-fold. Firstly, we aim to investigate whether the processing cost resulted from the establishment of a covert dependency in *wh*-questions can be found in both types of *wh*-questions. Secondly, although Rullmann and Beck (1998) concluded that complex *wh*-phrases ('which x') are definites in nature, some studies also analyzed complex *wh*-phrases

as indefinites (see Karttunen, 1977 for details). Although we are not particularly interested about the exact nature of complex *wh*-phrases (definites or indefinites), a comparison between complex *wh*-questions and their declaratives counterparts containing indefinites might shed new light on our understanding of the processing of complex *wh*-questions, which have the discourse-linking feature and require the establishment of a covert dependency between the in-situ *wh*-word and the interrogative operator (Q).

5.2 Methodology

5.2.1 Experiment 1: Processing simplex *wh*-questions

5.2.1.1 Participants

Forty-two native speakers of Mandarin (10 males, $\bar{x}$ age = 20.4) participated in this experiment. All participants came from the northern part of Mainland China and at that time were students at Tsinghua University in Beijing. Participants were reimbursed for their participation and provided written informed consent prior to beginning the experiment.

5.2.1.2 Materials and design

Twenty-four sets of stimuli were constructed for this experiment, with a stimuli string "Det + Adj + Noun (human) + Verb-*le* + *shéi* (who) / *rén* (someone) + *jiù* (and then) + Verb-*le*". As illustrated in Example (5), each set contained 2 conditions[24], the simplex *wh*-questions with the *wh*-word *shéi* ('who') at the critical region, and the declarative counterpart with the indefinite *rén* ('someone') at the critical region. The word between slashes represents the reading region that participants encountered. See Appendix C for the specific list of materials used in the experiment.

(5) a. Simplex *wh*-question

那个/	优雅的/	绅士/	帮了/	谁/	就 /	离去了？
Nàgè	yōuyǎde	shēnshì	bāng-le	shéi	jiù	líqùle ?
that.CL	gracious	gentleman	help-PERF	who	then	left

'Who has the gracious gentleman helped and then left?'

[24] Our study includes one more condition that will be reported in a different paper. We only reported the two conditions relevant to our research question here. The same applies to the other two experiments.

b. Declarative with indefinite

那个/	优雅的/	绅士/	帮了/	人/	就/	离去了。
Nàgè	yōuyǎde	shēnshì	bāng-le	rén	jiù	líqùle.
that.CL	gracious	gentleman	help-PERF	person	then	left

'The gracious gentleman helped someone and then left.'

These twenty-four sets of stimuli were distributed into 2 lists in a Latin Square design and each list consisted of 24 sentences. In addition, 48 filler items were used to prevent participants from developing test-taking strategies. Twenty-four were *wh*-questions that contained the questions words *shénme shíhou* 'when' (n = 8), *nǎlǐ* 'where' (n = 8) and *zěnyàng* 'how' (n = 8) and 24 were declaratives. Each participant therefore read 72 sentences in total.

5.2.1.3 Procedure

A word-by-word moving-window self-paced reading experiment (Just, Carpenter & Wooley, 1982) was run on a PC laptop using Linger software (Doug Rhode, MIT). Participants were instructed to read the sentences carefully at their own pace by pressing the space bar. The experiment was preceded by eleven practice trials to help participants familiarize with the procedure.

All critical sentences and fillers were followed by comprehension questions. The questions asked about different parts of the sentence so that the participants could focus equally on all parts of each sentence. Half of the comprehension questions had ''yes'' answers, and the other half had ''no'' answers. Participants were instructed to answer the questions by pushing the F key for ''yes'' and the J key for ''no''. The computer showed ''Wrong answer" in Chinese if the questions were incorrectly answered, but no feedback was given if the answers were correct. The whole experiment lasted approximately 20 minutes.

The experimental procedure was the same for the other two self-paced readings reported in section 5.2.3 and 5.2.4.

5.2.1.4 Data analysis and results

Of the 42 participants, one participant's data was excluded from analysis due to scoring a low reading rate. Therefore, 41 participants were included in the following analysis.

5.2.1.4.1 Comprehension task results

The overall accuracy of the critical sentences across all participants was 97.4%. Further, per condition, the accuracy was 97.9% for the questions, 97.0% for the declaratives. The high accuracy rate showed that participants were attentive in reading.

5.2.1.4.2 Reading Times (RTs) analysis

For the raw RTs in each region, we first log-transformed them to adjust the heavily skewed distribution, following existing Mandarin studies (Wu, Kaiser & Andersen, 2012; Xiang, Wang & Cui, 2015) and other East Asian language studies (e.g. Kwon & Sturt, 2013 for Korean). Furthermore, data points more than 3 Standard Deviations (SD) above or below the mean for each region were excluded from the analysis, affecting 0.9% data. The resulting log-RTs at each region were analyzed with linear mixed effects models in R using the lmerTest package (Kuznetsova, Brockhoff & Christensen, 2013) with condition as a fixed effect factor, participants and items as random factors, allowing by-participant and by-item random intercepts, and by-participant and by-item random slopes for conditions[25].

We followed the same protocol of the RTs analysis in the other two self-paced readings reported in section 5.2.3 and 5.2.4.

5.2.1.4.3 Results

The mean log RTs per region across the two conditions are shown in Figure 1, which demonstrates that simplex *wh*-questions were processed much faster than their declarative counterparts at the region immediately after the *wh*-word, the conjunction *jiù* (‘then’).

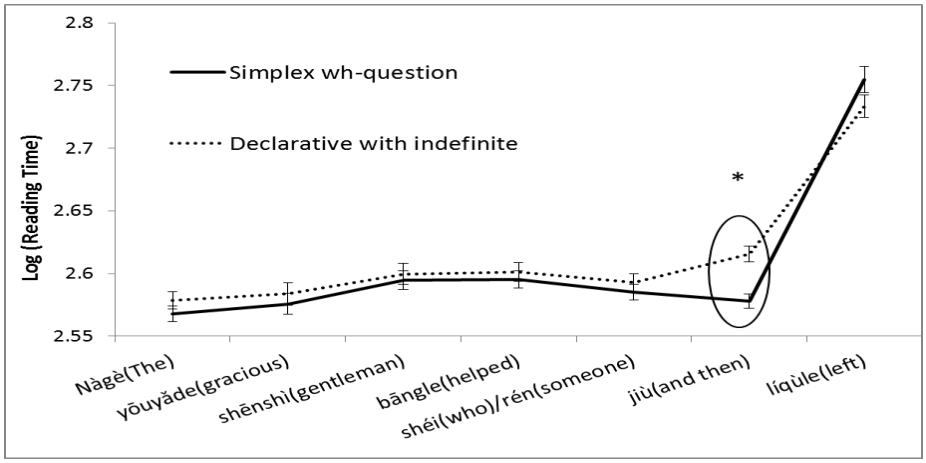


Figure 1. Mean log RTs per region with standard errors

[25] We allowed random slopes for all but after comparing models using the likelihood ratio test (Pinheiro & Bates, 2000; Bolker, Brooks, Clark, Geange, Poulsen, Stevens, & White, 2009), we found that adding random slopes didn’t improve the model. Hence the results we reported are based on the simpler model “Reading <- lmer (readingtime ~ condition + (1|subject) + (1|item), data=data)”.

A linear mixed effects model revealed a significant difference at the conjunction *jiù* ('then'). The reading time of questions is significantly shorter than that of the indefinite declarative condition ($\beta = -0.04$, $p < 0.001$). See Table 1 for the detailed results of the model.

Table 1. Summary of the linear mixed effects models at regions with significant reading time differences between conditions.

	Estimate β	Std. Error	t-value	p-value
conjunction (*jiù*)	-0.039	0.007	-5.582	< 0.001

To sum up, in the processing of simplex *wh*-questions and their declarative counterparts with indefinites, we found that questions were actually processed faster than declaratives. This processing pattern is contrary to the prediction that processing *wh*-questions should incur more cost than their declarative counterparts as a result of the establishment of a covert dependency in *wh*-questions. We will discuss the potential mechanism behind the unpredicted pattern in the following section.

5.2.2 Interim discussion

Our self-paced reading results show that simplex *wh*-questions with 'who' are processed with less cost than their declarative counterparts with 'someone', which was not expected based on the covert dependency theory discussed in section 5.1. Before we reach any conclusion, we must consider possible confounding factors such as word frequency or sentence stimuli in the current experiment.

Frequency. Word frequency has been shown to affect processing as readers tend to spend longer time when accessing infrequent words (Just & Carpenter, 1980; Kliegl, Grabner, Rolfs & Engbert, 2004, among others). In the current experiment, the simplex *wh*-questions and their declaratives counterparts only differ in the *wh*-word position. That is, the only differences are *shéi* 'who' and *rén* 'person'. In terms of frequency, both *shéi* and *rén* are highly frequently used words in Mandarin as shown by the Chinese online corpus (www.cncorpus.org). When used in isolation as one simple word, the word frequency of *shéi* and *rén* is 0.485‰ and 3.549‰ respectively. The fact that *rén* has a very high word frequency[26] exclude the possibility that the longer processing time at *rén* than *shéi* is due to their word frequency.

The interpretation of *rén* 'person'. Although *rén* in Mandarin normally receives an indefinite interpretation 'someone' at the object position, it can also obtain a definite interpretation under some conditions. For instance, as in our stimuli, when *rén* appeared within a bounded context (i.e., sentences containing a verb with a perfective marker *le*; see Sybesma, 1992 for details), *rén* can potentially be

[26] The *wh*-word *shéi* 'who' can have non-interrogative interpretations in other contexts (as introduced in section 1) and *rén* can also have other interpretations (see the discussion in the next paragraph). The reader should be aware that the word frequency reported here is a general word frequency without distinguishing the different usages or interpretations.

interpreted ambiguously between a definite and a specific indefinite interpretation (see Cheng & Sybesma, 1999). This ambiguity with respect to the interpretation on the noun phrase in declaratives may have incurred extra processing cost in the declarative condition. This can to some degree explain why the processing of declaratives with *rén* is slower than that of *wh*-questions, contrary to our expectations.

To restrict the interpretation of *rén* unambiguously to indefinite interpretation ('someone'), we created new stimuli that do not contain the perfective marker *le* or any bounded event in the sentence. With the new stimuli, we conducted the second self-paced reading study on *wh*-questions with simplex *wh*-phrases as compared with their declarative counterparts with indefinites, which will be reported in the next section.

5.2.3 Experiment 2: Processing simplex *wh*-questions with new stimuli

5.2.3.1 Participants

Thirty-six native speakers of Mandarin Chinese (20 males, $\bar{x}$ age = 20) participated in this experiment. They came from the northern part of Mainland China and were recruited from Tsinghua University (Beijing). None of them have participated in the other two self-paced reading experiments. Participants were reimbursed for their participation and provided written informed consent prior to beginning the experiment.

5.2.3.2 Materials and design

In this experiment, we compared in-situ *wh*-questions with *wh*-words *shéi* 'who' as in (6a), with declaratives that contained indefinites such as *rén* ('person') as in (6b). Crucially, by avoiding using the perfective marker '*le*' and by using intensional verbs[27] (e.g. 'want') instead, there is no bounded context that can bring ambiguous readings in declaratives. Hence *rén* ('person') can only be interpreted as an indefinite. Example (6) provides a set of sample stimuli. See Appendix D for the specific list of materials we used in the experiment.

(6) a. In-situ question with a simplex phrase

那个/	男生/	想要/	求/	谁/	解决/	问题?
Nàgè	nánshēng	xiǎngyào	qiú	shéi	jiějué	wèntí?
that.CL	boy	want	ask	who	solve	problem

'Who does the boy want to ask to solve the problem?'

[27] For the details of the discussions of the intensional verbs and their intensional readings, see Moltmann (1997).

b. Declarative with indefinite object noun phrase

那个/	男生/	想要/	求/	人/	解决/	问题。
Nàgè	nánshēng	xiǎngyào	qiú	rén	jiějué	wèntí.
that.CL	boy	want	ask	person	solve	problem

'The boy wants to ask someone to solve the problem.'

These 24 sets of stimuli were distributed into 2 lists in a Latin Square design and each list consists of 24 sentences. In addition, 72 filler items with 36 declaratives and 36 questions were used to prevent participants from developing test-taking strategies. Each participant therefore read 96 sentences in total.

5.2.3.3 Data analysis and results

All the 36 participants' data were included in the analysis.

5.2.3.3.1 Comprehension task results

The overall accuracy including fillers was 96.3%. Per condition, the accuracy was 97.6% for the simplex *wh*-questions and 97.9% for the declaratives. Again, the high accuracy rate demonstrated that participants were attentive in reading.

5.2.3.3.2 Results

Figure 2 presents the mean log RTs per region across the two conditions in (6a) and (6b). As we can see, simplex *wh*-questions were processed more slowly than their declarative counterparts immediately after the *wh*-phrase.

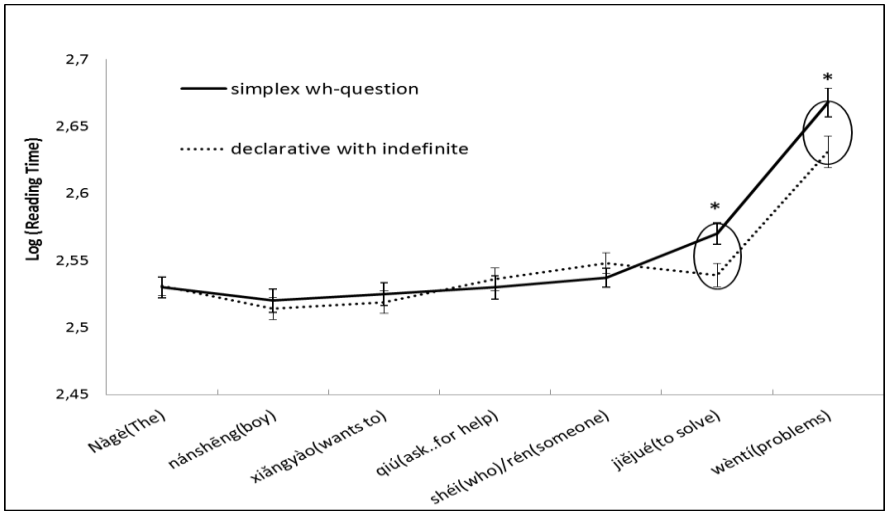


Figure 2. Mean log RTs per region with standard errors

A linear mixed effects model revealed a significant difference at the verb region immediately following the *wh*-word and at the final noun. In-situ questions with a simplex *wh*-phrase were processed with a significantly longer reading time than indefinite declarative conditions at the verb region (*jiějué* 'to solve') (β = 0.031, p < 0.01) and the sentence final noun region (*wèntí* 'problems'), (β = 0.037, p < 0.01). See Table 2 for the detailed results of the model.

Table 2. Summary of the linear mixed effects models at regions with significant reading time differences between conditions.

	Estimate β	Std. Error	t-value	p-value
verb (*jiějué*)	0.031	0.010	3.190	< 0.01
noun (*wèntí*)	0.037	0.012	3.104	< 0.01

In summary, by utilizing the new stimuli, we ruled out the definite interpretation of *rén*, restricting it to the indefinite interpretation ('someone'). And after addressing issues of the definiteness of noun phrases and the clause-final wrap-up issues we introduced in section 5.1, we found that at the region after the *wh*-word, simplex questions are processed significantly more slowly than declaratives containing indefinites. This points to a higher processing cost in *wh*-questions, which provides evidence for the establishment of a covert dependency in processing Mandarin *wh*-in-situ questions, in the case of simplex *wh*-questions. As mentioned in section 5.1, to obtain a comprehensive and a new perspective in understanding *wh*-questions with complex *wh*-phrases, in Experiment 3 we tested complex *wh*-questions using similar stimuli as in Experiment 2, which do not contain the perfective marker *le*.

5.2.4 Experiment 3: Processing complex *wh*-questions with new stimuli

5.2.4.1 Participants

54 native speakers of Mandarin Chinese (23 males, $\bar{x}$ age = 27) participated in this experiment. The participants came from the northern part of Mainland China and at that time were MA and PhD students studying at Leiden University, the Netherlands.[28] Participants were reimbursed for their participation and provided written informed consent prior to beginning the experiment.

5.2.4.2 Materials and design

The current experiment consisted of the same stimuli as that in Experiment 2, except for the *wh*-word region. In this experiment, we compared in-situ *wh*-questions with *wh*-phrases such as *nǎgè tóngxué* 'which classmate' in (7a), with declaratives that contained indefinite noun phrases such as *yígè tóngxué* 'a classmate' in (7b). Crucially, intensional verbs are used, and there is no bounded context. See Appendix E for the specific list of materials we used in the experiment.

[28] The students participating in our experiment had not been in the Netherlands for more than 3 years by the time of testing.

(7) a. In-situ question with a complex phrase

那个/	男生/	想要/	求/	哪个/	同学/	解决/	问题?
Nàgè	nánshēng	xiǎngyào	qiú	nǎgè	tóngxué	jiějué	wèntí?
that.CL	boy	want	ask	which	classmate	solve	problem

'Which classmate does the boy want to ask to solve the problem?'

b. Declarative with indefinite object noun phrase

那个/	男生/	想要/	求/	一个/	同学/	解决/	问题。
Nàgè	nánshēng	xiǎngyào	qiú	yígè	tóngxué	jiějué	wèntí.
that.CL	boy	want	ask	a	classmate	solve	problem

'The boy wants to ask a classmate to solve the problem.'

These 24 sets of stimuli were distributed into 2 lists in a Latin Square design and each list consisted of 24 sentences. In addition, the same 72 filler items as in Experiment 2 were used to prevent participants from developing test-taking strategies. Each participant therefore read 96 sentences in total.

5.2.4.3 Data analysis and results

All the 54 participants' data were included in the analysis.

5.2.4.3.1 Comprehension task results

The overall accuracy including fillers was 95.4%. Per condition, the accuracy was 94.2% for the complex *wh*-questions and 97.5% for the declaratives containing indefinites. The high accuracy rate showed that participants were attentive in reading.

5.2.4.3.2 Results

Figure 3 presents the mean log RTs per region across the two conditions. As shown, complex *wh*-questions were processed more slowly than declarative counterparts starting right at the determiner of the *wh*-phrase *nǎgè* 'which'.

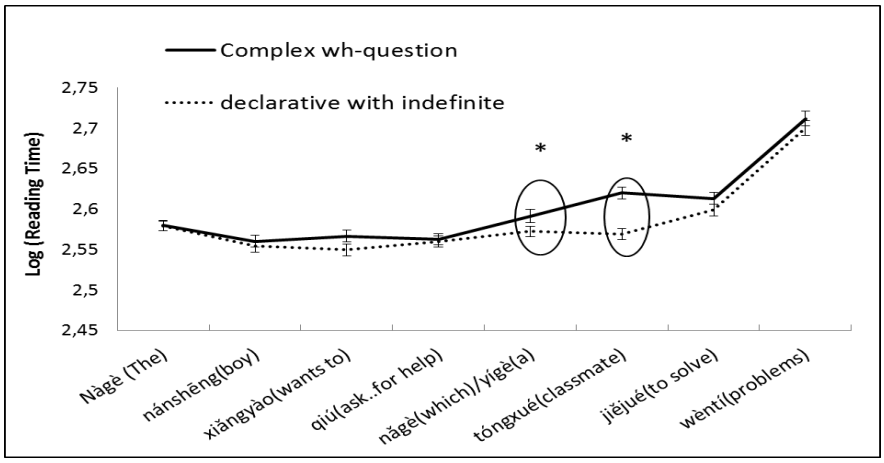


Figure 3. Mean log RTs per region with standard errors

A linear mixed model revealed a significant difference of the reading time at the region of the determiner (*nǎgè* 'which', *yígè* 'a'), where the in-situ questions with a complex *wh*-phrase were processed significantly more slowly than the indefinite declaratives (β = 0.019, $p < 0.05$). At the region of the noun (*tóngxué* 'classmate') immediately following the determiner, the results showed that in-situ questions were also processed significantly more slowly than declarative conditions (β = 0.051, $p < 0.001$). See Table 3 for the detailed results of the model.

Table 3. Summary of the linear mixed effects models at regions with significant reading time differences between conditions.

	Estimate β	Std. Error	t-value	p-value
nǎgè/yígè	0.019	0.008	2.434	< 0.05
noun (*tóngxué*)	0.051	0.008	6.530	< 0.001

To sum up, in the processing of complex *wh*-questions, we again found that *wh*-questions are processed with a significantly longer reading time than their declarative counterparts, consistent with our findings in the simplex *wh*-questions (Experiment 2). The results are also consistent with the findings in Xiang et al. (2013) that complex *wh*-questions are processed with more cost than declaratives, despite of the fact that our declaratives contain indefinites/indefinite noun phrase as the counterpart of *wh*-phrases.

5.3 Discussion and conclusion

The main focus of this chapter was to investigate whether processing Mandarin *wh*-in-situ questions in isolation incurs more processing cost than that in declaratives, by testing how Mandarin speakers process simplex and complex *wh*-questions with respect to their indefinite declarative counterparts. By designing new stimuli that eliminated the interpretation ambiguity of the indefinites, our studies on both

simplex and complex *wh*-questions reveal that in-situ *wh*-questions in Mandarin are processed with a longer time, suggesting a higher processing cost than their indefinite declarative counterparts.

The additional processing time/cost in *wh*-questions as opposed to declaratives provides evidence for the parser's establishing of a covert dependency between *wh*-words and the interrogative operator (Q) at a higher position in the clause. To be specific, when an in-situ *wh*-word is encountered during processing and when no other licensors can potentially license the *wh*-word, a covert dependency between the interrogative operator (Q) and the *wh*-word is established and the sentence is unambiguously a *wh*-question. Contrasting with question-words, indefinites like 'a classmate' or 'someone' in declaratives as in 'ask a classmate/someone' are locally selected by the verb and hence incur less processing cost than *wh*-questions.

In addition to the empirical evidence of the extra processing cost in *wh*-questions, we notice that the processing delay in complex questions and simplex questions as compared with their declarative counterparts starts at different regions. In particular, in the simplex ones, the processing delay starts immediately after the *wh*-word *shéi* 'who', covering two regions thereafter, whereas in the complex ones, the processing delay starts right at the *wh*-determiner, covering the two regions of the *wh*-phrase. In other words, the slowdown in processing complex *wh*-questions as compared with declaratives starts earlier than that in simplex questions.

For the different slowdown patterns observed between complex and simplex *wh*-questions, we assume that it is due to the fact that complex *wh*-phrase ('which x') involves discourse-linking, incurring extra processing cost. Complex *wh*-phrase (e.g. 'which classmate') is infelicitous when no set (of classmates) is pre-established in the discourse context whereas a simplex *wh*-word ('who') can be used without any existing contexts. Note that in our study, no previous context is provided to the participants. Hence, the additional processing cost in complex *wh*-questions because of discourse-linking can account for the slowdown timing differences between complex and simplex *wh*-questions when compared to their indefinite declarative counterparts.

To conclude, by comparing in-situ *wh*-questions containing simplex *wh*-phrases ('who') with declaratives containing indefinites ('someone'), we excluded other confounding factors such as definiteness and discourse-linking and scrutinized the processing evidence of establishing a covert dependency online between the in-situ question word and the interrogative operator at Spec-CP or C^o, supporting the theoretical account that the *wh*-word is licensed by the interrogative operator (Q) at Spec-CP or C^o, and obtains the interrogative quantificational force. In addition to the evidence of establishing a covert dependency in *wh*-questions, our studies also show that complex *wh*-questions are processed differently from simplex questions due to the discourse-linking of the former.

Chapter 6 Effects of Prosody on Clausal Typing in Mandarin — Evidence from ERP Studies

6.1 Introduction

Spoken language processing involves the use of different sources of information such as syntax, prosody, lexical knowledge and context. The role prosody plays in sentence processing has been intensively investigated in previous behavioral studies. For instance, some studies examine how prosody is employed in disambiguating sentence structures (i.e. how prosodic breaks or pauses help to resolve garden-path effects) (Marslen-Wilson, Tyler, Warren, Grenier & Lee, 1992; Speer, Kjelgaard & Dobbroth, 1996, among others). Some others investigate how prosodic prominence (e.g. pitch accent) interacts with information structure or reference resolution (Terken & Hirschberg, 1994; Dahan, Tanenhaus & Chambers, 2002; Weber, Braun & Crocker, 2006; Chen, den Os & de Ruiter, 2007, among others).

In addition, previous studies have also investigated the role of prosody in clausal typing; that is, its role in identifying questions or declaratives based on prosodic cues. These studies have shown that in some languages, clause type (question or declarative) is prosodically marked and listeners can identify the clause type by utilizing its prosodic cues such as F0 and duration (for behavioral studies, see Vion & Colas, 2006; Heeren, Bibyk, Gunlogson & Tanenhaus, 2015; Gryllia, Yang, Doetjes & Cheng, 2016; Shiamizadeh, Caspers & Schiller, 2017, a, b). Nevertheless, exactly when, in terms of the accurate time course, is the prosodic information detected and utilized in clausal typing is not a simple matter that can be easily detected in behavioral studies.

Different from the behavioral studies, event-related potentials (ERPs) provide a continuous measure of auditory processing with an excellent temporal resolution, and hence serve as an ideal measure for revealing the specific time course of auditory processing. Furthermore, several distinct ERP components are correlated with prosody-related processing, based on which we can investigate the role of prosody in clausal typing. In the next subsection, we first introduce those ERP components at length. Then we take a closer view at previous ERP studies on the effects of prosody on clausal typing and thereafter we raise our research questions concerning the effects of prosody on clausal typing in Mandarin, which motivate us to conduct two auditory ERP experiments.

6.1.1 Event-related potentials and prosody

In this subsection, we introduce five ERP components, which have been suggested to correlate with prosody-related processing.

Closure Positive Shift. The CPS is a positive waveform distributed often bilaterally (Steinhauer, Alter & Friederici, 1999; Steinhauer, 2003). It is a component elicited in synchronization to prosodic phrase boundaries, represented by pauses and boundary lengthening in spoken language. Take the two German

sentences in (1) from Steinhauer et al. (1999) as an example; CPS components have been observed in all intonation phrase boundaries (marked with IPH) in (1).

(1) a. [Peter verspricht Anna zu arbeiten]$_{IPH1}$[und das Büro zu putzen.]$_{IPH2}$
Peter promises Anna to work and the office to clean
'Peter promises Anna to work and to clean the office.'

b. [Peter verspricht]$_{IPH1}$ [Anna zu entlasten]$_{IPH2}$ [und das Büro zu putzen.]$_{IPH3}$
Peter promises Anna to unburden and the office to clean
'Peter promises to unburden Anna and to clean the office.'

Since CPS was revealed as a specific neural response to prosodic boundary processing, the elicitation of CPS can be used in the investigation of how prosodic boundaries help to resolve garden-path effects by inserting boundaries/pauses (see Bögels, Schriefers, Vonk & Chwilla, 2011; Pauker, Itzhak, Baum & Steinhauer, 2011 for detailed discussions).

RAN or Prosodic Negativity. Right Anterior Negativity (RAN) is an anterior negative component distributed in the right hemisphere, and prosodic negativity is a broadly distributed negative component, both of which are often elicited between about 300 to 500 ms after the onset of the critical word. Those two ERP components have been demonstrated to indicate a prosodic mismatch or the detection of unexpected intonation during auditory processing (Eckstein & Friederici, 2005, 2006; Honbolygo, Torok, Banreti, Hunyadi & Csepe, 2016). In Hungarian for instance, a relative clause (in bold) as illustrated in (2a) is normally marked with a rising contour. In contrast, a falling intonation contour as in (2b) is incongruent. The incongruent contour in (2b) as compared with the congruent one in (2a) elicited a RAN effect (Honbolygo et al., 2016).

(2) a. Congruent intonation contour
[A nagyapa,]$_{IPH1}$ [**aki belépett**,]↗$_{IPH2}$ [szomjas volt.]$_{IPH3}$
the grandfather who entered thirsty was
'The grandfather, **who entered**, was thirsty.'

b. Incongruent intonation contour
[A nagyapa,]$_{IPH1}$ [**aki belépett**,]↘$_{IPH2}$ [szomjas volt.]$_{IPH3}$
the grandfather who entered thirsty was
'The grandfather, **who entered***, was thirsty.'

Also, negative components were found in an ERP study on German (Eckstein & Friederici, 2006[29]). In a German declarative sentence, only the boundary tone is marked with a rise-fall contour and other constituents signaling sentence continuations (e.g. penultimate position) are normally marked with a fall-rise contour. In Eckstein and Friederici (2006), they manipulated the pitch contour of the

[29] There are other conditions in Eckstein and Friederici (2006) (i.e. the pure manipulations of word category based on affixes). We only discussed the relevant conditions about prosody (in)congruities.

penultimate position of declaratives. The critical word at the penultimate position with an incongruent rise-fall contour (signaling sentence end) as compared with the identical lexical word with a congruent fall-rise contour elicited a broadly distributed prosodic negativity since the onset of the critical word.

In short, although these studies manipulated prosody in different ways and found negativities with different topographical distributions, the negativities seem to be important indicators for the detection of prosodic incongruity.

P600. The P600 effect is a posterior positivity and often peaks around 600 ms after the onset of the critical word (Osterhout & Holcomb, 1992). It has been elicited by various kinds of syntactic violations and has been interpreted as a process of reanalysis (Friederici, 2002). In addition to RAN or prosodic negativity, P600 is sometimes found in prosodic mismatch studies. For instance, in Honbolygo et al. (2016) as illustrated in example (2), the prosodic incongruity of the relative clause (2b) elicited P600 effects as compared with (2a). Honbolygo et al. interpreted this P600 as follows: "the modification of the intonation contour forced the parser to resolve the mismatch between the well-formed syntactic structures and the incongruent intonational realization by exercising a larger processing effort" (Honbolygo et al., 2016, p. 31).

Early Negativity or N200. Early prosodic negativities are negative components that usually peak around 200 ms after the target onset, indicating the early detection of the incongruity of prosodic features, such as the unexpected F0 or duration cues. Studies have shown that similar to RAN or prosodic negativity, early negativities were also found in incongruent prosody, the latency differences probably being subject to the exact time when the distinctive prosodic differences emerge and are perceived. For instance, in a study on French by Magne, Astésano, Aramaki, Ystad, Kronland-Martinet and Besson (2007), the anomalous lengthened/stressed critical word *candidats* 'candidates' (the underlying syllable is stressed) in (3b) as compared with the correct form *candidats* 'candidates' in (3a) elicited a negativity between 250 to 450 ms after the onset of the critical word. It is worth nothing that the first syllable "can" of the critical word, already lasted approximately 100 ms and prosodic differences only started to emerge at the second syllable 'di'. If time-locked to the onset of the second syllable 'di', this observed prosodic negativity actually indicated an early detection of the anomalous prosody.

(3) a. Congruent condition
Le concours a regroupé mille candidats.
the competition has hosted thousand candidates
'The competition hosted a thousand candidates.'

b. Incongruent condition
Le concours a regroupé mille candidats.
the competition has hosted thousand candidates
'The competition hosted a thousand candidates.'

An early negativity was also observed in a German study in a condition with an infrequent prosodic structure, namely, a vocative or calling contour, whose intonation contour typically comprises a downstepped tonal sequence from one

stable pitch level to another (Mietz, Toepel, Ischebeck & Alter, 2008). Example (4) illustrates the vocative condition (4b) as compared with the normal coordination declarative condition (4a), time-locked at the onset of the second noun *Patricia*. Results from this study by Mietz and colleagues have shown that infrequently used prosodic structures as in (4b) evoke early negativities similar to those elicited by anomalous prosody.

(4) a. Coordination declarative
Anton beisst, Patricia petzt und Carola lügt.
Anton bite Patricia squeal and Carola lie
'Anton bites, Patricia squeals and Carola lies.'

b. Vocative (infrequent)
Anton beisst, Patricia.
Anton bite Patricia
'Anton bites, Patricia.'

N400. The N400 is a negative potential that peaks approximately 400 ms after the target onset with a centro-parietal distribution. The N400 has been considered as an electrophysiological signature of semantic anomaly indexing lexical-semantic integration (Kutas & Hillyard, 1980). The N400 has also been sometimes reported in auditory processing when the alteration of prosody affects the lexical-semantic integration or information structure status, or when the prosodic pattern is less frequently used. For instance, in the above-mentioned study Mietz et al. (2008), N400 effects were also evoked in the vocative condition in (4b) as compared with (4a), in addition to the early negativities mentioned above, since the infrequently used vocative contour in (4b) affected the semantic comprehension (Mietz et al., 2008).

6.1.2 Previous ERP studies on effects of prosody in clausal typing

The above literature review shows that although the manipulations of prosodic congruities can be different (word level as in Magne et al., 2007; clause level as in Honbolygo et al., 2016, among others), (early) prosodic negativities seem to be important indicators for the detection of prosodic incongruity. After reviewing the relevant ERP components, we now turn to a closer inspection of the role prosody plays in clausal typing. To the best of our knowledge, very few existing ERP studies have investigated the issue of clausal typing, let alone how and when exactly prosodic information was detected and employed in clausal typing. A relevant study concerning clause typing is Astésano, Besson and Alter (2004), an ERP study on French declaratives and yes-no questions, which are string-identical. We review this study in detail below.

In French, a yes-no question with a subject followed by an intransitive verb normally has a falling contour at the subject of the sentence and then a rising contour at the intransitive verb. In contrast, its declarative counterpart is marked with a rising contour at the subject followed by a falling contour at the intransitive verb (Di Cristo, 1998). Astésano et al. (2004) manipulated the prosodic congruity of

sentences "cross-splicing" (Gernsbacher & Jescheniak, 1995) from the onset of the verb. To be specific, they spliced the verb from the natural declarative (D) audio in (5ai) and combined it with the subject of the natural yes-no question audio (Q) in (5aii), and vice-versa. It thus led to two kinds of prosodic contour incongruent conditions[30], as illustrated in (5bi) and (5bii), respectively. According to Astésano et al. (2004), the cross-splicing procedure introduces an acoustic dissymmetry between the respective noun phrase (i.e. subject) and verb in both incongruent conditions in (5b), and the cross-spliced contours do not correspond to any existing congruent intonation contours. In the incongruent condition, D-Q as in (5bi), the pitch contour becomes a rising-rising contour, with an unusually huge F0 distance between the offset of the subject and the onset of the verb. In the incongruent condition, Q-D as in (5bii), the pitch contour becomes a falling-falling contour, again with an unusually huge F0 distance between the offset of the subject and the onset of the verb.

(5) a. Prosodic congruent condition[31]

i. (D)La	lumière	clignotait.	or	ii. (Q)La	lumière	clignotait?	
the	light	flash		the	light	flash	
'The light is flashing.'				'Is the light flashing?'			

b. Prosodic incongruent condition[32]

i. (D)La	lumière	(Q)clignotait	or	ii. (Q)La	lumière	(D)clignotait	
the	light	flash		the	light	flash	

By comparing prosodic incongruent conditions with congruent ones, Astésano and colleagues (2004) found a P800 effect at the critical word (the verb). In other words, the verb *clignotait* 'to flash' in the incongruent conditions in (5b) as compared with the verb in the congruent condition in (5a) elicited a P800 effect. Astésano et al. took the elicitation of the P800 component to indicate the detection of the clause type incongruity based on the detection of prosodic contour incongruity.

The study by Astésano et al. (2004) provides us with a good paradigm for investigating the neural correlates of clausal typing based on prosody by applying "cross-splicing" on declaratives and yes-no questions, which are string identical in French. Nevertheless, there are two potential issues in Astésano et al. (2004) that merit further discussions. First, as illustrated in (5), in their design the authors merged the two kinds of incongruities (D-Q as in (5bi) and Q-D as in (5bii)) into one condition. Although the authors specified that they did not aim at comparing the processing differences between declaratives and yes-no questions, merging the two kinds of incongruities into a single condition might mask the potential differences in

[30] In Astésano et al.'s (2004) study, there are other conditions such as semantic congruent and incongruent conditions. We don't report those conditions, as they are not relevant to the current discussion.

[31] Astésano et al. (2004) merged the two kinds of congruent stimuli (declarative and yes-no question) as one condition.

[32] Similarly, Astésano et al. (2004) merged the two kinds of incongruent stimuli (D-Q and Q-D) as one condition.

the detection of the prosodic violation in declaratives and questions. For instance, in the D-Q incongruity in (5bi), listeners expected a continuation of the declarative but they heard a verb with the question prosody instead. This violation of clause type based on prosodic contour might not be equally reflected in the Q-D incongruent condition in (5bii), where listeners expected a continuation of the question but heard a verb with declarative prosody instead. This merge of conditions as in (5bi) and (5bii) thus risks an overgeneralization of the neural correlate of the prosodic incongruity generated by different prosodic patterns. Hence, an accurate inspection on the potential processing differences between the two incongruity patterns (i.e., D-Q vs. Q-D) demands the separation of the two incongruity patterns into independent conditions.

The second issue concerns the P800 ERP component. As compared with the (early) negativities we discussed in section 6.1.1, this prosodic effect P800 emerged very late in the time course, in contrast to the conclusion that prosody plays an immediate role in sentence processing (Steinhauer, Alter & Friederici, 1999; Eckstein & Friederici, 2006, among others). To be more specific, in previous studies, the prosodic incongruity detection, whether in terms of incongruent intonation contour, prosodic prominence or prosodic boundary, as discussed in section 6.1.1, did not emerge as late as 800 ms after the onset of the prosodic incongruent constituent. Astésano et al. (2004) claimed that this late effect may reflect the reanalysis of the prosodic cue (F0) that violates the expected intonation contour in an attempt to integrate the prosodically incongruent information, like a late P600 effect. Nevertheless, even if this late effect reflects reanalysis, it remains unclear why participants cannot detect any prosodic incongruity as early as in the above-mentioned ERP studies. Given this, a further investigation is merited on the detailed time course of the neural correlates of a prosodic incongruity effect to test whether the detection of prosodic incongruity in clausal typing occurs early or late.

6.1.3 The current study

Existing studies on the effect of prosody in clausal typing are very rare and mainly restricted to yes-no questions (Astésano et al., 2004; Vion & Colas, 2006; Heeren et al., 2015). As discussed in previous chapters, with respect to the effect of prosody on clausal typing, Mandarin *wh*-questions and *wh*-declaratives (as illustrated in (6)) provide an ideal test case.

(6) a. 张三 买了 点儿 什么。 [*wh*-declarative]
Zhāng Sān mǎi-le diǎnr shénme.
Zhang San buy-PERF a.little SHENME
'Zhang San bought a little of something.'

b. 张三 买了 点儿 什么？ [*wh*-question]
Zhāng Sān mǎi-le diǎnr shénme?
Zhang San buy-PERF a.little SHENME
'What did Zhang San buy (a little of)?'

The advantage of using *wh*-questions and *wh*-declaratives in clausal typing is that they are string identical cases, such as those in Astésano et al. (2004), so that 1) apart from prosody and/or context, no syntactic or morphological cues can indicate their clause types in our speech; 2) there are no lexical differences involved in processing the two clause types. In other words, we can have a clear inspection of the role prosody plays in clausal typing.

In addition, our previous production study (Chapter 2) based on a design using similar stimuli as in (6) already shows that *wh*-questions and *wh*-declaratives are marked by different prosodic properties. In particular, the *wh*-word *shénme* is higher in pitch, expanded in pitch range and also longer in duration in *wh*-questions than in *wh*-declaratives. Besides, *wh*-declaratives differ from *wh*-questions in terms of duration from the beginning of the sentence; that is, *wh*-declaratives are longer than *wh*-questions starting from the sentence subject and the pattern only reverses at the *wh*-word. Moreover, our perception study based on similar stimuli as in (6) demonstrated that listeners can identify the clause type (*wh*-question or *wh*-declarative) accurately by hearing the *wh*-word *shénme* (over 90% in terms of accuracy); in addition, listeners can already have a preference for anticipating the correct clause type even after hearing the subject of the two clause types (Chapter 3). Thus, our production and perception studies provide the baseline for designing the current EEG experiments.

Although *wh*-questions and *wh*-declaratives provide us with a good test case for prosodic clausal typing in Mandarin during online sentence processing, there is still a lack of neurocognitive studies investigating the time course and neural basis of clausal typing based on prosody in this language. As discussed in section 6.1.2, the only existing ERP study on clausal typing (Astésano et al., 2004) did not provide us with a conclusive result, since merging D-Q and Q-D incongruities might potentially mask the processing differences that emerge from the incongruities of declaratives and those of questions; furthermore, the very late prosody-related effect (P800) challenges the claim that prosodic cues are immediately utilized in processing (Steinhauer, Alter & Friederici, 1999; Eckstein & Friederici, 2006, among others). In the ERP studies we conducted, we try to fill this gap and scrutinize the clausal typing of *wh*-questions and *wh*-declaratives during online processing by answering the following research questions: 1) What neural correlates can we find for clausal typing based on the detection of the prosodic marking incongruity in the two clause types, and what is the detailed time course of this prosodic incongruency detection in the brain? 2) How early in the sentence can we find the effects of prosody in clause type anticipation? 3) Is the detection of incongruous prosody symmetric in *wh*-questions and *wh*-declaratives?

We aim at answering these questions by conducting two auditory ERP studies on *wh*-questions and *wh*-declaratives preceded by contexts biasing each clause type and followed by conditions where we manipulate prosody. Experiment 1 tests *wh*-declaratives (D), *wh*-questions (Q) and their cross-spliced conditions D-Q and Q-D (audio recordings of *wh*-questions and *wh*-declaratives cross-spliced from the onset of *wh*-words onwards), all preceded by contexts biasing towards the clause type of the pre-*wh*-word region. This experiment investigates the clausal typing based on electrophysiological evidence of the detection of prosodic incongruity of clause types. For example, when listeners expect a *wh*-interrogative, as predicted by the

context and the prosody of the pre-*wh*-word region they already heard, an effect of clause type incongruity might occur if they hear a prosody associated with *wh*-existential from a *wh*-declarative sentence. Similarly, when they expect a *wh*-existential, as predicted by the context and the prosody of the pre-*wh*-word region, an effect of clause type incongruity might occur if they hear a prosody associated with *wh*-interrogative from a *wh*-question. As the cross-spliced *wh*-questions and *wh*-declaratives from the *wh*-word onwards should in principle give rise to prosodically incongruent clause types, the results of Experiment 1 can answer our research question 1) based on the electrophysiological evidence and its detailed time course. Experiment 2 manipulates the congruity between the contexts participants hear biasing *wh*-questions or *wh*-declaratives and the target sentences containing *wh*-questions or *wh*-declaratives that participants hear. It investigates whether we can find neural correlates for the early detection of prosodic incongruity which may start already at the subject position. The results of Experiment 2 can answer question 2). Finally, the results of both Experiments 1 and 2 can answer question 3) by means of creating two independent conditions for the two kinds of clause type incongruity patterns (i.e., D-Q vs. Q-D).

6.2 Experiment 1 Cross-splicings

6.2.1 Participants

Twenty-four students from Tianjin Normal University were paid to participate in the experiment. All were right-handed native speakers of Mandarin Chinese and were paid to participate in the experiment. They were undergraduate students with no known history of vision or hearing impairment or any cognitive or psychiatric disorder. Prior to testing, each participant gave written informed consent. Six participants were later excluded from further data analysis due to excessive eye or head movements or a failure in concentrating on the experiment according to the experimenter's observation and the participants' self-report. The remaining 18 participants (12 female) have a mean age of 21 years (SD = 2.8).

6.2.2 Materials

As introduced in the previous section, in this experiment we created target sentences of *wh*-questions, *wh*-declaratives and their cross-spliced conditions, all preceded by contexts that are biasing towards a question or a declarative. All target sentences and their preceding contexts were auditorily presented to participants. We first describe the four conditions in detail, starting from the target sentences and then their preceding contexts. Thereafter, we present the acoustic properties of the target sentences.

6.2.2.1 Four conditions

Altogether, we created 36 *wh*-questions, and 36 *wh*-declaratives which are string identical to the *wh*-questions. Examples (7a-b) illustrate a set consisting of a *wh*-

question and its string identical *wh*-declarative. All the *wh*-questions and *wh*-declaratives consist of 12 syllables (S), having the same word order as the stimuli created for production (Chapter 2) and perception studies (Chapter 3). We kept the number of syllables of all constituents constant across items and conditions.

(7) a. Congruent condition *wh*-declarative (D)

(D)林珊	昨天	寄了	点儿	什么	给	袁刚。
LínShān	zuótiān	jì-le	diǎnr	shénme	gěi	YuánGāng.
LinShan	yesterday	post-PERF	a.little	something	to	YuanGang

'Lin Shan posted a little something to Yuan Gang yesterday.'

b. Congruent condition *wh*-question (Q)

(Q)林珊	昨天	寄了	点儿	什么	给	袁刚?
LínShān	zuótiān	jì-le	diǎnr	shénme	gěi	YuánGāng?
LinShan	yesterday	post-PERF	a.little	what	to	YuanGang

'What did Lin Shan post a little to Yuan Gang yesterday?'

c. Incongruent condition (D-Q)

(D)林珊	昨天	寄了	点儿	**(Q)什么**	**给**	**袁刚**
LínShān	zuótiān	jì-le	diǎnr	**shénme**	**gěi**	**YuánGāng**
LinShan	yesterday	post-PERF	a.little	**what**	**to**	**YuanGang**

d. Incongruent condition (Q-D)

(Q)林珊	昨天	寄了	点儿	**(D)什么**	**给**	**袁刚**
LínShān	zuótiān	jì-le	diǎnr	**shénme**	**gěi**	**YuánGāng**
LinShan	yesterday	post-PERF	a.little	**something**	**to**	**YuanGang**

As the perception study (Chapter 3) showed that by hearing the *wh*-word *shénme* listeners can discriminate *wh*-questions and *wh*-declaratives very accurately, we cross-spliced *wh*-questions and *wh*-declaratives from the onset of *shénme* onwards, aiming for prosodically strong incongruent sentences. To be specific, we spliced the *wh*-word and the preposition phrase from a *wh*-question audio recording (Q) as in (7b) and combined it with the pre-*wh*-word fragment from the *wh*-declarative audio (D) as in (7a), resulting in an incongruent target sentence (D-Q) as illustrated in (7c) and vice-versa, resulting in the other incongruent target sentence (Q-D) as illustrated in (7d). We highlight the cross-spliced regions in bold in (7) for a clear illustration. Altogether we created 36 incongruent target sentences D-Q as in (7c) and 36 incongruent target sentences Q-D as in (7d).

In addition to the target sentences, we also created 36 contexts preceding *wh*-declaratives that only biased towards declaratives, and 36 contexts preceding *wh*-questions that only biased towards questions. As illustrated in example (8a-b), the contexts consist of two to three sentences and it is the final sentence in bold (e.g. 'Without asking, we know:' or 'We in a gossipy way ask:') that clearly biases the upcoming declarative or question. In the same item set, the context of the *wh*-question is in principle the same as that of the *wh*-declarative except for the final sentence.

(8) a. *Wh*-declarative (in bold) and its preceding (D) context
林珊和男友袁刚是异地恋。林珊经常准备一些礼物给袁刚，正好物流也方便。这不，昨天林珊又叫了快递。**我们不用问都知道：林珊昨天寄了点儿什么给袁刚。**
'Lin Shan and her boyfriend live in different cities. Lin Shan often prepares some gifts for Yuan Gang since express delivery is very convenient. Yesterday Lin Shan called the delivery service again. **Without asking, we know: Lin Shan posted something to Yuan Gang yesterday.**'

b. *Wh*-question (in bold) and its preceding (Q) context
林珊和男友袁刚是异地恋。林珊经常准备一些礼物给袁刚，正好物流也方便。这不，昨天林珊又叫了快递。**我们都八卦地问：林珊昨天寄了点儿什么给袁刚？**
'Lin Shan and her boyfriend live in different cities. Lin Shan often prepares some gifts for Yuan Gang since express delivery is very convenient. Yesterday Lin Shan called the delivery service again. **We in a gossipy way ask**: **What did Lin Shan post to Yuan Gang yesterday?**'

For the incongruent target sentences (D-Q and Q-D), we also added the contexts as in (8a-b) biasing declaratives and questions respectively. Example (9) illustrates an item set of the four conditions including both target sentences and their preceding contexts. Due to the limited space, we avoid repeating the contexts in each condition here; see Appendix F for the complete list of context-target sentences of all items and conditions. As illustrated in (9c-d), the unexpected *wh*-word *shénme* should give rise to a clear clause type incongruity based on prosody, as both the context and the pre-*wh*-word region point to the opposite clause type.

(9) a. *Wh*-declarative (D) preceded with context (D)

(D)Context (D)林珊	昨天	寄了	点儿	什么	给	袁刚。
LínShān	zuótiān	jì-le	diǎnr	shénme	gěi	YuánGāng.
LinShan	yesterday	post-PERF	a.little	something	to	YuanGang

'Lin Shan posted a little something to Yuan Gang yesterday.'

b. *Wh*-question (Q) preceded with context (Q)

(Q)Context (Q)林珊	昨天	寄了	点儿	什么	给	袁刚?
LínShān	zuótiān	jì-le	diǎnr	shénme	gěi	YuánGāng?
LinShan	yesterday	post-PERF	a.little	what	to	YuanGang

'What did Lin Shan post a little to Yuan Gang yesterday?'

c. Incongruent condition (D-Q) preceded with context (D)

(D)Context (D)林珊	昨天	寄了	点儿	**(Q)什么**	**给**	**袁刚**
LínShān	zuótiān	jì-le	diǎnr	**shénme**	**gěi**	**YuánGāng**
LinShan	yesterday	post-PERF	a.little	**what**	**to**	**YuanGang**

d. Incongruent condition (Q-D) preceded with context (Q)

(Q)Context (Q)林珊	昨天	寄了	点儿	**(D)什么**	**给**	**袁刚**
LínShān	zuótiān	jì-le	diǎnr	**shénme**	**gěi**	**YuánGāng**
LinShan	yesterday	post-PERF	a.little	**something**	**to**	**YuanGang**

Altogether, we have 144 target sentences with preceding contexts (4 conditions × 36 items). In addition, we created 144 filler sentences, including 72 declaratives, 36 yes-no questions and 36 questions with *wh*-adjuncts as the question words *shénme shíhou* ('when') or *nǎlǐ* ('where'), all preceded by contexts. Each participant listened to a total of 288 trials.

6.2.2.2 Acoustic properties of *wh*-questions and *wh*-declaratives

Targets, fillers and their preceding contexts as well were recorded by a female Mandarin speaker (age = 27) from Beijing in a sound-proof recording booth in the phonetics lab at Leiden University in Leiden, the Netherlands, with a sampling rate 44.1 kHz. To make sure that contexts affected target sentences evenly and listeners interpret target sentences as matrix clauses (instead of embedded clauses that are a continuation of the context), a 500 ms[33] silence was inserted between the offset of the context and the onset of the target or filler. Target sentences were manually segmented in Praat on the syllable level. We obtained the detailed acoustic properties of *wh*-questions and *wh*-declaratives in terms of duration on the word level and F0 on the syllable level. As we found that the two clause types differ in general in terms of word duration, we present the mean word duration across types, as presented in Figure 1[34].

[33] This length of pause did not affect the naturalness of sentences, according to the feedback of native speakers.

[34] We only presented duration information up to the *wh*-word *shénme*, as the prosodic cues of the post-*wh*-word region are not directly relevant for our study.

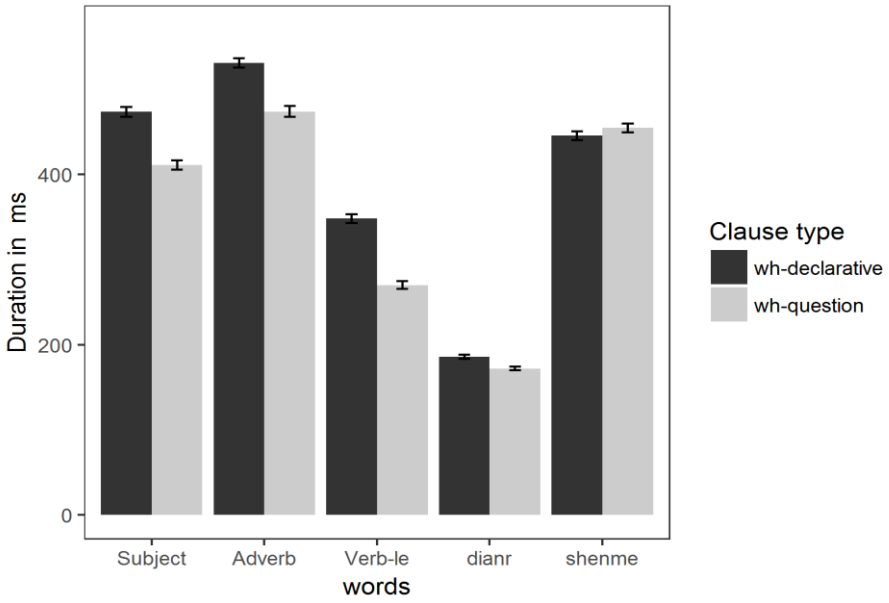


Figure 1. Mean word duration in ms with error bars showing standard error.

Linear mixed effects models in R were run using the lmerTest package (Kuznetsova, Brockhoff & Christensen, 2013), with each word duration as a dependent variable, clause type as a fixed effect factor, and item as a random factor. The results from linear mixed effects models showed that *wh*-declaratives are in general significantly longer than *wh*-questions except for the *wh*-word *shénme*; the acoustic properties of the two clause types are in general consistent with our production results in Chapter 2 from 34 speakers. The statistical results of each word duration are summarized in Table 1.

Table 1. Summary of the linear mixed effects models at words with significant duration differences between *wh*-questions and *wh*-declaratives.

	Estimate β	Std. Error	t-value	p-value
subject	62.448	5.095	12.257	< 0.001
adverb	57.069	4.301	13.269	< 0.001
verb-*le*	77.944	5.201	14.986	< 0.001
diǎnr	13.738	2.411	5.698	< 0.001

Further, the F0 difference between *wh*-questions and *wh*-declaratives are mainly found at the *wh*-word *shénme*. For a clearer illustration about the F0 difference at s*hénme* between the two clause types, we present an exemplar waveform and spectrogram of *wh*-declaratives and *wh*-questions together with their textgrid, as illustrated in Figure 2. As shown, *shénme* (highlighted) in *wh*-questions is boosted in F0 and expanded in F0 range as opposed to that in *wh*-declaratives.

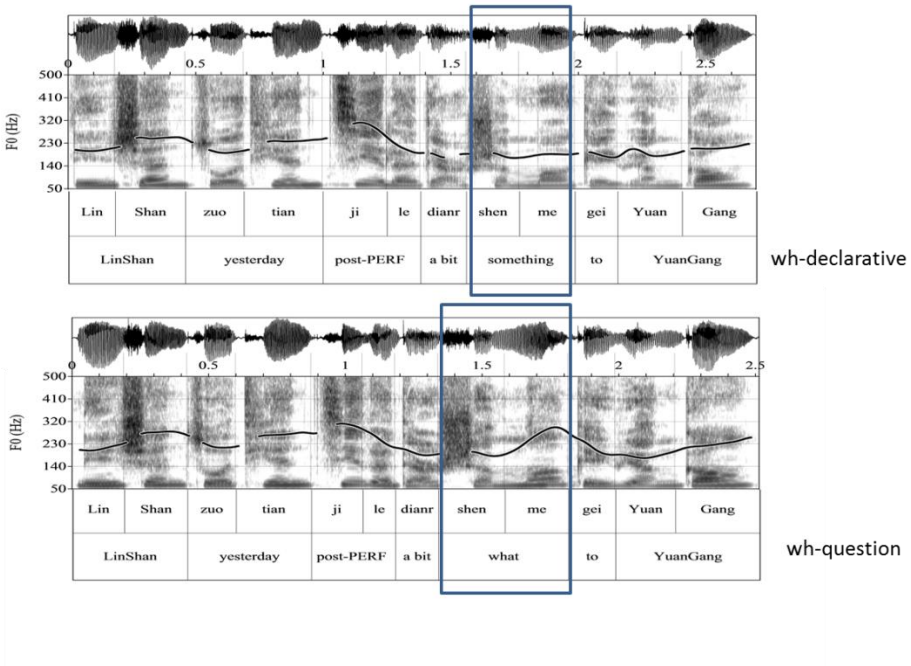


Figure 2. Exemplar waveforms and spectrograms of *wh*-declarative (9a) and *wh*-question (9b) with superimposed F0-contours, syllables and glosses, with *shénme* being highlighted.

In addition to the prosodic properties of the target sentences, we also investigated the properties of the final syllable in the context. Given that the final syllable (the verb) is T4 across all items and conditions (i.e. 道 (*dào*) 'know' in context (D) and 问 (*wèn*) 'ask' in context (Q)), for its F0 measurement we first obtained the F0-maximum and then the F0-minimum. Figures 3 and 4 present the average duration of the final syllable and its F0 measurement in the two contexts. Their acoustic differences in the final syllable between the two clause types are very trivial.

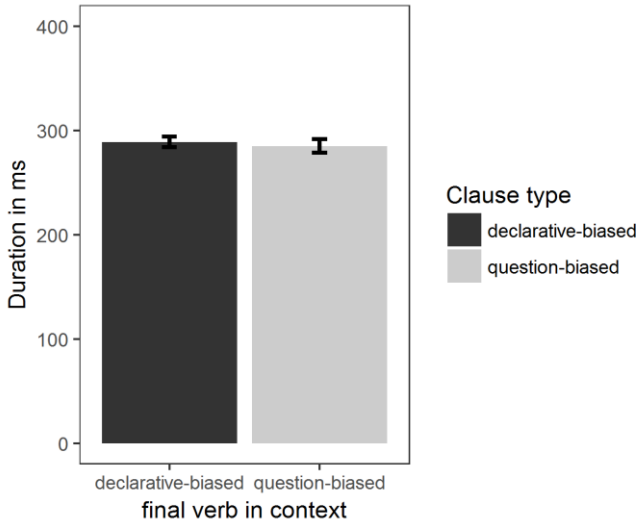


Figure 3. Mean word duration in ms of the final syllable in the context with error bars showing standard error.

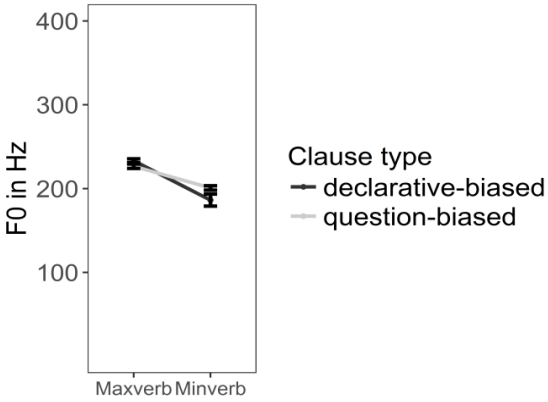


Figure 4. Stylized means of F0 curves of the final syllable in the context with error bars showing standard errors.

In short, *wh*-questions and *wh*-declaratives differ significantly at the *wh*-word *shénme* in F0 and differ significantly before the *wh*-word mainly in duration. These prosodic differences are aligned with results of our production study in Chapter 2. They serve as a baseline for the cross-spliced conditions in (9c) and (9d). For a clear illustration of cross-splicing of *wh*-questions and *wh*-declaratives from the *wh*-word *shénme* onwards, we present the two cross-spliced conditions D-Q as in (9c) and Q-D as in (9d) in Figure 5.

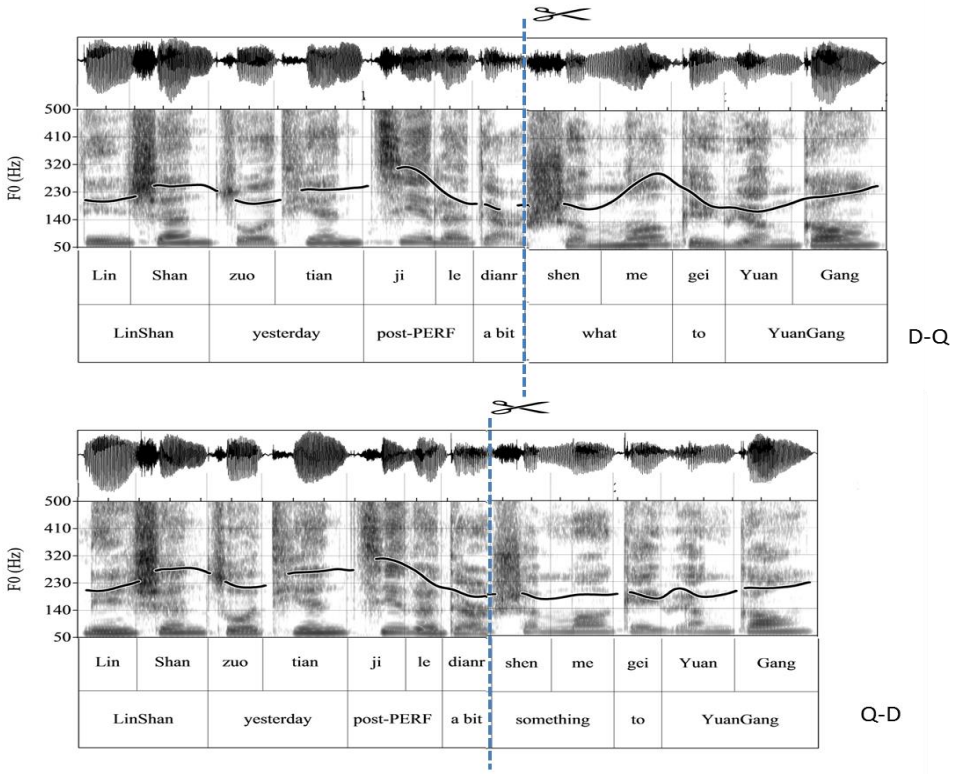


Figure 5. Exemplar waveforms and spectrograms of the two cross-spliced conditions D-Q (9d) and Q-D (9d) with superimposed F0-contours, syllables and glosses; starting from the scissors and the vertical lines, the audios are cross-spliced.

6.2.3 The predictions of the experiment

In this experiment, the *wh*-word *shénme* is the critical word, and we set a trigger at the onset of *shénme* for each target sentence. Based on the ERP components on prosody-related processing that we discussed in section 6.1.1, the predictions of the experiment are as follows: by comparing the cross-spliced incongruent condition D-Q as in (9c) with congruent *wh*-declaratives as in (9a) and by comparing the incongruent condition Q-D as in (9d) with *wh*-questions as in (9b), we expect to find prosodic mismatch effects such as (early) prosodic negativities at the critical word *shénme* (Magne et al., 2007; Mietz et al., 2008) or RAN effects (Eckstein & Friederici, 2005; Honbolygo, Torok, Banreti, Hunyadi & Csepe, 2016). In addition to effects related to prosody, we may also find N400 effects (Li, Yang & Hagoort, 2008) in (9c) and (9d), if the semantic violation based on prosody affects the lexical-semantic integration: for instance, when the *wh*-word *shénme* in condition (9c) with high pitch and expanded pitch range is associated with the interrogative

interpretation ('what') is in violation with the expected non-question interpretation ('something').

Alternatively, we may find asymmetry in detecting the two types of incongruities (D-Q in (9c) and Q-D in (9d)). For instance, we may find the neural correlates of detection of prosodic incongruity in one condition and not in the other, if listeners expect or accommodate the two clause types (D and Q) differently. If we can really find an asymmetry in the detection of incongruity, this asymmetry further supports that the participants make use of prosodic cues to detect the incongruity of clause types. Remember that in both incongruities (D-Q vs. Q-D) the unexpected pitch or duration are present; if we only find effects in one, it indicates that the detection is not just a superficial detection of incongruity based on acoustic cues alone (i.e. unexpected pitch or duration only) but a detection in reference to clause types.

6.2.4 Procedure

Participants sat in a comfortable chair approximately 90 cm in front of a computer screen in a dim and sound-proof room at an ERP lab in the Psychology department of Tianjin Normal University. Auditory stimuli were played from two loudspeakers placed on either side of the screen using the presentation software PsychoPy. A fixation cross (" + ") remained at the center of the screen before, during and after the auditory presentation of all stimuli/trials. To be specific, the highlighting of the fixation cross (changing to a bigger size) indicated the start of a trial. Each trial was separated by a 2 ms interval. Participants were instructed to keep their eyes fixated on the cross during the auditory presentation, to not move their head and body and to blink as little as possible.

Participants were instructed to listen to all trials attentively for comprehension. To make sure that participants were attentive during the testing, about one third of the trials were followed by a comprehension question (80 out of 288 trials) and around a quarter of the trials were followed by a question regarding arithmetic operations (55 out of 288 trials). The comprehension questions fall on both target sentences and contexts to avoid participants from guessing the aim of the experiment. The comprehension question or arithmetic operation remained at the center of the screen, and the two answers ('Yes' or 'No' in comprehension questions and two numbers in arithmetic operation questions) were also given below the question, on the left and on the right. Participants were instructed to respond to the question by pressing the "F" key on a keyboard for the answer on the left or the "J" key for the answer on the right with the keyboard. The answers were counter-balanced on the screen to avoid any left or right preference by participants.

The total 288 trials each participant received were presented in four blocks each containing 72 trials in total. More specifically, of the 288 trials, the 144 target trials were divided into four blocks using a Latin Square design so that there were 36 target trials (9 × 4 conditions) in each block. The 144 filler trials were also divided into four blocks with each block having 36 filler trials including 18 declaratives, 9 yes-no questions and 9 *wh*-adjunct questions, preceded by their contexts. In each block, the target trials and filler trials were randomized. Participants could take a break as long as they wanted between blocks. Prior to the experiment, each participant received 12 practice trials. The whole experiment lasted for

approximately 2 hours excluding electrode conduction preparation and participants' hair cleaning.

6.2.5 EEG recording

Continuous EEGs were recorded from 64 electrodes in a secured elastic cap (Neuroscan, Quick~cap, Ag/AgCl) localized at the following sites: AF7, AF3, FP1, FPZ, FP2, AF4, AF8, F7, F5, F3, F1, Fz, F2, F4, F6, F8, FT7, FC5, FC3, FC1, FCZ, FC2, FC4, FC6, FT8, T7, C5, C3, C1, CZ, C2, C4, C6, T8, TP7, CP5, CP3, CP1, CPZ, CP2, CP4, CP6, TP8, P7, P5, P3, P1, PZ, P2, P4, P6, P8, PO7, PO5, PO3, POZ, PO4, PO6, PO8, CB1, O1, Oz, CB2 and O2. EEGs on these electrodes were referenced online to the left mastoid and were re-referenced offline to the mean of the left and right mastoids. The vertical electro-oculogram (VEOG) was recorded from electrodes placed above and below the left eye. The horizontal EOG (HEOG) was recorded from electrodes placed at the outer cantus of left and right eyes. The biosignals were amplified with a band pass from 0.016 to 100 Hz and digitized on-line with a sampling frequency of 250 Hz.

6.2.6 ERP analysis

ERPs were filtered with 45-55 Hz band rejection to eliminate the signals from the fixed-frequency electricity followed by a 0.1-30 Hz high and low pass filtering[35]. The ocular artifacts were corrected automatically, with both VEOG and HEOG as common reference and blink detection by algorithms implemented in Brain Vision Analyzer™. ERPs were computed for each condition, electrode site, and participant. Trials containing excessive movement artifacts (mean voltage exceeding ± 150 μV) were excluded before averaging. The overall rejection rate was 4.9% and the mean rejection rate for each condition was 5.2% (SD = 0.084) for condition (9a), 4.3% for condition (9b) (SD = 0.065), 5.1% (SD = 0.080) for condition (9c) and 4.8% (SD = 0.084) for condition (9d).

Analyses were based on the ERPs extracted from the onset of the critical word *shénme*. For baseline correction, we used a baseline covering 100 ms pre-critical-word-onset. Subsequent analyses were based on 800 ms-epochs post critical word onset.

6.2.7 Behavioral results

The overall response accuracy rate for all behavioral tasks was 95.6% (SD = 0.204); the comprehension questions have an overall accuracy of 94.1% (SD = 0.236) and the arithmetic questions have an overall accuracy of 98.0% (SD = 0.142). The high accuracy rate shows that participants were attentive in listening and in responding to the comprehension and arithmetic questions.

[35] See *Brain Vision analyzer user manual* 10.1.4 for the detailed discussions of band-rejection filters.

6.2.8 ERP results

We investigated ERPs starting from the onset of the critical word *shénme*. Four-way repeated measures analyses of variance (ANOVAs) were conducted with the following within-subjects factors: congruity (congruent, incongruent) of target sentences, pre-*wh*-word clause type (question, declarative), hemisphere (left, middle and right) and position (anterior, central and posterior). Crossing the factors of hemisphere and position produced nine regions of interest (ROI), each with 6, 5, or 2 electrodes, including left anterior (F1, F3, F5, FC1, FC3, FC5), left central (C1, C3, C5, CP1, CP3, CP5), left posterior (P1, P3, P5, PO3, PO5), middle anterior (FZ, FCZ), middle central (CZ, CPZ), middle posterior (PZ, POZ), right anterior (F2, F4, F6, FC2, FC4, FC6), right central (C2, C4, C6, CP2, CP4, CP6) and right posterior (P2, P4, P6, PO4, PO6).

Opting for a non-biased analysis, omnibus ANOVAs were performed repeatedly using sliding 200 ms long windows to localize the effect with respect to the onset time of *shénme*. We checked all the main effects and interactions of congruity and reported the significant ones. As illustrated in Table 2, there is a three-way interaction of Position, Congruity and Pre-*wh*-word clause type in the time window of 200-400 ms, F (2, 34) = 3.543, $p < 0.05$. Table 3 provides the follow-up interaction evaluation for the time window with significant interactions observed in Table 2. Results showed that in the Anterior position, there is an interaction between Congruity and Pre-*wh*-word clause type, F (1, 17) = 6.913, $p < 0.05$. Table 4 shows a summary of the main effects and pair-wise comparisons in the breakdown of the significant interactions observed in Table 3 and results show a significant main effect of congruity when the pre-*wh*-word clause type is declarative, F (1, 17) = 5.408, $p < 0.05$. In other words, in the time window of 200-400 ms and at the Anterior position, significantly different ERPs were found between the *wh*-declarative condition as in (9a) and the incongruent condition D-Q as in (9c), where the pre-*wh*-word clause type is declarative.

Table 2. Multiple window ANOVA three-way interaction results (p-value reported)

	Time window (ms)					
	100-300	200-400	300-500	400-600	500-700	600-800
Position × Congruity × Pre-wh-word clausetype	0.160	0.040*	0.412	0.843	0.894	0.635

*$p < 0.05$

Table 3. Interactions broken down by Position.

		Time window (ms)
		200-400
Anterior	Congruity	0.432
	Hemisphere × Congruity	0.554
	Congruity × Pre-wh-word clausetype	0.018*
	Hemisphere × Congruity × Pre-wh-word clausetype	0.364
Central	Congruity	0.981
	Hemisphere × Congruity	0.800
	Congruity × Pre-wh-word clausetype	0.087
	Hemisphere × Congruity × Pre-wh-word clausetype	0.169
Posterior	Congruity	0.710
	Hemisphere × Congruity	0.785
	Congruity × Pre-wh-word clausetype	0.687
	Hemisphere × Congruity × Pre-wh-word clausetype	0.446

*$p < 0.05$

Table 4. Simple comparisons in Anterior position.

		Time window (ms)
		200-400
Pre-wh-word clausetype		
Declarative	Congruity	0.033*
	Hemisphere × Congruity	0.622
Question	Congruity	0.406
	Hemisphere × Congruity	0.360

*$p < 0.05$

Figure 6 shows the grand average ERPs at 9 electrodes from the 9 ROIs with the conditions in (9a) and (9c) that showed significant differences; 200-400 ms time windows in the ROIs from the Anterior position are highlighted. Figure 7 shows the topographic distributions of the mean ERP differences between (9a) and (9c) at the 200-400 ms time window.

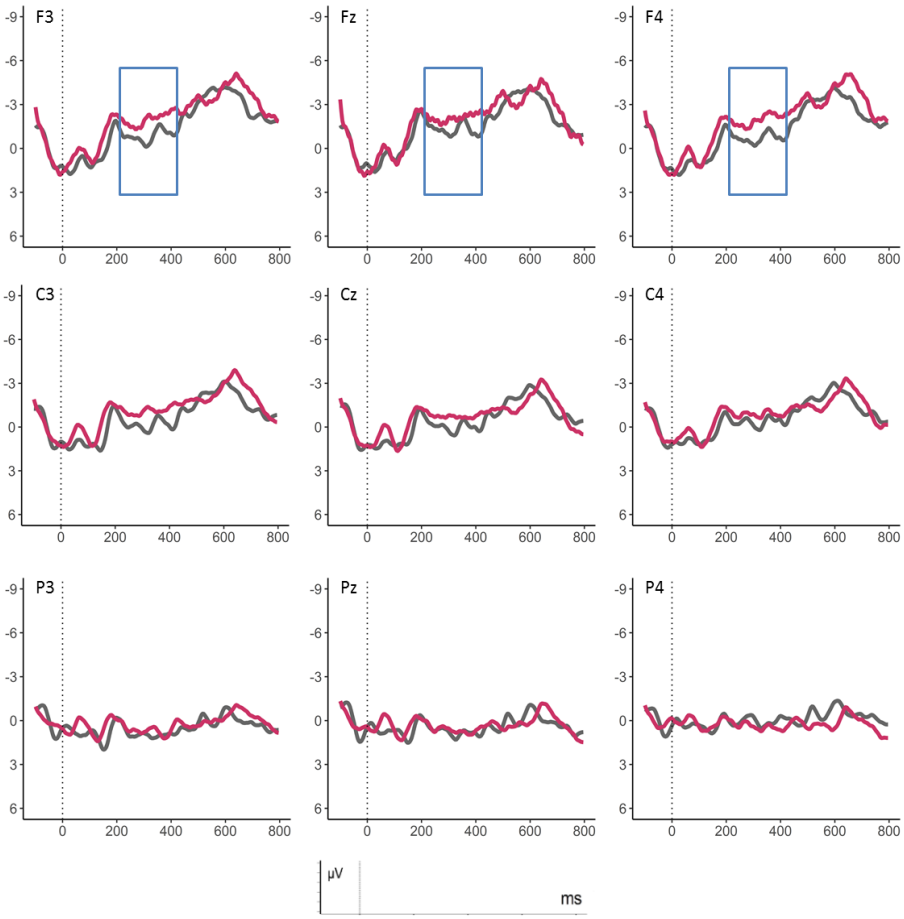


Figure 6. Grand average ERPs at 9 exemplar electrodes for each ROI time-locked to the onset of the critical word *shénme* (condition (9a) *wh*-declarative is in black, and condition (9c) (D-Q) in red).

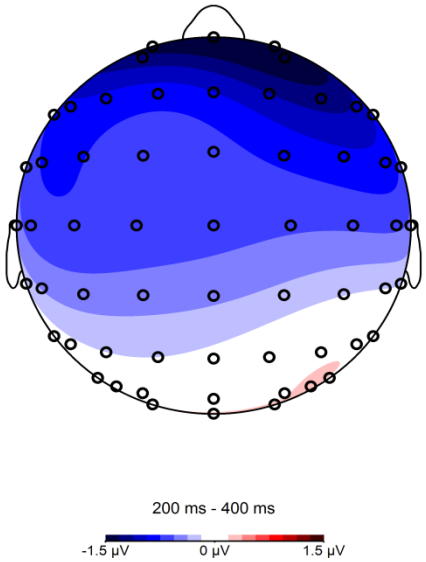


Figure 7. Topographic distributions of the mean ERP differences at the 200-400 ms time window. Condition (9c) (D-Q) was compared with condition (9a) (*wh*-declarative).

As Figures 6 and 7 show, the clause type incongruent condition (9c) (D-Q) as compared with the congruent condition (9a) *wh*-declarative evoked an anterior negativity in the 200-400 ms time window. Opting for a clearer illustration, we plotted average ERPs of the 14 electrodes in the anterior regions (F1, F3, F5, FC1, FC3, FC5, FZ, FCZ, F2, F4, F6, FC2, FC4, FC6), as presented in Figure 8.

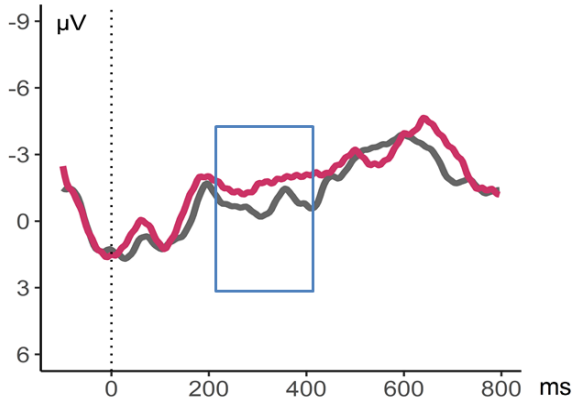


Figure 8. Grand average ERPs at anterior position time-locked to the onset of the critical word *shénme* (condition (9a) *wh*-declarative in black and condition (9c) (D-Q) in red).

It is worth noting that the anterior negativity in the 200-400 ms time window exists between congruent condition and incongruent condition only when the pre-*wh*-word clause type is a declarative, namely between (9a) *wh*-declarative and (9c) (D-Q), while no significant ERP differences were found between condition (9b) *wh*-question and condition (9d) (Q-D). Recall that our research question 3) concerns whether the detection of Mandarin clause type incongruity based on prosody is the same between *wh*-questions and *wh*-declaratives. The existence of an anterior negativity in (9c) (D-Q) but its absence in (9d) (Q-D) seems to imply that the mechanism of clause type detection based on prosody is different between the two clause types. This issue will be discussed in detail in section 6.4.

6.2.9 Interim discussion

The ERP results we found are in general consistent with our predictions in 6.2.3. By comparing congruent conditions of *wh*-declaratives and *wh*-questions with the incongruent cross-spliced conditions D-Q and Q-D, we found that the prosodically incongruent *wh*-word *shénme* in condition D-Q in (9c) evoked an anterior negativity (200-400 ms time window) when compared with the congruent declarative condition (9a), indicating an early detection of the clause type prosodic incongruity. The prosodic negativity is in general consistent with the results of other auditory ERP studies reporting the detection of prosodic incongruities (Magne et al., 2007; Mietz, 2008, among others), although the topographical distribution of the effect is not exactly the same. It further demonstrates that prosodic cues like F0 at *shénme* were immediately used in clausal typing when the prosodic violation is from the direction of a declarative (D) to a question (Q), that is, when a declarative prosody is expected and a question prosody is heard instead. Nevertheless, no significant effects were found in the opposite direction Q-D in (9d) when compared to the congruent question condition (9b) at *shénme*, indicating an asymmetry in detecting the incongruities.

We know from the acoustic properties of our own stimuli described in section 2.2.2 that *wh*-declarative and *wh*-questions also differ in the pre-*wh*-word region. Although the prosodic differences (mainly duration) at the beginning of the sentence are less salient than those at *shénme*, our audio-gating study in Chapter 3 demonstrated that listeners are able to utilize the prosodic differences and to anticipate the clause type even when hearing only the subject. Our next research question thus arises, which is how early in the sentence we can find the effects of prosody in clause type anticipation, in other words, how early clausal typing based on prosody can take place in the sentence during processing. To address this question, and to investigate whether participants can make use of the "less salient" prosodic cues at the beginning of the sentence in anticipating clause types based on the electrophysiological evidence, we conducted Experiment 2, which is reported in the next section.

6.3 Experiment 2 Contexts-*wh*-sentences incongruities

6.3.1 Participants

Twenty-four students from Tianjin Normal University were paid to participate in the experiment and none of them participated in similar experiments before. The participants were right-handed native speakers of Mandarin Chinese and were paid to participate in the experiment. They were undergraduate students with no known history of vision or hearing impairment or any cognitive or psychiatric disorder. Prior to testing, each participant gave written informed consent. Five participants were later excluded from further data analysis due to excessive eye or head movements or a failure in concentrating on the experiment according to the experimenter's observation and the participants' self-report. The remaining 19 participants (14 female) have a mean age of 21 years (SD = 2.19).

6.3.2 Materials

The materials we used as a basis for creating conditions are the same as in section 6.2.2.1. We combined the two contexts (question-biased context and declarative-biased context) and the two target sentences (*wh*-declarative and *wh*-question), yielding four conditions altogether, as illustrated in example (10). The target sentences in incongruent conditions (10c-d) are highlighted in bold for a clearer illustration. More specifically, the incongruent condition represents the instances where the clause type predicted by context is a declarative (D), listeners heard a *wh*-question (Q) instead as in (10c), marked as D-Q, while when the clause type predicted by context is a question (Q), listeners heard a *wh*-declarative (D) instead (marked as Q-D) in 10 (d). Here condition D-Q or Q-D only indicates the incongruenty of context and target sentences, not the cross-splicing of the target sentences as in Experiment 1.

(10) a. Context (D) *Wh*-declarative (D) (D-D)

(D)Context (D)林珊 昨天 寄了 点儿 什么 给 袁刚。
LínShān zuótiān jì-le diǎnr shénme gěi YuánGāng.
LinShan yesterday post-PERF a.little something to YuanGang
'Lin Shan posted a little something to Yuan Gang yesterday.'

b. Context (Q) *Wh*-declarative (Q) (Q-Q)

(Q)Context (Q)林珊 昨天 寄了 点儿 什么 给 袁刚?
LínShān zuótiān jì-le diǎnr shénme gěi YuánGāng?
LinShan yesterday post-PERF a.little what to YuanGang
'What did Lin Shan post a little to Yuan Gang yesterday?'

c. Context (D) *Wh*-question (Q) (D-Q)
(D)Context (**Q**)**林珊 昨天 寄了 点儿 什么 给 袁刚?**
LínShān zuótiān jì-le diǎnr shénme gěi YuánGāng？
LinShan yesterday post-PERF a.little what to YuanGang
'What did LinShan post a little to YuanGang yesterday?'

d. Context (Q) *Wh*-declarative (D) (Q-D)
(Q)Context (**D**)**林珊 昨天 寄了 点儿 什么 给 袁刚。**
LínShān zuótiān jì-le diǎnr shénme gěi YuánGāng.
LinShan yesterday post-PERF a.little something to YuanGang
'LinShan posted a little something to YuanGang yesterday.'

The subject, the verb and the *wh*-word *shénme* in (10) are the three critical words in the design and we set a trigger at the onset of each critical word. Altogether, we created 144 target sentences preceded by their contexts (36 items × 4 conditions). We used the same 144 filler sentences with their contexts as in Experiment 1. Altogether, each participant was presented with 288 trials.

6.3.3 Predictions of the experiment

In the current experiment, we try to localize where exactly in a sentence can we find ERP correlates for clausal type anticipation based on prosody during oneline sentence processing. As our previous perception study in Chapter 3 showed, clause type anticipation seems to take place early in the sentence (i.e. subject), thus, our predictions are as follows: at the critical word subject, by comparing the context-target incongruent condition D-Q as in (10c) with the congruent condition D-D as in (10a), and by comparing the incongruent condition Q-D as in (10d) with the congruent condition Q-Q as in (10b), we expect to find prosodic mismatch effects such as (early) prosodic negativities (Magne et al., 2007; Mietz et al., 2008) or RAN effects (Eckstein & Friederici, 2005; Honbolygo et al., 2016) at the first critical region, i.e., the subject. At the other critical two regions, i.e., verb and *shénme*, the predictions are twofold. 1) If as speech unfolds, listeners find that the incongruent clause type becomes "expected", it is less likely to keep on evoking similar effects as those elicited at the subject position. In other words, once they detect the incongruous prosody at the subject, they might not react to the incongruous prosody in the upcoming constituents any more as this prosody is somehow expected. 2) If, on the contrary, listeners still find the upcoming incongruous prosody unexpected as speech unfolds, we expect to find similar prosodic negativities as in the subject position at the verb and *shénme* position in conditions (10c) and (10d).

Again, we may find asymmetry in detecting the two conditions of incongruities (D-Q in (10c) and Q-D in (10d)) at above positions, if participants expect or accommodate clause types differently.

6.3.4 Procedure

The procedure is the same as in Experiment 1. The total 288 trials (144 target trials

and 144 filler trials) each participant received were presented in four blocks each containing 72 trials in total (36 target trials and 36 filler trials). The 144 target trials were divided into four blocks following a Latin Square design. Prior to the experiment, each participant also received 12 practice trials. The whole experiment lasted for approximately 2 hours excluding electrode conduction preparation and participants' hair cleaning.

6.3.5 EEG recording

The EEG recordings followed the same protocol as in Experiment 1.

6.3.6 ERP analysis

EEGs were additionally filtered with 45-55 Hz band rejection and 0.1-30Hz high and low pass. The ocular artifacts were corrected automatically, with both VEOG and HEOG as common reference and blink detection by algorithms implemented in Brain Vision Analyzer™. ERPs were computed for each condition, electrode site, and participant. Trials containing excessive movement artifacts (mean voltage exceeding ±150 μV) were excluded before averaging. The overall rejection rate was 6.9% and the mean rejection rate for each condition was 7.5% (SD = 0.135) for condition (a) D-D, 6.9% (SD = 0.113) for condition (b) Q-Q, 5.8% (SD = 0.119) for condition (c) D-Q and 7.3% (SD = 0.099) for condition (d) Q-D.

Analyses were based on the ERPs extracted from the onset of each critical word. For baseline correction, we used a baseline covering 100 ms pre-critical-word-onset. Subsequent analyses were based on 800 ms-epochs post onset of the critical words.

6.3.7 Behavioral results

The overall response accuracy rate was 95.8% (SD = 0.200); the comprehension questions had an overall accuracy of 94.1% (SD = 0.236) and the arithmetic questions had an overall accuracy of 98.3% (SD = 0.129). The high accuracy rate demonstrates that participants were attentive in listening and in responding to the comprehension and arithmetic questions.

6.3.8 ERP results

We investigated the ERPs at 3 positions, namely, the subject position, the verb position and the *wh*-word *shénme*. Below we report their ERP results respectively.

At the subject position

We first investigated ERPs starting from the onset of the subject, since this is the first possible identification point for the prosodic incongruity. Four-way repeated measures analyses of variance (ANOVAs) were conducted with the following within-subjects factors: congruity between context and target sentence (congruent, incongruent), context (question, declarative), hemisphere (left, middle and right) and position (anterior, central and posterior). Crossing the factors of hemisphere and region produced nine regions of interest (ROI), each with 6, 5, or 2 electrodes,

including left anterior (F1, F3, F5, FC1, FC3, FC5), left central (C1, C3, C5, CP1, CP3, CP5), left posterior (P1, P3, P5, PO3, PO5), middle anterior (FZ, FCZ), middle central (CZ, CPZ), middle posterior (PZ, POZ), right anterior (F2, F4, F6, FC2, FC4, FC6), right central (C2, C4, C6, CP2, CP4, CP6) and right posterior (P2, P4, P6, PO4, PO6).

Opting for a non-biased analysis, omnibus ANOVAs were performed repeatedly using sliding 200 ms long windows to localize the potential effect with respect to the onset time of the subject, as illustrated in Table 5. Although no four-way or three-way interaction was found, there is a two-way interaction between hemisphere and congruity in the time window of 300-500 ms, F (2, 36) = 3.310, $p < 0.05$.

Table 5. Multiple window ANOVA two-way interaction results (p value reported).

	Time window (ms)					
	100-300	200-400	300-500	400-600	500-700	600-800
Hemisphere × Congruity	0.558	0.219	0.048*	0.088	0.122	0.220

* $p < 0.05$

We then investigated the interactions and main effects of congruity in 300-500 ms time window in the breakdown of hemisphere, as summarized in Table 6.

Table 6. Main effects and interactions broken down by hemisphere.

		Time window (ms)
		300-500
Hemisphere		
Left	Congruity	0.178
	Position × Congruity	0.884
	Congruity × Context	0.082*
	Position × Congruity × Context	0.921
Middle	Congruity	0.116
	Position × Congruity	0.411
	Congruity × Context	0.264
	Position × Congruity × Context	0.792
Right	Congruity	0.912
	Position × Congruity	0.421
	Congruity × Context	0.486
	Position × Congruity × Context	0.644

* $p < 0.1$

Given that there is a marginal significant interaction between congruity and context in the left hemisphere, F (1, 18) = 3.400, $p < 0.1$, we then further investigated the effects in the breakdown of context, as summarized in Table 7.

Table 7. Main effects and Interactions in the left hemisphere broken down by context.

		Time window (ms)
		300-500
Context		
Declarative	Congruity	0.031*
	Position × Congruity	0.926
Question	Congruity	0.716
	Position × Congruity	0.870

* $p < 0.05$

The statistical analysis of the 300-500 ms time window revealed significant differences between congruent and incongruent conditions only when the context is a declarative, F (1, 18) = 5.443, $p < 0.05$. In other words, the ERP differences only exist in 300-500 ms time window between condition (10a) D-D and (10c) D-Q. In contrast, no differences were found between condition (10b) Q-Q and (10d) Q-D.

Figure 9 illustrates the grand average ERPs at 9 electrodes from the 9 ROIs with the conditions that showed significant differences, namely, condition (10a) D-D and (10c) D-Q; 300-500 ms time windows in the ROIs from the left hemisphere are highlighted. Figure 10 shows the topographic distributions of the mean ERP differences at the 300-500 ms time window between the two condtions, D-D as in (10a) and D-Q as in (10c), respectively.

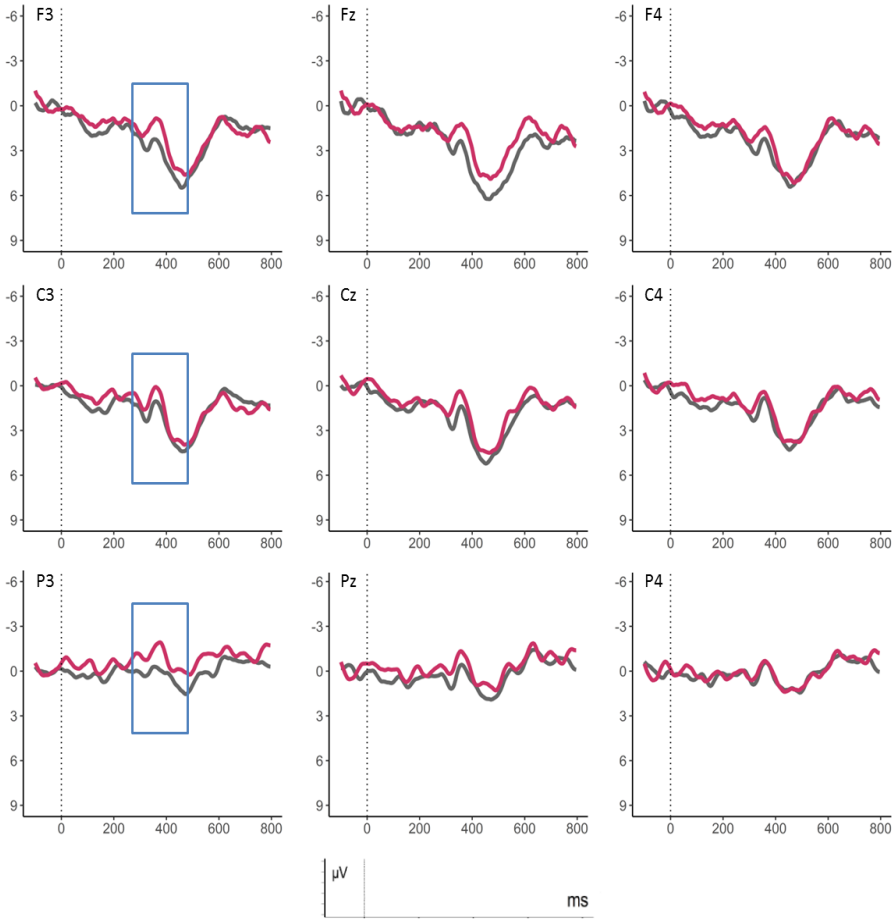


Figure 9. Grand average ERPs at 9 exemplar electrodes time-locked to the onset of the subject between condition (10a) D-D in black, and condition (10c) D-Q in red.

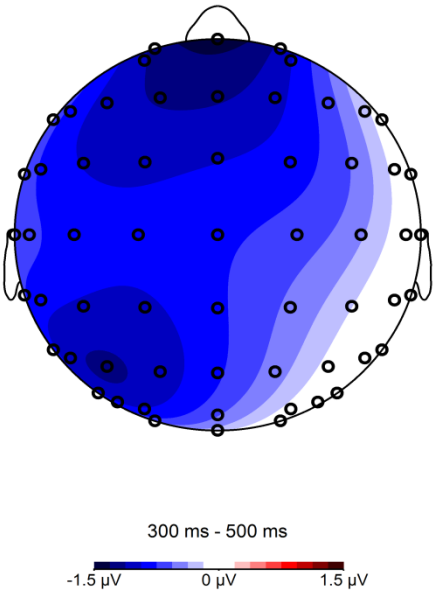


Figure 10. Topographic distributions of the mean ERP differences at the 300-500 ms time window. Condition (10c) (D-Q) was compared with condition (10a) (D-D).

As presented in Figures 9 and 10, a left-lateralized negativity in the 300-500 ms time-window was evoked at the subject in the incongruent condition (10c) D-Q as compared with the congruent condition (10a) D-D. We plotted average ERPs of all the electrodes in the left-hemisphere (F1, F3, F5, FC1, FC3, FC5, C1, C3, C5, CP1, CP3, CP5, P1, P3, P5, PO3, PO5), as shown in Figure 11.

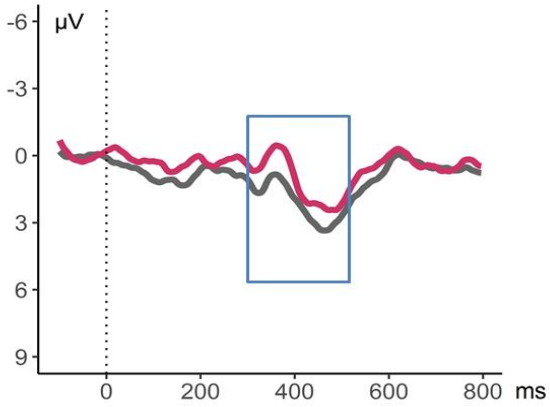


Figure 11. Grand average ERPs for left hemisphere electrode sites time-locked to the onset of the subject with condition (10a) D-D in black and condition (10c) D-Q in red.

At the verb position

We then investigated the ERPs starting from the onset of the verb. Given that before the critical word (verb) the prosody is already different between (10a) and (10c) and between (10b) and (10d), to keep the pre-critical word prosodically the same, we used different pairs of comparisons, namely, comparing the incongruent condition (10d) with the congruent condition (10a), and comparing the incongruent condition (10c) with the congruent condition (10b). Then four-way repeated measures analyses of variance (ANOVAs) were conducted with the following within-subjects factors: congruity (congruent, incongruent), target clause type (question, declarative), hemisphere (left, middle and right) and position (anterior, central and posterior).

Omnibus ANOVAs were performed repeatedly using sliding 200 ms long windows to localize any potential effect with respect to the onset time of the verb as illustrated in Table 8. Although no four-way or three-way interactions were found, there is a two-way interaction between hemisphere and congruity in the time window of 100-300 ms, F (2, 36) = 3.696, $p < 0.05$.

Table 8. Multiple window ANOVA two-way interaction results (p value reported).

	Time window (ms)					
	100-300	200-400	300-500	400-600	500-700	600-800
Hemisphere × Congruity	0.035*	0.124	0.115	0.257	0.259	0.454

* $p < 0.05$

We then investigated the effects and interactions of congruity in 100-300 ms time window in the breakdown of hemisphere, nevertheless, no significant main effects or interactions were found.

At *shénme* position

Finally, we investigated ERPs starting from the onset of *shénme*. The steps we took to carry out the statistical analysis are the same as those carried out for the critical word verb. Omnibus ANOVAs were performed repeatedly using sliding 200 ms long windows to localize the potential effect with respect to the onset time of *shénme*. Table 9 summarizes the omnibus ANOVAs. Although no four-way interactions were found, there is a three-way interaction of position, congruity and the target clause type in the time window of 200-400 ms, F (2, 36) = 6.350, $p < 0.05$ and 300-500 ms time windows, F (2, 36) = 4.263, $p < 0.05$; also, there is a three-way interaction of hemisphere, congruity and the target clause type in the time window of 600-800 ms, F (2, 36) = 3.861, $p < 0.05$.

Table 9. Multiple window ANOVA three-way interaction results (p value reported)

	Time window (ms)					
	100-300	200-400	300-500	400-600	500-700	600-800
Position × Congruity × Target	0.063	0.004**	0.022*	0.361	0.965	0.966
Hemisphere × Congruity × Target	0.875	0.626	0.140	0.175	0.142	0.030*

* $p < 0.05$, ** $p < 0.01$

When we further broke down the interactions by position or by hemisphere in those time windows, however, there were no significant main effects or interactions of congruity.

6.3.9 Interim discussion

Consistent with our predictions, by comparing context-target incongruent conditions D-Q (10c) with congruent conditions of D-D (10a), we found a left-lateralized negativity in the 300-500 ms time-window. The results are in general consistent with other auditory ERP studies that have detected prosodic incongruity (Eckstein & Friederici, 2006; Li, Chen & Yang, 2011), although the topographical distribution of the negativity is a bit different. This negativity can be interpreted as the detection of the prosodic incongruity in the clause type with respect to the predictions of the context. It further implies that even at the subject position, where only very limited prosodic cues are given, participants of our ERP study indeed utilize prosody to predict the clause type; they detected the incongruity between the question prosody they hear at the subject and the declarative prosody they expected based on the previous context.

It is worth noting that similar to Experiment 1, we observed ERP effects, only when the prosodic violation is in the direction of D-Q as in condition (10c). The prosodic negativity exists only when listeners were expecting a declarative while they heard a question instead. No significant effects were found in the opposite direction Q-D as in condition (10d), which again implies that clause type detection based on prosody is not the same for these two clause types. The discussion about this asymmetry in processing will be continued in the next section.

It should be noted that we observed the incongruity effects at the subject position, while at the other critical regions, namely the verb and *shénme*, we did not find any effects. The absence of effects at the verb and *shénme* is interpreted as follows. In our experiment, the auditory stimuli were presented continuously. In continuous speech, we assume that the prosodic incongruity is relatively easier to be perceived when the unexpected prosody is encountered for the first time (i.e. at the subject). As speech unfolds, the incongruent clause type becomes "expected", as it is congruent with the prosody of the subject, and hence it is less likely to keep on evoking similar effects as the subject position does at upcoming positions such as the verb and the *wh*-word.

6.4 Discussion

By conducting two auditory ERP experiments on *wh*-questions and *wh*-declaratives preceded by contexts and their manipulated incongruent conditions, we found negative components in the incongruent condition D-Q. More specifically, we found that in the cross-spliced condition D-Q as in (9c) in Experiment 1, the incongruent *wh*-word *shénme* elicited an early frontal negativity in the 200-400 ms time-window when compared to (9a); in the D-Q context-target incongruent condition as in (10c) in Experiment 2, the unexpected subject of the *wh*-question elicited a left-lateralized negativity in the 300-500 ms time-window when compared to (10a).

The direct conclusions we may draw based on our ERP experiments are as

follows. First, the processing time course on the detection of the prosodic incongruity in clausal typing shows an early time window (i.e., 200-400 ms) in Experiment 1 and a slightly later time window (i.e., 300-500 ms) in Experiment 2, different from the very late time window in Astésano et al. (2004). Second, listeners are actively making use of prosodic cues and detecting the incongruity of clause type prosody as early as the beginning of a clause.

Furthermore, there are three issues connected to the results that merit further consideration. First, we discovered that the detection of the prosodic incongruity is not the same between *wh*-questions and *wh*-declaratives, addressing our research question 3) raised in section 1. In particular, the incongruity is only detected electrophysiologically when listeners expect a declarative while a question prosody is heard, but not vice-versa. Further, this asymmetry occurs both in the cross-spliced cases in Experiment 1 and in the context-target incongruent cases in Experiment 2. As stated in section 6.2.3, although one might assume that the detection of prosodic incongruity is not necessarily equal to the detection of clause type incongruity, this asymmetry in the detection of incongruity actually demonstrates that participants indeed make use of prosodic cues in detecting the incongruity of clause types. This asymmetry provides evidence that the negative components observed in D-Q conditions of both experiments did not result from a superficial detection of incongruity based on acoustic aspects alone (i.e. unexpected pitch or duration only), since in both incongruities (D-Q vs. Q-D) the unexpected pitch or duration is present whereas we only observed negative components in D-Q conditions. Actually, the negative components we observed are understood as a reflection of participants' detection of the clausal type incongruences based on prosodic properties.

We hypothesize that the presence of the negative components in D-Q and its absence in Q-D can be attributed to the prosodic marking differences between *wh*-questions and *wh*-declaratives. In general, Mandarin declaratives are quite fixed in their intonation. Mandarin declarative intonation was described by Chao (1968) as normal intonation and has a succession of tones with a tendency for the pitch to trail off to a lower key towards the end (Chao, 1968: 40). This is consistent with most languages: sentences bearing the meaning of completion, termination, finality or assertion are associated with a low or falling pitch (Ladd, 2008). In contrast, *wh*-questions in many languages including Mandarin may not necessarily be marked with the typical question prosody (i.e. high F0). As Shen (1990) generalized and Ni and Kawai (2004) supported, the interrogative intonation of *wh*-questions may end with a low pitch, similar to declaratives. In other words, when questions have a non-typical question prosody, they may still be perceived as *wh*-questions, especially when the context or other information is biased towards questions. This allows us to explain why listeners did not react to the incongruent *wh*-declarative prosody when they expected a *wh*-question; they probably accommodated the *wh*-declarative prosody to a non-typical *wh*-interrogative one and processed the incongruent Q-D condition as a congruent Q-Q condition.

The second point concerns the different time-course of the negativities observed in the two experiments. In Experiment 1, *shénme* in the cross-spliced condition D-Q as in (9c) elicited an anterior negativity in the 200-400 ms time-window, while in Experiment 2, the subject in the context-target incongruent condition D-Q as in (10c) elicited a left-lateralized negativity in the 300-500 ms time-window. Although

both negativities can be understood as violations on a particularly expected prosody, their different latencies may be attributed to the different prosodic cues detected in the critical regions between the two clause types.

To be more precise, the detection of the incongruity in Experiment 1 is based on the F0 difference of the *wh*-word *shénme*, which showed early and salient F0 differences between *wh*-questions and *wh*-declaratives, as reported in section 6.2.2. F0 has been identified as the primary acoustic correlate of both tone and intonation in Mandarin (Ho, 1977; Wu, 1982; Xu & Wang, 2001), hence, the incongruent F0 is a reliable cue for detecting the incongruent prosody. Furthermore, the F0 differences between the two clause types start from the onset of the *wh*-word *shénme*. This is illustrated in Figures 12 and 13, where the stylized mean F0 at the syllables *shén* (S8) and *me* (S9) is shown.

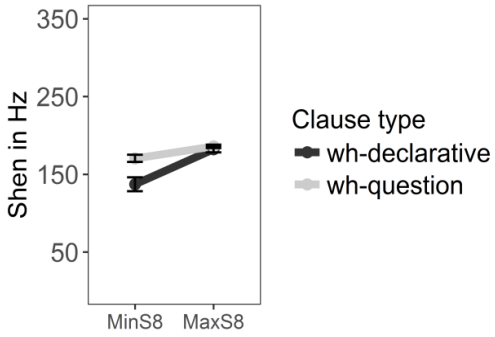


Figure 12. Stylized mean F0 at *shén* between *wh*-declaratives and *wh*-questions with error bars showing standard errors.

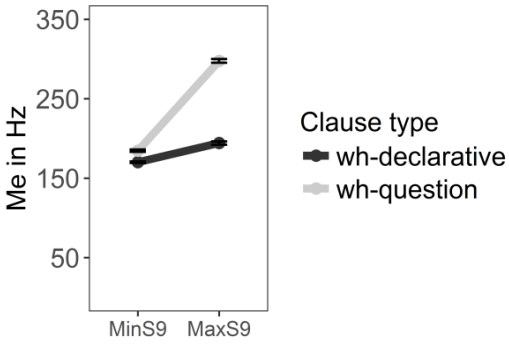


Figure 13. Stylized mean F0 at *me* between *wh*-declaratives and *wh*-questions with error bars showing standard errors.

Conversely, the detection of the incongruity at the subject in Experiment 2 is based on less salient prosodic differences between *wh*-questions and *wh*-declaratives (in the audio-gating study, the clause type anticipation accuracy by hearing the subject

only is 54.6% for *wh*-declaratives and 59% for *wh*-questions) and the start of prosodic differences is relatively later than that at *shénme*. As reported in the prosodic properties of the subject in section 6.2.2, *wh*-questions differ from *wh*-declaratives in terms of duration and the F0 of the second syllable[36]. This is shown in Figures 14 and 15, where S1 and S2 represent the two syllables in the subject. As illustrated in Figure 13 where the mean duration of S1 and S2 in the subject is portrayed, S1 lasts 197 ms in *wh*-questions and 225 ms in *wh*-declaratives. This means that in the first 197 ms after the onset of the subject, listeners heard the same segment and it is very hard to detect any duration difference between *wh*-questions and *wh*-declaratives, which would explain why the negative effects observed at the subject position in Experiment 1 starts later than that at *shénme* in Experiment 2.

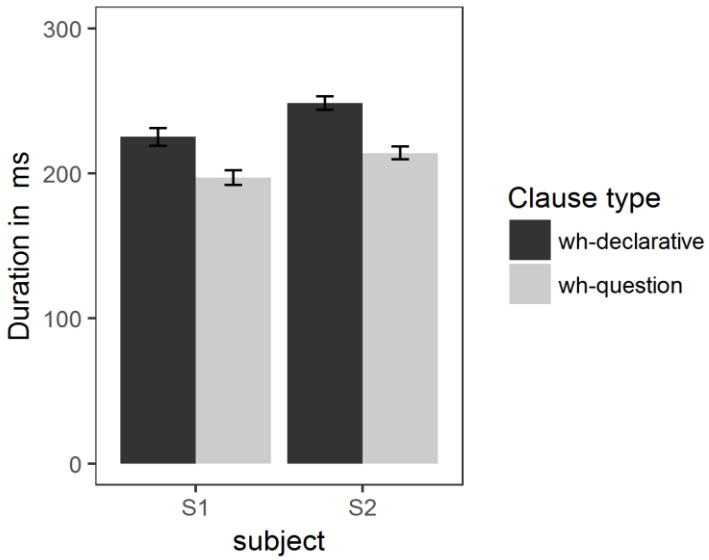


Figure 14. Mean syllable duration of the subject in ms with error bars showing standard errors between *wh*-declaratives and *wh*-questions.

[36] Different from the *wh*-word *shénme* which is consistent across items and conditions, subjects vary across items of target sentences.

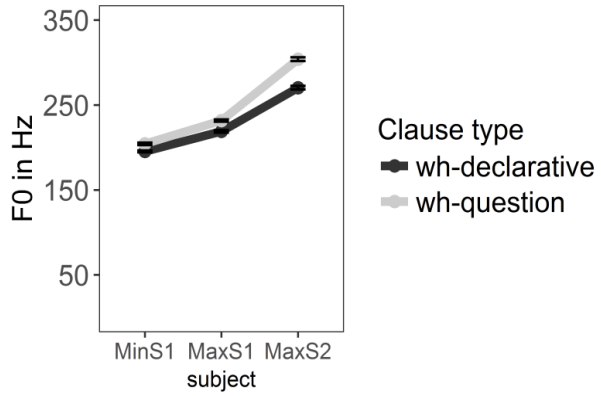


Figure 15. Stylized mean F0 at the subject between *wh*-declaratives and *wh*-questions with error bars showing standard errors.

Further, as illustrated in Figure 15 which shows the stylized mean F0 at the subject, the F0 difference between clause types starts after S1, confirming that in the first 197 ms after the subject's onset, the prosodic differences between clause types did not emerge yet, and hence could not be detected at the subject position until a later time window.

The detailed timing with respect to when prosodic differences emerge can thus account for the relatively early prosodic negativity at *shénme* in Experiment 1, and the later prosodic negativity at the subject in Experiment 2. In short, negativities can indicate the detection of prosodic incongruity, reflecting the detection of clausal typing incongruity, and the difference in their latencies depends on when precisely the given prosodic cues are enough for listeners to make clause type distinctions.

The third point concerns the fact that in both experiments, we only find (early) prosodic negativities, while no other effects such as CPS, N400 or P600 effects have been found. The lack of CPS in both experiments is not surprising, as neither the cross-splicing nor context-target incongruities leads to the violation of prosodic boundaries, as shown in Figure 5. Furthermore, boundary-elicited CPS is often found in relatively big prosodic boundaries such as intonation phrase boundary as in Steinhauer et al. (1999) and Honbolygo et al. (2016), marked with pauses or pre-boundary lengthening; hence in our target sentences (mono-clause only), it is less plausible to detect CPS effects. Moreover, we did not find any N400 effects or P600 effects in the two experiments. As introduced above, N400 effects and P600 effects often indicate the process of lexical-semantic integration and structural reanalysis respectively (Kutas & Hillyard, 1980; Friederici, 2002). One of the possibilites on the absence of these two effects is that participants can accommodate the further interpretation or integration of the sentence after perceiving the prosodic incongruities. To be specific, participants are very sensitive to the incongruity of the clause type based on prosody (as reflected in the ERP effects), however, participants can still accommodate the incongruities and continue their original processing without too much difficulty; to note that in our study the *wh*-sentences preceded by *diǎnr* can in principle have both *wh*-question and *wh*-declarative clause types and

hence the accommodation may be easier for participants. Alternatively, one may interpret the lack of N400 effect and P600 effect in the prosodic incongruent conditions as a kind of blocking of further semantic integrations or structural reanalysis brought by the difficulty/failure of prosody; this line of thinking follows Friederici, Gunter, Hahne and Mauth (2004) that a failure of an earlier stage processing may block the following processings. We leave the assessment of these possibilities and the resolution for the future studies.

Before closing this chapter, we summarize our findings and the significance of our studies briefly. Localizing temporally the clausal typing incongruity detection based on prosodic cues, our ERP studies revealed early negativities (200-400 ms) or relatively early negativities (300-500 ms) when expecting *wh*-declaratives while the prosody of questions are given (i.e. D-Q condition). To our knowledge, our study is the first electrophysiological study that explores the role of prosody on clausal typing in Mandarin during online processing. The online neural correlates (early negativities) and the detailed time course of the detection of the prosodic incongruities (200-400 ms or 300-500 ms) provide more fine-grained and direct evidence than any offline behavioral studies (e.g. audio-gating) on the effects of prosody on clausal typing. The results indicate that in terms of sentence processing, the prosodic information is immediately utilized online in the detection of the clause type incongruities of the sentence. In other words, prosody plays an immediate role on clausal typing as reflected in the detailed time course of the neural correlates.

In terms of clause type anticipation, the similar electrophysiological evidence found in the beginning of clauses (i.e. subject) in Experiment 2 as in the *wh*-word in Experiment 1 indicates that during online processing, participants can make efficient use of limited prosodic cues (i.e. duration information at the subject) they hear to anticipate clause types and detect the prosodic incongruities. Again, the electrophysiological evidence speaks more powerfully than a traditional study (i.e. audio-gating) that mainly provides an offline judgement of the clause type, and provides support for the claim that prosody plays an early role in clausal typing and clause type anticipation even at the clause initial position (i.e. subject).

Chapter 7 General Conclusion

This dissertation investigates the clausal typing in Mandarin and the licensing of Mandarin *wh*-indeterminates (*wh*-interrogatives and *wh*-existentials), by presenting diverse empirical evidence ranging from fine-grained prosodic studies to online sentence processing. In this chapter, I address the two research questions put forth in Chapter 1 by summarizing the main findings in each chapter. In section 7.1, I propose an extended clausal typing hypothesis in Mandarin. In section 7.2, I discuss the licensing environments of *wh*-existentials and *wh*-interrogatives. In the final section, I conclude this dissertation and discuss the relevance and implications of the current research for the future work.

7.1 The clausal typing mechanism

7.1.1 The main findings about clausal typing in Mandarin

As a *wh*-in-situ language, Mandarin has no *wh*-movement or reliable *wh*-particles to type clauses and hence we investigate whether prosody can function as a clausal typer. We scrutinize the prosodic analysis of *wh*-questions and *wh*-declaratives, which are string identical[37]. The results of the production study (Chapter 2) demonstrate that prosody is utilized in differentiating the two clause types in Mandarin. To be specific, *wh*-declaratives differ from *wh*-questions in terms of prosodic properties already from the clause onset in duration. *Wh*-declaratives are prosodically longer than *wh*-questions in terms of word duration starting from the subject and the pattern reverses at the *wh*-word; *wh*-declaratives are lower in F0 and smaller in F0 range than *wh*-questions at the *wh*-word and there is a F0 range compression in the post-*wh*-word region in *wh*-questions. Furthermore, the results of the perception and the audio-gating study (Chapter 3) show that listeners not only make use of the prosodic cues to identify the two clause types accurately but also anticipate clause types by only listening to the clause onset (e.g. the gate only containing sentence subject). It suggests that the prosodic marking of different clause types functions like *wh*-particles or *wh*-movement, which we can count on in making an assessment and prediction about the clause types.

The findings about the effects of prosody on clausal typing are further supported by electrophysiological evidence found in the two auditory ERP studies on *wh*-questions and *wh*-declaratives and their manipulated incongruent conditions (Chapter 7). The (early) negative components found in the incongruent conditions (Declarative-Question) of both experiments confirm that the prosodic cues are

[37] So far the findings that prosody can mark the clause type and thus listeners can identify and anticipate clause types are only based on the investigations of the ambiguous sentences, *wh*-questions and *wh*-declaratives respectively. One mighg wonder, will the same findings be extended to *wh*-questions and ordinary declarative containing noun phrases? Gryllia, Yang, Pablos, Doetjes & Cheng (submitted) has confirmed this by showing that prosody can differentiate *wh*-questions and declarative containing noun phrases and listeners can anticipate clause types based on prosody, just as in the case of *wh*-questions and *wh*-declaratives.

utilized in online sentence processing and play an important role in clausal typing. Furthermore, it provides evidence that even at the sentence beginning (i.e. subject), listeners can already detect the clause typing incongruity by utilizing the limited prosodic cues, confirming that the findings of the audio-gating studies in Chapter 3 about the clause type anticipation have a neural correlate basis.

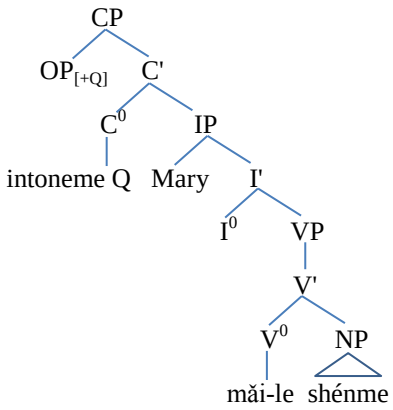
7.1.2 The extended clausal typing mechanism

According to Chomsky and Lasnik (1977: 445), "each clause must be identified as declarative (a declarative or relative clause) or interrogative (a direct or indirect question)." Given all the above findings on the effects of prosody on clausal typing in Mandarin, we make an extension based on the traditional Clausal Typing Hypothesis on *wh*-questions, repeated here "Every clause needs to be typed. In the case of typing a *wh*-question, either a *wh*-particle in C^0 is used or else fronting of a *wh*-word to the Spec of C^0 is used, thereby typing a clause through C^0 by spec-head agreement." (Cheng, 1991: 29). Our extended Clausal Typing Hypothesis is given in (1).

(1) **Extended clausal typing hypothesis**
Every clause needs to be typed. In the case of typing a *wh*-question, languages can have different ways to realize that. Languages can front a *wh*-word to the Spec of CP (*wh*-movement languages), or utilize a *wh*-particle (*wh*-in-situ languages with *wh*-particles), or/and utilize prosody to type the clause.

The above extended hypothesis raises another theoretical question. As prosody is often perceived as the PF level realization, how can it function as a clausal typer and be included in the early process before spell-out? We hypothesize that it is indeed in the process of narrow syntax that prosody starts to function and in this process, prosody is a relatively abstract concept that types the clause, without having all the details of the prosodic realizations as our production study shows. We term the clausal typer at this process as an "intonational Q-morpheme" following Cheng and Rooryck (2000) or simply an "intoneme" Q following Feng (2015) and Tang (2015). In the process of narrow syntax as represented in (2), we hypothesize that the intonational Q-morpheme/intoneme Q is inserted in C^0, and has Spec-head agreement with the interrogative operator at Spec-CP.

(2)



But only at the interface level (syntax-prosody), the rules coming from phonology will interact with the information syntax conveys (including intonational Q-morpheme /intoneme Q), leading to the well-formed PF realization/prosodic markings. Further studies are needed for the complete theoretical analysis on the intonational Q-morpheme/intoneme Q before spell-out.

7.2 The licensing of *wh*-existentials and *wh*-interrogatives

Wh-words in Mandarin Chinese are "*wh*-indeterminates" as they can have interrogative, existential, universal, or free choice interpretations. The interpretations of *wh*-words are restricted to certain contexts where there is a licensor that can license the specific interpretations. In this dissertation, I mainly investigate the licensing environments and licensing evidence of two frequently used interpretations of Mandarin *wh*-words, *wh*-existentials and *wh*-interrogatives respectively, and the results are summarized as below.

7.2.1 The licensing of *wh*-existentials in Mandarin

With respect to licensing of *wh*-indeterminates, Mandarin has no reliable particles like Japanese and Korean to license *wh*-indeterminates. Previous studies claim that Mandarin *wh*-existentials are only licensed by nonveridical contexts such as contexts containing negation or epistemic modality (Lin, 1998; Xie, 2007; Lin, Weerman & Zeijlstra, 2014; Huang, 2017). We provide evidence in Chapter 4 that *wh*-existentials can also be licensed in nonveridical contexts under certain conditions. To be specific, in Chapter 4 we find that if a *wh*-word in a sentence can be licensed as a *wh*-existential, the sentence normally has the following environments:

(3) i. Either the sentence contains a nonveridical operator (negation, questions, conditionals and epistemic modalities) that can license the *wh*-existential.
or ii. When the nonveridical contexts (i.e. containing future environments) are not able to license *wh*-existential or when the sentence is simply veridical, it needs a last resort licensor like *diănr*. *Diănr* provides the existential force in licensing the *wh*-existential in the VP when *diănr* cliticizes to the verb.

7.2.2 The licensing of *wh*-interrogatives in Mandarin

In Mandarin, when there is no overt licensor/operator, a sentence containing a *wh*-word is a *wh*-question with the in-situ *wh*-word licensed by the null interrogative operator (Q) at Spec-CP or C^0, and thus obtains the interrogative force. This covert dependency between the null interrogative operator (Q) and the in-situ *wh*-word is investigated in Chapter 5, in which three self-paced reading studies on Mandarin *wh*-questions as compared with declaratives with indefinite noun phrase are reported. By comparing *wh*-questions containing simplex *wh*-phrases ('who') with declaratives containing indefinites ('someone'), and by comparing *wh*-questions containing complex *wh*-phrases ('which x') with declaratives containing indefinites ('an x'), we provide processing evidence of constructing a covert dependency between the in-situ *wh*-word and the interrogative operator (Q) at Spec-CP or C^o. The higher processing cost we find in *wh*-questions than in declaratives support the hypothesis that a covert dependency is needed between the *wh*-phrase and the interrogative operator (Q) at Spec-CP or C^o.

7.3 Conclusions and future work

This dissertation explores the licensing of *wh*-indeterminates in Mandarin, with a focus on the clausal typing of sentences containing *wh*-existentials and *wh*-interrogatives and the licensing of them, from the perspective of prosody and processing. The results of the production study and audio-perception/gating studies demonstrate that prosody can mark and type *wh*-questions, and listeners can anticipate clause types by utilizing prosody since the clause onset. Moreover, this dissertation uses electrophysiological measurement to investigate the fine-grained time course of clausal typing based on prosody during online sentence processing, the results of which show that prosodic information is immediately utilized in detecting clause type incongruities. On the basis of the empirical evidence, we propose an extended clausal typing hypothesis, explicitly recognizing the role of prosody in clausal typing and hence making the original clausal typing hypothesis more complete, cross-linguistically speaking.

With respect to the licensing of *wh*-existentials, this dissertation challenges the general assumption in previous studies that *wh*-existentials are licensed in nonveridical contexts only (Lin, 1998; Xie, 2007; Lin et al., 2014; Huang, 2017); we provide evidence and analysis that *wh*-existentials can be licensed in veridical contexts, when the preceding licensor of *diǎnr* (or *gè*) cliticizes to the verb as a clitic group. Regarding the licensing of *wh*-interrogatives/question words, the higher processing cost in reading *wh*-questions as compared with declaratives containing indefinites provides empirical evidence that a covert dependency is required between the *wh*-interrogative and the interrogative operator (Q) at Spec-CP (or C^o), in processing a *wh*-in-situ question in Mandarin. Cross-linguistically speaking, a dependency is always needed in interpreting a *wh*-question, be it a *wh*-in-situ question like in Mandarin or a *wh*-movement question like in English.

For future studies, it is important to investigate whether prosody also plays a role in the clausal typing of *wh*-questions in other *wh*-in-situ languages (e.g. Bangla,

Vietnamese, among others). It is also important to give a detailed theoretical analysis about how intonational Q-morpheme/intoneme Q takes part in the derivation process before the spell-out of a sentence. Furthermore, although this dissertation investigates the licensing of *wh*-existentials and *wh*-interrogatives, the methodology and the research perspective (prosody and processing) can be extended to the licensing of other interpretations of *wh*-indeterminates (e.g. universal and free-choice interpretations).

References

Aaronson, D., & Scarborough, H. S. (1976). Performance theories for sentence coding: Some quantitative evidence. *Journal of Experimental Psychology: Human Perception and Performance*, *2*(1), 56.

Allan, K. (1977). Classifier. *Language*. *53*, 285–311.

Aoun, J., & Li, A. Y. H. (1993). Wh-elements in-situ: syntax or LF? *Linguistic Inquiry*, *24*:199-238.

Avrutin, S. (2000). Comprehension of discourse-linked and non-discourse-linked questions by children and Broca's aphasics. *Language and the Brain: Representation and Processing*, 295-313.

Astésano, C., Besson, M., & Alter, K. (2004). Brain potentials during semantic and prosodic processing in French. *Cognitive Brain Research*, *18*(2), 172-184.

Baltazani, M. (2007). Intonation of polar questions and the location of nuclear stress in Greek. In Gussenhoven & Riad (ed.), *Tones and Tunes, 2* (pp.387-405). Berlin, Germany: Mouton de Gruyter.

Beach, C. M., Katz, W. F., & Skowronski, A. (1996). Children's processing of prosodic cues for phrasal interpretation. *The Journal of the Acoustical Society of America*, *99*(2), 1148-1160.

Boersma, P., & Weenink, D. (2016). Praat: doing phonetics by computer [Computer program]. Version 6.0.19, retrieved from http://www.praat.org/.

Bolinger, D. (1978). Intonation across languages. *Universals of Human Language, 2*, 471-524.

Bolker, B. M., Brooks, M. E., Clark, C. J., Geange, S. W., Poulsen, J. R., Stevens, M. H. H., & White, J. S. S. (2009). Generalized linear mixed models: a practical guide for ecology and evolution. *Trends in Ecology & Evolution*, *24*(3), 127-135.

Bögels, S., Schriefers, H., Vonk, W., & Chwilla, D. J. (2011). Prosodic Breaks in Sentence Processing Investigated by Event-Related Potentials. *Language and Linguistics Compass*, *5*(7), 424-440.

Braun, B., & Chen, A. (2012). Now for something completely different: Anticipatory effects of intonation. In O. Niebuhr (Ed.), *Understanding Prosody: The Role of Context, Function and Communication* (pp. 289-311). Berlin: de Gruyter.

Buxton, H. (1983). Temporal predictability in the perception of English speech. In A. Coulter & D. R. Ladd (Eds.), *Prosody: Models and Measurements* (pp. 111-121). Berlin: Springer.

Cangemi, F., & D'Imperio, M. (2013). Tempo and the perception of sentence modality. *Laboratory Phonology*, *4*(1), 191-219.

Carlson, K., Clifton Jr, C., & Frazier, L. (2001). Prosodic boundaries in adjunct attachment. *Journal of Memory and Language*, *45*(1), 58-81.

Chao, Y. R. (1932). A preliminary study of English intonation (with American variants) and its Chinese equivalents. *BIHP Supplement* No.1.

Chao, Y. R. (1968). *A Grammar of Spoken Chinese*. Berkeley: University of California Press.

Chen, A., den Os, E., & de Ruiter, J. P. (2007): Pitch accent type matters for online

processing of information status: Evidence from natural and synthetic speech. *The Linguistic Review*, *24*(2-3), 317-344.
Chen, M. Y. (2000). *Tone Sandhi: Patterns Across Chinese Dialects* (Vol. 92). Cambridge University Press.
Chen, S. H. (2005). The effects of tones on speaking frequency and intensity ranges in Mandarin and Min dialects. *The Journal of the Acoustical Society of America*, *117*(5), 3225-3230.
Chen, S. W., Wang, B., & Xu, Y. (2009). Closely related languages, different ways of realizing focus. In *Interspeech*, 1007-1010.
Chen, Y. (2010). Post-focus F0 compression—Now you see it, now you don't. *Journal of Phonetics*, *38*(4), 517-525.
Cheng, L. L. S. (1991). *On the Typology of WH-Questions*, PhD dissertation, MIT.
Cheng, L. L. S. (1994). "Wh-words as Polarity Items," in R. Li (Eds.), *Chinese Languages and Linguistics II*, (pp. 615-640). Taipei: Academia Sinica.
Cheng, L. L. S. (2009). On every type of quantificational expression in Chinese. In: M. Rather and A. Giannakidou (eds.), *Quantification, Definiteness,and Nominalization*, 53-75. Oxford: Oxford University Press.
Cheng, L. L. S., & Giannakidou, A. (2013). The non-uniformity of wh-indeterminates with polarity and free choice in Chinese. In K.-H. Gil, S. Harlow and G. Tsoulas (Eds.), *The Nature of Quantification* (pp. 123-151). Oxford: Oxford University Press.
Cheng, L. L. S., & Rooryck, J. (2000). Licensing wh-in-situ. *Syntax*, *3*(1), 1-19.
Cheng, L. L. S., & Sybesma, R. (1999). Bare and not-so-bare nouns and the structure of NP. *Linguistic Inquiry*, *30*, 509–542.
Cho, Y. Y. (1990). Syntax and phrasing in Korean. In S. Inkelas & D. Zec (Eds.), *The Phonology-syntax Connection* (pp.47-62). Chicago: University of Chicago Press.
Chomsky, N., & Lasnik, H. (1977). Filters and control. *Linguistic Inquiry*, *8*(3), 425-504.
Chuang, Y. Y., & Fon, J. (2010). The effect of prosodic prominence on the realizations of voiceless dental and retroflex sibilants in Taiwan Mandarin spontaneous speech. *Proceedings of Speech Prosody 2010*.
Constant, N. (2014). *Contrastive Topic: Meanings and Realizations*. PhD dissertation, University of Massachusetts, Amherst.
Cruttenden, A. (2006). The de-accenting of given information: A cognitive universal? *Empirical Approaches to Language Typology*, *20*(8), 311.
Dahan, D., Tanenhaus, M.K., & Chambers, C.G., (2002): Accent and reference resolution in spoken-language comprehension. *Journal of Memory and Language*, *47*, 292-314.
De Francis, J. (1963). *Beginning Chinese*. New Haven: Yale University Press.
Deguchi, M., & Kitagawa, Y. (2002). Prosody and Wh-questions. *NELS 32*, 73-92.
De Vincenzi, M. (1996). Syntactic analysis in sentence comprehension: Effects of dependency types and grammatical constraints. *Journal of Psycholinguistic Research*, *25*, 117-133.
Di Cristo, A. (1998). Intonation in French. In D. Hirst & A. Di Cristo (Eds.), *Intonation Systems: A Survey of Twenty Languages*. Cambridge: Cambridge University Press.

Dong, H. (2009). *Issues in the Semantics of Mandarin Questions*. PhD dissertation, Cornell University.

Donkers, J., Hoeks, J. C. J., & Stowe, L. A. (2013). D-linking or set-restriction? Processing Which-questions in Dutch. *Language and Cognitive Processes*, *28*(1-2), 9-28.

Duanmu, S. (2000). *The Phonology of Standard Chinese*. New York: Oxford University Press.

Duanmu, S. (2004). Tone and non-tone languages: An alternative to language typology and parameters. *Language and Linguistics*, *5*(4), 891-923.

Ducrot, O. (1972). *Dire et ne pas dire, Principes de Sémantique Linguistique*. Paris: Hermann.

Eckstein, K., & Friederici, A. D. (2005). Late interaction of syntactic and prosodic processes in sentence comprehension as revealed by ERPs. *Cognitive Brain Research*, *25*(1), 130-143.

Eckstein, K., & Friederici, A. D. (2006). It's early: event-related potential evidence for initial interaction of syntax and prosody in speech comprehension. *Journal of Cognitive Neuroscience*, *18*(10), 1696-1711.

Face, T. L. (2004). The intonation of absolute interrogatives in Castilian Spanish. *Southwest Journal of Linguistics*, *23*(2), 65-80.

Feng, S. (2015). Tone, intonation and sentence final particles in Chinese. *Essays on Linguistics*, *51*, 52-79.

Féry, C., & Kügler, F. (2008). Pitch accent scaling on given, new and focused constituents in German. *Journal of Phonetics*, *36*(4), 680-703.

Frazier, L., & Clifton Jr, C. (1989). Successive cyclicity in the grammar and the parser. *Language and Cognitive Processes*, *4*(2), 93-126.

Friederici, A.D. (2002). Towards a neural basis of auditory sentence processing. *Trends in Cognitive Sciences*, *6*(2), 78-84.

Friederici, A. D., Gunter, T. C., Hahne, A., & Mauth, K. (2004). The relative timing of syntactic and semantic processes in sentence comprehension. *NeuroReport*, *15*, 165-169.

Frota, S. (2002). Nuclear falls and rises in European Portuguese: a phonological analysis of declarative and question intonation. *Probus*, *14*(1), 113-146.

Gårding, E. (1987). Speech act and tonal pattern in Standard Chinese: constancy and variation. *Phonetica*, *44*(1), 13-29.

Gérard, C., & Clément, J. (1998). The structure and development of French prosodic representations. *Language and Speech*, *41*(2), 117-142.

Gernsbacher, M. A., & Jescheniak, J. D. (1995). Cataphoric devices in spoken discourse. *Cognitive Psychology*, *29*(1), 24-58.

Giannakidou, A., & Cheng, L. L.S. (2006). (In) definiteness, polarity, and the role of wh-morphology in free choice. *Journal of Semantics*, *23*(2), 135-183.

Grosjean, F. (1980). Spoken word recognition processes and the gating paradigm. *Perception & Psychophysics*, *28*(4), 267-283.

Gryllia, S., Yang, Y., Doetjes, J., & Cheng, L. L.S. (2016). *Prosody encodes clause type anticipation: evidence from Mandarin* [Powerpoints slides]. Talk given in Tone and Intonation in Europe (TIE), September 1-3, 2016, Kent.

Ho, A. T. (1977). Intonation variation in a Mandarin sentence for three expressions: interrogative, exclamatory and declarative. *Phonetica*, *34*(6), 446-457.

Honbolygó, F., Török, Á., Bánréti, Z., Hunyadi, L., & Csépe, V. (2016). ERP correlates of prosody and syntax interaction in case of embedded sentences. *Journal of Neurolinguistics*, *37*, 22-33.

Hu, F. (2002). A prosodic analysis of wh-words in Standard Chinese. In proceedings of *Speech Prosody2002*.

Heeren, W. F., Bibyk, S. A., Gunlogson, C., & Tanenhaus, M. K. (2015). Asking or Telling–Real-time Processing of Prosodically Distinguished Questions and Statements. *Language and Speech*, *58*(4), 474-501.

Huang, C. T. J. (1982). Move wh in a language without wh-movement. *Linguistic Review*, *1*(4). 369-416.

Huang, C. T. J. (1984). Phrase structure, lexical integrity, and Chinese compounds. *Journal of the Chinese Language Teachers Association*, *19*(2), 53-78.

Huang, H. L. (2017). On the role of classifiers in licensing Mandarin existential wh-phrases. To appear in *Proceedings of Sinn und Bedeutung 21*.

Iljic, R. (1994). Quantification in Mandarin Chinese: Two markers of plurality. *Linguistics*, *32*, 91–116.

Ishihara, S. (2002). Invisible but audible wh-scope marking: Wh-constructions and deaccenting in Japanese. In *Proceedings of WCCFL* (Vol. 21, pp. 180-193).

Ishihara, S. (2003). *Intonation and Interface Conditions*, PhD dissertation, MIT.

Jiang, P., & Chen, A. (2011). Representation of Mandarin intonations: boundary tone revisited. In Proceedings of 23rd North American Conference on Chinese Linguistics, (1), 97-109.

Jun, S. A., & Oh, M. (1996). A prosodic analysis of three types of wh-phrases in Korean. *Language and Speech*, *39*(1), 37-61.

Just, M. A., & Carpenter, P. A. (1980). A theory of reading: From eye fixations to comprehension. *Psychological Review*, *87*(4), 329.

Just, M. A., Carpenter, P.A., & Wooley, J.D. (1982). Paradigms and Processes in Reading Comprehension. *Journal of Experimental Psychology: General*, *111* (2), 228-238.

Karttunen, L. (1977). Syntax and semantics of questions. *Linguistics and Philosophy*, *1*(1), 3-44.

Kitagawa, Y., & Fodor, J. D. (2003). Default prosody explains neglected syntactic analyses of Japanese. *Japanese/Korean Linguistics*, *12*, 267-279.

Kjelgaard, M. M., & Speer, S. R. (1999). Prosodic facilitation and interference in the resolution of temporary syntactic closure ambiguity. *Journal of Memory and Language*, *40*(2), 153-194.

Kliegl, R., Grabner, E., Rolfs, M., & Engbert, R. (2004). Length, frequency, and predictability effects of words on eye movements in reading. *European Journal of Cognitive Psychology*, *16*(1-2), 262-284.

Kohn, K., & Kalina, S. (1996). The strategic dimension of interpreting. *Meta: Journal des raducteurs/Meta: Translators' Journal*, *41*(1), 118-138.

Kratzer, A., & Shimoyama, J. (2002). Indeterminate pronouns: The view from japanese. In Y. Otsu (Ed.), *Proceedings of the Third Tokyo Conference on Psychoinguistics, Tokyo*, pp.1-25. Hituzi Syobo.

Kuo, Y. C., Xu, Y., & Yip, M. (2007). The phonetics and phonology of apparent cases of iterative tonal change in Standard Chinese. In Gussenhoven & Riad (ed.), *Tones and Tunes*, *2*, 211-237.

Kuroda, S. Y. (1965). *Generative Grammatical Studies in the Japanese Language*. PhD dissertation, MIT.

Kutas, M., & Hillyard, S. A. (1980). Event-related brain potentials to semantically inappropriate and surprisingly large words. *Biological Psychology*, *11*(2), 99-116.

Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. B. (2013). lmerTest: Tests for Random and Fixed Effects for Linear Mixed Effect Models (Lmer Objects of lme4 Package). R Package Version 2.0–3[Computer Software].

Kwon, N., & Sturt, P. (2013). Null pronominal (pro) resolution in Korean, a discourse-oriented language. *Language and Cognitive Processes*, *28*(3), 377-387.

Ladd, D. R. (2008). *Intonational Phonology*, second edition. Cambridge: Cambridge University Press.

Lambrecht, K., & Michaelis, L. A. (1998). Sentence accent in information questions: Default and projection. *Linguistics and Philosophy*, *21*(5), 477-544.

Lee, O. J. (2005). *The Prosody of Questions in Beijing Mandarin*. PhD dissertation, The Ohio State University.

Li, A. (2002). Chinese prosody and prosodic labeling of spontaneous speech. *In proceedings of Speech Prosody 2002*.

Li, B. (2006). *Chinese Final Particles and the Syntax of the Periphery*. PhD dissertation, Leiden University.

Li, K. (2009). *The Information Structure of Mandarin Chinese: Syntax and Prosody*. PhD dissertation, University of Washington.

Li, Q., & Chen, Y. (2012).Trisyllabic tone sandhi in Tianjin Mandarin. *Proceedings of the third International Symposium on Tonal Aspects of Languages*.

Li, X., Chen, Y., & Yang, Y. (2011). Immediate integration of different types of prosodic information during on-line spoken language comprehension: An ERP study. *Brain Research*, *1386*, 139-152.

Li, X., Yang, Y., & Hagoort, P. (2008). Pitch accent and lexical tone processing in Chinese discourse comprehension: an ERP study. *Brain Research*, *1222*, 192-200.

Li, Y. H. A. (1990). *Order and Constituency in Mandarin Chinese*. Dordrecht, the Netherlands: Kluwer.

Li, Y. H. A. (1992). Indefinite wh in Mandarin Chinese. *Journal of East Asian Linguistics*, *1*(2), 125-155.

Lin, J. W. (1998). On existential polarity-wh-phrases in Chinese. *Journal of East Asian Linguistics*, *7*(3), 219-255.

Lin, J., Weerman, F., & Zeijlstra, H. (2014). Mandarin SHENME as a superweak NPI. *Black Book: a festschrift in honor of Frans Zwarts*, 229-251.

Lin, J. W. (2014). Wh-expressions in Mandarin Chinese. In A. Simpson, J. C. T. Huang, and A. Y. H. Li (Eds.), *Handbook of Chinese Linguistics*. Wiley Blackwell.

Lindsey, G. A. (1985). *Intonation and Interrogation: Tonal Structure and the Expression of a Pragmatic Function in English and other Languages*. PhD dissertation, UMI.

Liu, F. (2009). *Intonation Systems of Mandarin and English: A Functional Approach*. PhD Dissertation, University of Chicago.

Liu, X., Li, A., & Jia, Y. (2016). How does prosody distinguish wh-statement from wh-question? A case study of standard Chinese. *Proceedings of Speech Prosody 2016*, 1076-1080.

Lü, S. (1999). *Xiandai Hanyu Babai Ci* [800 Words in Chinese]. Beijing: Shangwu Press.

Magne, C., Astésano, C., Aramaki, M., Ystad, S., Kronland-Martinet, R., & Besson, M. (2007). Influence of syllabic lengthening on semantic processing in spoken French: behavioral and electrophysiological evidence. *Cerebral Cortex*, *17*(11), 659-2668.

Marslen-Wilson W.D., Tyler, L.K., Warren, P., Grenier, P., & Lee, C.S. (1992). Prosodic effects in minimal attachment. *The Quarterly Journal of Experimental Psychology*. 45, 73-87.

Mietz, A., Toepel, U., Ischebeck, A., & Alter, K. (2008). Inadequate and infrequent are not alike: ERPs to deviant prosodic patterns in spoken sentence comprehension. *Brain and Language*, *104*(2), 159-169.

Moltmann, F. (1997). Intensional verbs and quantifiers. *Natural Language Semantics*, *5*(1), 1-52.

Ni, J., & Kawai, H. (2004). Pitch targets anchor Chinese tone and intonation patterns. In *Proceedings of Speech Prosody 2004*.

Nishigauchi, T. (1990). *Quantification in the Theory of Grammar*. Dordrecht: Kluwer.

Osterhout, L., & Holcomb, P.J. (1992). Event-related brain potentials elicited by syntactic Anomaly. *Journal of Memory and Language*, *31*(6), 785-806.

Ohala, J. J. (1983). Cross-language use of pitch: an ethological view. *Phonetica*, *40*(1), 1-18.

Ohala, J. J. (1984). An ethological perspective on common cross-Language utilization of F0 of Voice. *Phonetica*, *41*(1), 1-16.

Ouyang, I. C., & Kaiser, E. (2015). Prosody and information structure in a tone language: an investigation of Mandarin Chinese. *Language, Cognition and Neuroscience*, *30*(1-2), 57-72.

Pauker, E., Itzhak, I., Baum, S. R., & Steinhauer, K. (2011). Effects of cooperating and conflicting prosody in spoken English garden path sentences: ERP evidence for the boundary deletion hypothesis. *Journal of Cognitive Neuroscience*, *23*(10), 2731-2751.

Pesetsky, D. (1987). Wh-in-situ: movement and unselective binding. *In The Representation of (In)definiteness*, E. J. Reuland, and A. ter Meulen (eds.), 98-129. Cambridge, MA: MIT Press.

Petrone, C., & Niebuhr, O. (2014). On the intonation of German intonation questions: The role of the prenuclear region. *Language and Speech*, *57*(1), 108-146.

Pinheiro, J. C., & Bates, D. M. (2000). Linear mixed-effects models: basic concepts and examples. *Mixed-effects Models in S and S-Plus*, 3-56.

Rayner, K., Sereno, S. C., Morris, R. K., Schmauder, A. R., & Clifton Jr, C. (1989). Eye movements and on-line language comprehension processes. *Language and Cognitive Processes*, *4*(3-4), SI21-SI49.

Richards, N. (2010). *Uttering Trees*. Cambridge: MIT Press.

Richards, N. (2016). *Contiguity Theory*. Cambridge: MIT Press.

Rullmann, H., & Beck, S. (1998). Presupposition projection and the interpretation of which-questions. *Semantics and Linguistic Theory*, *8*, 215-232.

Seeber, K. G. (2001). Intonation and anticipation in simultaneous interpreting. *Cahiers de Linguistique Française*, *23*, 61-97.

Shen, J. (1994). Intonation structure and intonation types of Chinese. *Dialect*, 4, 221-228.

Shen, X. N. S. (1990). *The Prosody of Mandarin Chinese*. CA: UC Press.

Shen, X. N. S. (1993). Relative duration as a perceptual cue to stress in Mandarin. *Language and Speech*, *36*(4), 415-433.

Shi, P. (1980). Intonation variations in four types of Mandarin sentences. *Language Teaching and Research*, *2*, 71-81.

Shiamizadeh, Z., Caspers, J., & Schiller, N. O. (2017a). The role of F0 and duration in the identification of wh-in-situ questions in Persian. *Speech Communication*, *93*, 11-19.

Shiamizadeh, Z., Caspers, J., & Schiller, N. O. (2017b). The role of prosody in the identification of Persian sentence types: declarative or wh-question?. *Linguistics Vanguard*, *3*(1).

Shields, J. L., McHugh, A., & Martin, J. G. (1974). Reaction time to phoneme targets as a function of rhythmic cues in continuous speech. *Journal of Experimental Psychology*, *102*(2), 250.

Shih, C. (1997). Mandarin third tone sandhi and prosodic structure. *Linguistic Models*, 20, 81-124.

Shin, J. Y. (2005). Wh-constructions in Korean: A lexical account. *Toronto Working Papers in Linguistics*, 25.

Snedeker, J., & Trueswell, J. (2003). Using prosody to avoid ambiguity: Effects of speaker awareness and referential context. *Journal of Memory and Language*, *48*(1), 103-130.

Speer, S.R., Kjelgaard, M.M. & Dobroth, K.M. (1996). The influence of prosodic structure on the resolution of temporary syntactic closure ambiguities. *Journal of Psycholinguistic Research*, *25*(2), 249-271.

Steinhauer, K. (2003). Electrophysiological correlates of prosody and punctuation. *Brain and Language*, *86*(1), 142-164.

Steinhauer, K., Alter, K., & Friederici, A. D. (1999). Brain potentials indicate immediate use of prosodic cues in natural speech processing. *Nature Neuroscience*, *2*, 191-196.

Stowe, L. A. (1986). Parsing wh-constructions: Evidence for on-line gap location. *Language and Cognitive Processes*, *1*(3), 227-245.

Sudo, Y. (2013). Biased polar questions in English and Japanese, in D. Gutzmann and H-M. Gaertner [eds], *Beyond Expressives: Explorations in Use-Conditional Meaning*. Leiden: Brill,

Sybesma, R. (1992). *Causatives and Accomplishments: The Case of Chinese Ba*. PhD dissertation, HIL/Leiden University.

Tai, J. H. Y. (1992). Variation in classifier systems across Chinese dialects: Towards a cognition-based semantic approach. Symposium Series of the Institute of History and Philology. *Academia Sinica*, *2*, 587–608.

Tai, J. H. Y. & Wang, L. Q. (1990). A semantic study of the classifier tiao. *Journal of the Chinese Language Teachers Association*, *25*, 35–56.

Tang, S. W. (2015). A generalized syntactic schema for utterance particles in Chinese. *Lingua Sinica*, *1*(1), 1-23.

Terken, J., & Hirschberg, J. (1994): Deaccentuation of words representing 'given' information: effects of persistence of grammatical function and surface position. *Language and Speech*, *37*(2), 125–145.

Tsai, W. T. D. (1994). *On Economizing the Theory of A-bar Dependencies*. PhD. Dissertation. MIT.

Tsai, W. T. D. (1999). On lexical courtesy. *Journal of East Asian Linguistics*, *8*(1), 39-73.

Tsai, W. T. D. (2010). On the Syntax-Semantics Correspondences of Chinese Modals. *Zhongguo Yuwen*, 3, 208-221.

Van Heuven, V. J., & Van Zanten, E. (2005). Speech rate as a secondary prosodic characteristic of polarity questions in three languages. *Speech Communication*, *47*(1-2), 87-99.

Vion, M., & Colas, A. (2006). Pitch cues for the recognition of yes-no questions in French. *Journal of Psycholinguistic Research*, *35*(5), 427-445.

Wakefield, J. C. (2016). Sentence-final particles and intonation: two forms of the same thing. In *Proceedings of Speech Prosody 2016*.

Warren, T., & Gibson, E. (2002). The influence of referential processing on sentence complexity. *Cognition*, *85*, 79–112.

Watanabe, A. (1992). Wh-in-situ, subjacency and chain formation (No. 2). *MIT Working Papers in Linguistics Department of Linguistics and Philosophy*, Massachusetts Institute of Technology.

Weber, A., Braun, B., & Crocker, M., (2006): Finding referents in time: eyetracking evidence for the role of contrastive accents. *Language and Speech*, *49*(3), 367-392.

Wu, F., Kaiser, E., & Andersen, E. (2012). Animacy Effects in Chinese Relative Clause Processing. *Language and Cognitive Processes*, *27*(10), 1489-1524.

Wu, Z. (1982). Rules of intonation in Standard Chinese. Papers for the Working Group on Intonation ICL, 13, 95-103.

Wu, Z. (1982). Putonghua yuju zhong de shengdiao bianhua [Tonal variations in Mandarin sentences]. *Zhongguo Yuwen*, 439-450.

Xiandai hanyu cidian (2005). *A Contemporary Chinese Dictionary*. Beijing: Commercial Publishing House.

Xiang, M. (2008). Plurality, maximality and scalar inferences: A case study of Mandarin Dou. *Journal of East Asian Linguistics*, *17*(3), 227-245.

Xiang, M., Dillon, B., Wagers, M., Liu, FQ., & Guo, T.M. (2013). Processing covert dependencies: An SAT study on Mandarin wh-in-situ questions. *Journal of East Asian Linguistics*, *23*(2), 207-232.

Xiang, M., Wang, S.P., & Cui, Y.L. (2015). Constructing covert Dependencies — the case of wh-in-situ processing. *Journal of Memory and Language*, *84*, 139-166.

Xie, Z. (2007). Nonveridicality and existential polarity wh-phrases in Mandarin. In *Proceedings from the Annual Meeting of the Chicago Linguistic Society*, *43*, 121–135.

Xu, B. R., & Mok, P. (2011). Final rising and global raising in Cantonese intonation. In *Poceedings of the Seventeenth International Congress of the Phonetic Sciences*. 2173- 2176.

Xu, Y. (1999). Effects of tone and focus on the formation and alignment of F0 contours. *Journal of Phonetics*, *27*(1), 55-105.

Xu, Y. (2001). Fundamental frequency peak delay in Mandarin. *Phonetica*, *58*(1-2), 26-52.

Xu, Y., Chen, S. W., & Wang, B. (2012). Prosodic focus with and without post-focus compression: A typological divide within the same language family?. *The Linguistic Review*. *29*(1), 131-147.

Xu, Y., & Wang, Q. E. (2001). Pitch targets and their realization: Evidence from Mandarin Chinese. *Speech Communication*, *33*(4), 319-337.

Yoshida, K., & Yoshida, T. (1997). Question marker drop in Japanese. *Language Research Bulletin*, *11*, 37-54.

Yuan, J. (2004). *Intonation in Mandarin Chinese: Acoustics, Perception, and Computational Modeling*. Unpublished doctoral thesis, Cornell University, NY.

Yuan, J. (2006). Mechanisms of question intonation in Mandarin. In *Chinese Spoken Language Processing* (pp.19-30). Berlin-Heidelberg, Germany: Springer.

Yun, J. (2012). The deterministic prosody of indeterminates. In *Proceedings of the 29th West Coast Conference on Formal Linguistics* (pp. 285-93).

Zhang, L. (2014). Segmentless sentence-final particles in Cantonese: An experimental study. *Studies in Chinese Linguistics*, *35*(2), 47-60.

Zhang, N. N. (2013). *Classifier Structures in Mandarin Chinese*. Berlin: Walter de Gruyter.

Zwarts, F. (1995). Nonveridical contexts. *Linguistic Analysis*. *25*(3-4), 286–312.

Summary

Mandarin is a *wh*-in-situ language, in which *wh*-words remain at their base position just as their declarative counterparts do, and hence Mandarin has no *wh*-movement to type the clause as a *wh*-question. Mandarin is also known to be a *wh*-indeterminate language, in which *wh*-words like *shénme* can have both interrogative ('what') and non-interrogative interpretations (e.g. existential interpretation, 'something'), but unlike other *wh*-indeterminates languages like Japanese and Korean, Mandarin has no particles to license the interpretations of the *wh*-indeterminates. Due to the *wh*-in-situ and *wh*-indeterminates nature of Mandarin, clausal typing in Mandarin and the licensing of Mandarin *wh*-indeterminates have long been two intriguing topics in the fields of syntax and semantics. Different from traditional studies, this dissertation investigates the licensing of Mandarin *wh*-indeterminates and the clausal typing of sentences containing them from the perspective of prosody and processing. To better explore the topics, a variety of methodologies are adopted in this dissertation, including production / acoustic study, audio-perception /gating paradigm, self-paced reading and electrophysiological means. These diversified empirical studies provide evidence for the role of prosody in clausal typing in Mandarin, based on which an extended clausal typing hypothesis is proposed; the licensing evidence and constraints of *wh*-interrogatives (e.g. 'what') and *wh*-existentials (when *wh*-words have existential interpretations, e.g. 'something') are also explored in Mandarin, the two most frequently used interpretations of *wh*-indeterminates.

This dissertation consists of seven chapters. **Chapter 1** briefly introduces the background of the main research issues to be discussed, lays out the research questions and provides an overview of the dissertation.

Chapter 2 investigates the prosodic markings of *wh*-questions and *wh*-declaratives (declaratives containing *wh*-words interpreted as *wh*-existentials, which are string identical to *wh*-questions in our study) through a production experiment on native Mandarin speakers from Beijing. The fine-grained analysis of the audio recordings demonstrates that *wh*-declaratives differ from *wh*-questions in terms of different prosodic properties such as duration, F0, F0 range and intensity range. It shows that already from the clause onset the two clause types already show prosodic differences. In particular, *wh*-declaratives are longer than *wh*-questions in terms of duration starting from the subject and the pattern reverses at the *wh*-word; *wh*-declaratives are lower in F0 and smaller in F0 range than *wh*-questions at the *wh*-word and there is a F0 range compression in the post-*wh*-word region in *wh*-questions. *Wh*-declaratives show larger intensity range than *wh*-questions at the verb and the pattern reverses at the *wh*-word. The production study in this chapter demonstrates that prosody is indeed utilized in clausal typing.

Based on the results of Chapter 2, **Chapter 3** further explores whether listeners can identify as well as anticipate the clause types (*wh*-questions and *wh*-declaratives) based on prosodic cues only. This chapter reports an audio-perception experiment, in which participants were asked to listen to a sentence (containing *wh*-questions or *wh*-declaratives) and complete a dialogue. The audio-perception study includes an

audio-gating experiment, where the audios of sentences (*wh*-questions or *wh*-declaratives) are segmented into different lengths of fragments / gates for listeners. The audio-gating experiment investigates whether participants can anticipate clause types before hearing the *wh*-word and if so, at which gate they succeed in doing that. The results of the audio-perception and gating experiment show that 1) Participants can make use of prosody to differentiate the two clause types; 2) Starting from the onset of the first word of the sentence (*wh*-question or *wh*-declarative), participants already demonstrate a preference for the clause type that was intended by the speaker. Moreover, the more listeners hear of the sentence, the more accurate their anticipation for the clause type becomes.

Chapter 4 discusses the licensing of the existential interpretation of *wh*-indeterminates (i.e., *wh*-existentials) in Mandarin. *Wh*-existentials have been considered to be licensed only in non-veridical contexts in previous literature, for instance, in contexts containing non-veridical operators such as negation, questions, conditionals and epistemic modalities, in which the truth of a proposition cannot be entailed in the sentence. In this chapter I challenge this claim, by demonstrating that Mandarin *wh*-existentials can also be licensed in veridical contexts. I focus on the investigation of veridical *wh*-sentences containing *diănr* 'a little' and discuss the licensing of *wh*-existentials with *diănr*. I conducted a sentence reading experiment where participants read veridical sentences containing *wh*-indeterminates with and without *diănr* and the results show that *wh*-sentences containing *diănr* are ambiguous between *wh*-questions and *wh*-declaratives. This provides empirical evidence that *wh*-existentials can also be licensed in veridical contexts. In order to further shed light on the licensing conditions of *diănr* on *wh*-existentials, a special gate in the gating study was designed. The result of this gating study shows that listeners are reluctant to interpret a *wh*-word as a *wh*-existential when the verb and *diănr* are separated into different prosodic domains.

I discuss the licensing of Mandarin *wh*-interrogatives in **Chapter 5**. I conducted self-paced reading studies on Mandarin simplex *wh*-questions (containing simplex *wh*-words such as 'who') in comparison with their declarative counterparts containing 'someone', and on complex *wh*-questions (containing *wh*-words such as 'which x'; also known as discourse-linked *wh*-words) in comparison with their declarative counterparts containing 'an x'. The results show that questions are processed with a higher cost than declaratives, providing empirical evidence for covert dependency between the in-situ *wh*-word and the Spec-CP or C^0 domain. This study also implies that complex *wh*-questions are processed differently from simplex *wh*-questions due to the discourse-linking property of the former.

In **Chapter 6** I report the results of two ERP experiments and provide electrophysiological evidence for the role of prosody in clausal typing during online sentence processing. As the audio-perception study (Chapter 3) shows that listeners can identify *wh*-questions and *wh*-declaratives accurately by hearing the *wh*-word, in the first ERP Experiment, I cross-splice audio-recordings of *wh*-questions and *wh*-declaratives from *wh*-words onwards and test *wh*-declaratives (D), *wh*-questions (Q) and their cross-spliced conditions D-Q and Q-D, all preceded by contexts biasing towards the clause type of the pre-*wh*-word region. Furthermore, as our audio-gating study (Chapter 3) shows that listeners can anticipate clause types by hearing the sentence beginning (e.g. subject), the second ERP experiment manipulates the

congruity between contexts biasing *wh*-questions (Q) / *wh*-declaratives (D) and target sentences of *wh*-questions (Q) / *wh*-declaratives (D). Electrophysiological results of both experiments reveal (early) prosodic negativities, indicating an early detection of clause type incongruity based on prosody during online senence processing; but the negativities are only found in the incongruent condition of D-Q, which implies a different accommodation in processing the incongruities.

Lastly, **Chapter 7** returns to the research questions raised in Chapter 1 and summarizes the main findings of the dissertation. At the end of this chapter, I conclude this dissertation and provide suggestions for future work.

Samenvatting

Het Mandarijn is een *wh*-in-situ taal, waar vraagwoorden op hun oorspronkelijke positie blijven, net zoals hun declaratieve tegenhangers doen. Daarom heeft het Mandarijn geen vraagwoordverplaatsing om een zin als een vraagzin te markeren. Het Mandarijn staat ook bekend als een vraagwoord-onbepaalde taal, waarin vraagwoorden zoals *shénme* zowel een vraagwoordsinterpretatie ('wat') alsook een niet-vraagwoordelijke interpretatie (bijv. een existentiële interpretatie, 'iets'). Echter, in tegenstelling tot andere vraagwoord-onbepaalde talen zoals het Japans en het Koreaans, heeft het Mandarijn geen partikels om de interpretatie van deze zogenaamde *wh*-indeterminates te fiatteren. Vanwege de *wh*-in-situ en *wh*-indeterminate kenmerken van het Mandarijn, zijn zinsbepaling en de fiattering van *wh-indeterminates* al lange tijd intrigerende onderwerpen in de syntaxis en de semantiek. Anders dan traditionele studies, onderzoekt deze dissertatie de fiattering van Mandarijnse *wh*-indeterminates en de zinsbepaling van zinnen met deze vraagwoorden vanuit het perspectief van prosodie en taalverwerking. Om de onderwerpen beter te kunnen verkennen, zijn er diverse methodes gebruikt in deze dissertatie, waaronder analyses van productie- en audio-opnames, een audioperceptie/gating paradigma, zelf-gereguleerd lezen (self-paced reading), en elektrofysiologische middelen. Deze diversiteit aan empirische studies leveren bewijs voor de rol die prosodie speelt in zinsbepaling in het Mandarijn, op basis waarvan een uitgebreide hypothese over zinsbepaling wordt voorgesteld. De fiatteringscondities van vraagwoorden (bijv. 'wat') en vragende voornaamwoorden in onbepaald gebruik, de zogenaamde *wh-existentials* (vraagwoorden met een existentiële interpretatie, bijv. 'iets'), de twee meest gebruikte interpretaties, worden ook onderzocht.

Deze dissertatie bestaat uit zeven hoofdstukken. **Hoofstuk 1** introduceert kort de achtergrond van de voornaamste onderzoekskwesties, geeft de onderzoeksvragen weer en een overzicht van de dissertatie.

Hoofdstuk 2 onderzoekt de prosodische markering van open vraagzinnen en vraagwoord-declaratieven (declaratieve zinnen waarin een vraagwoord als een onbepaald voornaamwoord geïnterpreteerd dient te worden. In onze studie hebben deze zinnen dezelfde woorden en woordvolgorde als de vraagzinnen) middels een productie-experiment met moedertaalsprekers van het Mandarijn uit Beijing. Een gedetailleerde analyse van de audio-opnames laat zien dat de vraagwoord-declaratieven verschillen van de open vraagzinnen wat betreft hun prosodische kenmerken, en dat verschil is al meetbaar vanaf het begin van de zin. Vraagwoord-declaratieven zijn akoestisch langer dan open vraagzinnen vanaf het onderwerp van de zin, maar dit patroon keert zich om vanaf het vraagwoord; vraagwoord-declaratieven hebben een lagere F0 en een kleiner F0-bereik dan het vraagwoord van vraagzinnen. Er is bovendien een compressie van het bereik van de F0 in de rest van de vraagzin na het vraagwoord. De werkwoorden van vraagwoord-declaratieven hebben een groter intensiteitsbereik dan de woordwoorden van vraagzinnen, maar het omgekeerde geldt voor het vraagwoord; bovendien verschillen open vraagzinnen van vraagwoord-declaratieven op het gebied van prosodische frasering en focus. De

productiestudie in dit hoofdstuk laat zien dat prosodie inderdaad wordt ingezet bij zinsbepaling.

Op basis van de bevindingen van Hoofdstuk 2 onderzoekt **Hoofdstuk 3** of luisteraars in staat zijn om de twee zinstypes (open vraagzinnen en vraagwoord-declaratieven) te identificeren alsook te anticiperen, waarbij zij alleen gebruik kunnen maken van prosodische aanwijzingen. Dit hoofdstuk bespreekt een audioperceptie-experiment, waarin proefpersonen werd gevraagd om naar een zin te luisteren (een vraagzin of een vraagwoord-declaratief) en een dialoog aan te vullen. De audioperceptiestudie bevat een audio-gating experiment, waarbij de opnames van beide typen zinnen gesegmenteerd zijn in fragmenten van verschillende lengtes. Het audio-gating experiment onderzoekt of proefpersonen zinstypes kunnen anticiperen voordat ze het vraagwoord horen, en, als dat zo is, vanaf welk fragment ze daarin slagen. De resultaten van het audioperceptie- en gating experiment laten zien dat 1) proefpersonen in staat zijn om gebruik te maken van prosodie om de twee zinnen van elkaar te onderscheiden; 2) vanaf het begin van het eerste woord van de zin laten proefpersonen al een voorkeur zien voor het zinstype dat door de spreker bedoeld was. Bovendien, hoe meer luisteraars van een zin horen, hoe accurater hun anticipatie van het zinstype wordt.

Hoofdstuk 4 bespreekt de fiattering in het Mandarijn van vraagwoorden gebruikt als onbepaalde voornaamwoorden, de *wh-existentials*. In de literatuur is altijd aangegeven dat dit woordtype alleen gefiatteerd kan worden in contexten met non-veridicale operatoren zoals negatie, vraagzinnen, en conditionele en epistemisch modale zinnen, waarin de waarheid van een propositie niet kan volgen uit de zin zelf. In dit hoofdstuk weerleg ik deze claim, door te demonstreren dat Mandarijnse *wh*-existentials ook in veridicale contexten gefiatteerd kunnen worden. Ik richt me op veridicale vraagzinnen met *diănr* 'een beetje' en bespreek de fiattering van *wh*-existentials. Ik heb een leesexperiment afgenomen waar proefpersonen veridicale vraagzinnen met en zonder *diănr* lazen en de resultaten laten zien dat vraagzinnen met *diănr* ambigu zijn tussen open vraagzinnen en vraagwoord-declaratieven. Dit empirische bewijs laat zien dat *wh*-existentials ook in veridicale omgevingen gefiatteerd kunnen worden. Om meer inzicht te krijgen in de fiatteringscondities van *diănr* op *wh*-existentials is er een speciale gating-studie ontworpen. De resultaten van deze studie laten zien dat luisteraars terughoudend zijn om een vraagwoord te interpreteren als een *wh*-existential als het werkwoord en *diănr* zich in verschillende prosodische domeinen bevinden.

Ik bespreek de fiattering van Mandarijnse vraagwoorden in **Hoofdstuk 5**. Ik heb zelf-gereguleerde leesexperimenten uitgevoerd op Mandarijnse simplexe open vraagzinnen (die simplexe vraagwoorden zoals 'wie' bevatten) en die vergeleken met hun declaratieve tegenhangers met 'iemand', en met complexe open vraagzinnen (die vraagwoorden zoals 'welke x' bevatten; deze staan ook wel bekend als discours-gekoppelde vraagwoorden) in vergelijking met hun declaratieve tegenhangers die 'een x' bevatten. De resultaten laten zien dat vragen met meer moeite worden verwerkt dan declaratieve zinnen, wat empirisch bewijs levert voor een coverte afhankelijkheid tussen het *in-situ* vraagwoord en het Spec-CP of C^0 domein. Deze studie impliceert ook dat complexe vraagzinnen anders verwerkt worden dan simplexe vraagzinnen vanwege de discourse-linking eigenschappen van de eerste.

In **Hoofdstuk 6** rapporteer ik de resultaten van twee ERP experimenten en geef ik elektrofysiologisch bewijs voor de rol van prosodie in zinsbepaling. Aangezien de audio-perceptiestudie (Hoofdstuk 3) laat zien dat luisteraars in staat zijn om open vraagzinnen en vraagwoord-declaratieven accuraat te identificeren zodra ze het vraagwoord horen, heb ik in het eerste ERP experiment audio-opnames van vraagzinnen en vraagwoord-declaratieven gecross-spliced vanaf het vraagwoord. Ik heb vervolgens gekeken naar vraagwoord-declaratieven (D), open vraagzinnen (Q), en hun cross-spliced condities D-Q en Q-D. Alle zinnen werden voorafgegaan door contexten die de luisteraar laten neigen naar een interpretatie van een vraagzin in het geval van de Q en Q-D condities, en naar de interpretatie van een declaratieve zin in het geval van de D en D-Q condities. Aangezien onze audio-gating studie (Hoofdstuk 3) laat zien dat luisteraars het zinstype kunnen anticiperen als ze het begin van de zin horen (bijvoorbeeld het onderwerp), manipuleert het tweede ERP experiment de mate van congruentie tussen contexten die de luisteraar laten neigen naar een interpretatie van vraagzinnen (Q) / vraagwoord-declaratieven (D) en de doelzinnen van vraagzinnen (Q) / vraagwoord-declaratieven (D). Elektrofysiologische resultaten van beide experimenten laten een (vroege) prosodische negativiteit zien, wat een indicatie is van een vroege realisatie van een incongruent zinstype, gebaseerd op prosodische informatie; maar de negatieve signalen zijn alleen te vinden in de incongruente conditie D-Q, wat impliceert dat de incongruenties op een andere manier worden verwerkt.

Hoofdstuk 7 komt ten slotte terug bij de onderzoeksvragen van Hoofdstuk 1 en vat de belangrijkste bevindingen van deze dissertatie samen. Aan het einde van dit hoofdstuk sluit ik de dissertatie af met een conclusie en suggesties voor toekomstig onderzoek.

摘要

本博士论文旨在探讨句法学和语义学中的两个重要议题：1）汉语疑问句和陈述句是如何标句的(clausal typing)；2）汉语 *wh*-不定代词(*wh*-indeterminates)的意义是如何在句中得到确定和允准的(licensing)。汉语的疑问词之所以能引起学者的兴趣，原因在于汉语在表示疑问时，疑问词不发生移位(*wh*-in-situ)，即汉语疑问句的语序与陈述句一致，因此汉语无法像英语一样通过疑问词的移位与否来标句。同时，汉语的疑问词是 *wh*-不定代词，其有非疑问的用法，却又无法像日语和韩语一样通过丰富的助词(particles)来允准 *wh*-不定代词的意义。与传统研究不同，该博士论文从韵律和句子加工的视角出发，采用多种实验方法，如声学分析、听觉感知/gating 范式，自控步速阅读(self-paced reading)和脑电(electrophysiological)研究等手段，从而能更全面地探讨上述两个议题。基于我们的研究证据，我们证实了韵律是汉语标句中的关键；以此为基础，我们还对标句假设理论(clausal typing hypothesis)进行了相应的讨论及扩展。对于第二个议题，本论文探讨了 *wh*-不定代词的两种最常见的意义，即疑问义(*wh*-interrogatives)和存在义(*wh*-existentials)在句子中的允准，研究证据支持了疑问句存在隐性依存关系(covert dependency)来允准疑问义的说法，而存在义的允准则不局限于 nonveridical 语境(命题真值不能确定的语境)。

本论文主要包括七章内容。

第一章介绍了研究议题的相关背景，提出了本论文要解决的研究问题，并罗列了行文结构。

第二章通过一个声学实验探讨了汉语特殊疑问句和 *wh*-declaratives (包含 *wh*-不定代词，但作存在义理解所以句子为陈述句；这里的 *wh*-declaratives 与特殊疑问句除标点和韵律以外表面形式上完全相同)的韵律区别。声学分析结果显示，特殊疑问句和 *wh*-declaratives 从句首开始就呈现出不同的韵律特征。从句子每个词的时长来看， *wh*-declaratives 都要比特殊疑问句更长；但在疑问词自身上，情况刚好相反：疑问句中的疑问词要比 *wh*-declaratives 句中作存在义的 *wh*-不定代词时长更长。从音高和音域上来看，在疑问句中的疑问词，相比在 *wh*-declaratives 句中的 *wh*-不定代词而言，音高更高，音域更宽，且疑问词在疑问句中呈现音域压缩(F0-range compression)的趋势。从音强差值上来看，疑问句中的动词比 *wh*-declaratives 句中的动词音强差值要小，而在疑问词上，该情况正好相反。除此之外，疑问句与 *wh*-declaratives 句在韵律结构和焦点的位置分布上也不相同。声学实验的分析结果表明，发音人使用不同的韵律来进行标句。

第三章进一步研究了听话者能否仅凭韵律特征就对疑问句和 *wh*-declaratives 句进行区分和预测。为此我们进行了两个听觉感知实验：第一个实验要求被试听句子(疑问句或者 *wh*-declaratives)后, 借助对话情景辨别句型(clause type)；第二个实验则基于 gating 范式进行，我们人为将疑问句和 *wh*-declaratives 句切分成不同长度音频/gates，以此作为实验材料，测试被试能否在听到疑问词之前的 gate 就预测出句子的类型，如果能够听辨, 具体是在哪个

gate。两个实验的结果分别表明: 1) 被试能够通过韵律特征来区别两种句型; 2) 从句子的第一个词开始，被试就已经能够预测说话者的目标句型了。

第四章讨论了 *wh*-不定代词的存在义(*wh*-existential)在汉语中的允准(licensing)。前贤认为，*wh*-不定代词的存在义用法只能在 nonveridical 语境(命题真值不能确定的语境)得到允准，即包含否定、疑问、条件或者认知情态词的语境。在本章中，我们对该观点进行了反驳。通过对包含“点儿”的 veridical 语境, 以及其对句中 *wh*-不定代词的存在义的允准的讨论，我们认为，汉语的 *wh*-不定代词的存在义也可以在 veridical 语境中得到允准。具体来说，本章包含一个句子阅读实验，实验材料为包含 *wh*-不定代词的 veridical 句子, 其中一部分带有“点儿”。实验结果表明，被试更倾向将不带“点儿”的句子判断为疑问句 (即 *wh*-不定代词只有疑问义)，而对带“点儿”的句子则会出现两可判断，即可以是疑问句也可以是陈述句。这些结果表明，在 veridical 句子中，*wh*-不定代词的存在义也可以得到允准。为了进一步探索“点儿”对 *wh*-不定代词的存在义的允准，在第三章 gating 实验的基础上，本章又进行了一个特殊的 gate 实验，即把动词和“点儿”切分在不同的 gate 中。实验结果表明：“点儿”对 *wh*-不定代词存在义的允准是有一定韵律限制的，当动词和“点儿”不在一个 gate 中时，听话者就不再倾向将 *wh*-不定代词理解为存在义了。

第五章讨论了汉语 *wh*-不定代词的疑问义(*wh*-interrogatives)的允准。本章通过自控步速阅读实验，讨论了两个问题：一是，对比了汉语非语篇连接疑问句(即 non-discourse-linking, 包含疑问词“谁”)和与之对应的陈述句(对应疑问句“谁”的位置为 “人”)在阅读加工上的区别；二是，语篇连接疑问句(即 discourse-linking, 包含疑问词“哪个 XX”)和与之对应的陈述句(对应疑问句“哪个 XX”的位置为 “一个 XX”)在阅读加工上的区别。实验结果表明，加工疑问句比加工对应的陈述句需要更多的加工资源。这为语法上疑问句中疑问词和 Spec-CP or C^0 上的疑问算子之间的隐形依存和允准关系提供了实证支持(陈述句没有这种依存/允准关系，所以不需要额外的加工资源)。此外，该研究还显示，语篇连接疑问句因为其连接语篇的特征，在加工模式上与非语篇连接疑问句存在一定的区别。

第六章汇报了两个 ERP 实验，为在线句子加工中韵律对标句的作用提供了电生理证据。第三章中的感知实验已经证实，听话者能够在听到疑问词时就非常准确地进行判断。据此，第一个 ERP 实验以疑问词为界限，将疑问句的录音和对应的 *wh*-declaratives 的录音交叉剪接，创造两种句型不连贯的条件(D-Q 和 Q-D)，然后测试两个未剪接和两个剪接过的的条件（共计四个条件）的脑电加工过程。第二个 ERP 实验的设计基于第三章 gating 实验的结果，该实验已表明听话者在句子开始的部分(例如主语)就能进行句子类型的预测。因此，本 ERP 实验将 *wh*-declaratives (D)和疑问句(Q)与其各自的上下文(the preceding context)进行交叉组合，形成四个条件，测试这四个条件的脑电加工过程。两个实验结果一致观测到了韵律不连贯所引发的(早期)负波((early) prosodic negativities)。这表明在句子的在线加工中，被试可以较早地觉察到韵律主导的句子标句冲突。但这种(早期)负波只在 D-Q 中发现，而未在 Q-D 中发现，即两种不连贯的句型可能有不同的加工和适应机制。

基于上述研究，第七章讨论了第一章所提出的研究问题，并总结了本论文的主要发现。最后，作者总结全文，并对未来相关研究的议题提出了建议。

Appendices

Appendix A

Summary of all the fitting linear mixed effects models on measurements in Chapter 2 where the clause type has an effect.

Measurement	**Model**
Utterance duration	Utterance <- lmer (Utterance ~ condition + (condition|subject) + (1|item), data=data)
Word duration	
subject	Word <- lmer (duration ~ condition + (1|subject) + (1|item), data=data)
verb-le	Word <- lmer (duration ~ condition + (condition|subject) + (condition|item), data=data)
diǎnr	Word <- lmer (duration ~ condition + (condition|subject) + (1|item), data=data)
shénme	Word <- lmer (duration ~ condition + (condition|subject) + (1|item), data=data)
F0	
F0-min *shén*	F0 <- lmer (F0 ~ condition + (condition|subject) + (1|item), data=data)
F0-min *me*	F0 <- lmer (F0 ~ condition + (condition|subject) + (1|item), data=data)
F0-max *me*	F0 <- lmer (F0 ~ condition + (condition|subject) + (1|item), data=data)
F0-min verb (T2)	F0 <- lmer (F0 ~ condition + (1|subject) + (condition|item), data=data)
F0-max verb (T2)	F0 <- lmer (F0 ~ condition + (1|subject) + (1|item), data=data)
F0-min *gěi*	F0 <- lmer (F0 ~ condition + (condition|subject) + (1|item), data=data)
F0-min indirect object(1st syllable)	F0 <- lmer (F0 ~ condition + (condition|subject) + (1|item), data=data)
F0-max indirect object (1st syllable)	F0 <- lmer (F0 ~ condition + (condition|subject) + (1|item), data=data)
F0-max indirect object (2nd syllable)	F0 <- lmer (F0 ~ condition + (condition|subject) + (1|item), data=data)
F0 range	
me	F0 range <- lmer (F0Range ~ condition + (condition|subject) + (1|item), data=data)
preposition phrase	F0 range <- lmer (F0Range ~ condition + (condition|subject) + (1|item), data=data)
Intensity range	
verb	Intensity range <- lmer (IntensityRange ~ condition + (1|subject) + (1|item), data=data)
le	Intensity range <- lmer (IntensityRange ~ condition + (1|subject) + (condition|item), data=data)
diǎnr	Intensity range <- lmer (IntensityRange ~ condition + (1|subject) + (1|item), data=data)
shén	Intensity range <- lmer (IntensityRange ~ condition + (condition|subject) + (1|item), data=data)
me	Intensity range <- lmer (IntensityRange ~ condition + (condition|subject) + (1|item), data=data)
gěi	Intensity range <- lmer (IntensityRange ~ condition + (condition|subject) + (1|item), data=data)

Appendix B

Summary of all the fitting linear mixed effects models in each gate in Chapter 3.

Gate	**Model**
a	ParticipantResponse <- glmer (response~ condition + (1 + condition | item) + (1 | subject), data=data,family=binomial())
b	ParticipantResponse <- glmer (response~ condition + (1 + condition | item) + (1 | subject), data=data,family=binomial())
d	ParticipantResponse <- glmer (response~ condition + (1 + condition | item) + (1 | subject), data=data,family=binomial())

Appendix C

#1a 那个勇敢的男生救了谁就离开了？
救人的是那个勇敢的女生吗？ N
#1b 那个勇敢的男生救了人就离开了。
救人的是那个勇敢的女生吗？ N
#2a 那个傲慢的上司骂了谁就回家了？
骂人的是那个傲慢的上司吗？ Y
#2b 那个傲慢的上司骂了人就回家了。
骂人的是那个傲慢的上司吗？ Y
#3a 那个优雅的绅士帮了谁就离去了？
那个绅士帮了人就留下了吗？ N
#3b 那个优雅的绅士帮了人就离去了。
那个绅士帮了人就留下了吗？ N
#4a 那个严厉的老师训了谁就回去了？
那个老师回去了吗？ Y
#4b 那个严厉的老师训了人就回去了。
那个老师回去了吗？ Y
#5a 那个严肃的交警罚了谁就走掉了？
是交警罚了人就走掉了吗？ Y
#5b 那个严肃的交警罚了人就走掉了。
是交警罚了人就走掉了吗？ Y
#6a 那个机智的警察抓了谁就撤走了？
是那个没经验的消防员抓了人就撤了吗？ N
#6b 那个机智的警察抓了人就撤走了。
是那个没经验的消防员抓了人就撤了吗？ N
#7a 那个可恶的郎中骗了谁就开溜了？
那个郎中开溜了吗？ Y
#7b 那个可恶的郎中骗了人就开溜了。
那个郎中开溜了吗？ Y
#8a 那个凶狠的罪犯杀了谁就逃走了？
那个罪犯自首了吗？ N
#8b 那个凶狠的罪犯杀了人就逃走了。
那个罪犯自首了吗？ N
#9a 那个超速的司机撞了谁就逃跑了？
司机是正常速度驾驶吗？ N
#9b 那个超速的司机撞了人就逃跑了。
司机是正常速度驾驶吗？ N
#10a 那个顽劣的痞子打了谁就溜走了？
痞子打了人就自首了吗？ N
#10b 那个顽劣的痞子打了人就溜走了。

痞子打了人就自首了吗？ N
#11a 那个好心的女士扶了谁就离开了？
去扶人的是个好心的女士吗？ Y
#11b 那个好心的女士扶了人就离开了。
去扶人的是个好心的女士吗？ Y
#12a 那个草率的领导裁了谁就后悔了？
那个草率的领导后悔了吗？ Y
#12b 那个草率的领导裁了人就后悔了。
那个草率的领导后悔了吗？ Y
#13a 那个热心的学长帮了谁就放心了？
帮人的是那个热心的学姐吗？ N
#13b 那个热心的学长帮了人就放心了。
帮人的是那个热心的学姐吗？ N
#14a 那个无私的医生救了谁就累倒了？
救了人就累倒的是医生吗？ Y
#14b 那个无私的医生救了人就累倒了。
救了人就累倒的是医生吗？ Y
#15a 那个年长的学者见了谁就高兴了？
那个学者见了人就生气了吗？ N
#15b 那个年长的学者见了人就高兴了。
那个学者见了人就生气了吗？ N
#16a 那个冲动的少年杀了谁就自首了？
那个少年杀了人就潜逃了吗？ N
#16b 那个冲动的少年杀了人就自首了。
那个少年杀了人就潜逃了吗？ N
#17a 那个鲁莽的员工惹了谁就辞职了？
辞职的是那个鲁莽的员工吗？ Y
#17b 那个鲁莽的员工惹了人就辞职了。
辞职的是那个鲁莽的员工吗？ Y
##18a 那个年轻的村民害了谁就自杀了？
自杀的是那个年长的村支书吗？ N
#18b 那个年轻的村民害了人就自杀了。
自杀的是那个年长的村支书吗？ N
#19a 那个冲动的被告打了谁就道歉了？
那个冲动的被告道歉了吗？ Y
#19b 那个冲动的被告打了人就道歉了。
那个冲动的被告道歉了吗？ Y
#20a 那个和蔼的导师夸了谁就回去了？
那个和蔼的导师回去了吗？ Y
#20b 那个和蔼的导师夸了人就回去了。
那个和蔼的导师回去了吗？ Y

#21a 那个年轻的主妇雇了谁就轻松了？
那个主妇年龄很大吗？ N
#21b 那个年轻的主妇雇了人就轻松了。
那个主妇年龄很大吗？ N
#22a 那个挑剔的姑娘嫁了谁就后悔了？
那个姑娘挑剔吗？ Y
#22b 那个挑剔的姑娘嫁了人就后悔了。
那个姑娘挑剔吗？ Y
#23a 那个单纯的孩子见了谁就失踪了？
那个孩子接着就回家了吗？ N
#23b 那个单纯的孩子见了人就失踪了。
那个孩子接着就回家了吗？ N
#24a 那个新任的班长骂了谁就道歉了？
那个新任班长道歉了吗？ Y
#24b 那个新任的班长骂了人就道歉了？
那个新任班长道歉了吗？ Y

Appendix D

#1a 那个男生想要求谁解决问题？
是男生想要求人解决问题吗？ Y
#1b 那个男生想要求人解决问题。
是男生想要求人解决问题吗？ Y
#2a 那个上司准备选谁接替职位？
这个职位是要被接替吗？ Y
#2b 那个上司准备选人接替职位。
这个职位是要被接替吗？ Y
#3a 那个商贩想要骗谁购买假货？
这是一批真货吗？ N
#3b 那个商贩想要骗人购买假货。
这是一批真货吗？ N
#4a 那个同事打算找谁修理电脑？
要修的是电脑吗？ Y
#4b 那个同事打算找人修理电脑。
要修的是电脑吗？ Y
#5a 那个老板准备招谁开展业务？
是员工要招人开展业务吗？ N
#5b 那个老板准备招人开展业务。
是员工要招人开展业务吗？ N
#6a 那个警察想要逼谁交出证据？
是嫌犯想要逼人交出证据吗？ N

#6b 那个警察想要逼人交出证据。
是嫌犯想要逼人交出证据吗？N
#7a 那个画家打算请谁装修画室？
要装修的是卧室吗？N
#7b 那个画家打算请人装修画室。
要装修的是卧室吗？N
#8a 那个员工准备接谁出席会议？
是员工要去接人吗？Y
#8b 那个员工准备接人出席会议。
是员工要去接人吗？Y
#9a 那个店长准备聘谁打理事物？
是店长要聘人吗？Y
#9b 那个店长准备聘人打理事物。
是店长要聘人吗？Y
#10a 那个管家打算叫谁清理房间？
要清理的是庭院吗？N
#10b 那个管家打算叫人清理房间。
要清理的是庭院吗？N
#11a 那个女士打算托谁代购补品？
要托人代购的是补品吗？Y
#11b 那个女士打算托人代购补品。
要托人代购的是补品吗？Y
#12a 那个大妈想要讹谁赔偿费用？
是大叔想要讹别人吗？N
#12b 那个大妈想要讹人赔偿费用。
是大叔想要讹别人吗？N
#13a 那个学长准备陪谁面试工作？
是学姐准备陪人面试吗？N
#13b 那个学长准备陪人面试工作。
是学姐准备陪人面试吗？N
#14a 那个 老人想要劝谁珍惜感情？
老人想要劝人放弃感情吗？N
#14b 那个 老人想要劝人珍惜感情。
老人想要劝人放弃感情吗？N
#15a 那个 老师准备教谁学习语言？
是要学习语言吗？Y
#15b 那个 老师准备教人学习语言。
是要学习语言吗？Y
#16a 那个客人想要让谁弹奏曲子？
客人是想让人修改曲子吗？N
#16b 那个客人想要让人弹奏曲子。

客人是想让人修改曲子吗？N
#17a 那个对手打算派谁散布谣言？
是对手打算派人造谣吗？Y
#17b 那个对手打算派人散布谣言。
是对手打算派人造谣吗？Y
#18a 那个警官准备带谁执行任务？
是要去执行任务吗？Y
#18b 那个警官准备带人执行任务。
是要去执行任务吗？Y
#19a 那个导师打算挑谁接手项目？
是院长打算挑人吗？N
#19b 那个导师打算挑人接手项目。
是院长打算挑人吗？N
#20a 那个邻居准备雇谁打扫院子？
是邻居准备雇人吗？Y
#20b 那个邻居准备雇人打扫院子。
是邻居准备雇人吗？Y
#21a 那个房东打算留谁享用晚餐？
是要享用午饭吗？N
#21b 那个房东打算留人享用晚餐。
是要享用午饭吗？N
#22a 那个主编想要催谁完成专栏？
是社长催人完成专栏吗？N
#22b 那个主编想要催人完成专栏。
是社长催人完成专栏吗？N
#23a 那个厨师打算喊谁品尝佳肴？
要品尝的是佳肴吗？Y
#23b 那个厨师打算喊人品尝佳肴。
要品尝的是佳肴吗？Y
#24a 那个教练想要罚谁练习发球？
是教练想要罚人吗？Y
#24b 那个教练想要罚人练习发球。
是教练想要罚人吗？Y

Appendix E

#1a 那个男生想要求哪个同学解决问题？
是男生想要求人解决问题吗？Y
#1b 那个男生想要求一个同学解决问题。
是男生想要求人解决问题吗？Y
#2a 那个上司准备选哪个员工接替职位？

这个职位是要被接替吗？Y
#2b 那个上司准备选一个员工接替职位。
这个职位是要被接替吗？Y
#3a 那个商贩想要骗哪个顾客购买假货？
这是一批真货吗？N
#3b 那个商贩想要骗一个顾客购买假货。
这是一批真货吗？N
#4a 那个同事打算找哪个师傅修理电脑？
要修的是电脑吗？Y
#4b 那个同事打算找一个师傅修理电脑。
要修的是电脑吗？Y
#5a 那个老板准备招哪个职员开展业务？
是员工要招人开展业务吗？N
#5b 那个老板准备招一个职员开展业务。
是员工要招人开展业务吗？N
#6a 那个警察想要逼哪个证人交出证据？
是嫌犯想要逼人交出证据吗？N
#6b 那个警察想要逼一个证人交出证据。
是嫌犯想要逼人交出证据吗？N
#7a 那个画家打算请哪个工人装修画室？
要装修的是卧室吗？N
#7b 那个画家打算请一个工人装修画室。
要装修的是卧室吗？N
#8a 那个员工准备接哪个领导出席会议？
是员工要去接人吗？Y
#8b 那个员工准备接一个领导出席会议。
是员工要去接人吗？Y
#9a 那个店长准备聘哪个助手打理事物？
是店长要聘人吗？Y
#9b 那个店长准备聘一个助手打理事物。
是店长要聘人吗？Y
#10a 那个管家打算叫哪个佣人清理房间？
要清理的是庭院吗？N
#10b 那个管家打算叫一个佣人清理房间。
要清理的是庭院吗？N
#11a 那个女士打算托哪个熟人代购补品？
托人代购的是补品吗？Y
#11b 那个女士打算托一个熟人代购补品。
托人代购的是补品吗？Y
#12a 那个大妈想要讹哪个路人赔偿费用？
是大叔想要讹别人吗？N

#12b 那个大妈想要讹一个路人赔偿费用。
是大叔想要讹别人吗？N
#13a 那个学长准备陪哪个学弟面试工作？
是学姐准备陪人面试吗？N
#13b 那个学长准备陪一个学弟面试工作。
是学姐准备陪人面试吗？N
#14a 那个 老人想要劝哪个晚辈珍惜感情？
老人想要劝人放弃感情吗？N
#14b 那个 老人想要劝一个晚辈珍惜感情。
老人想要劝人放弃感情吗？N
#15a 那个 老师准备教哪个学生学习语言？
是要学习语言吗？Y
#15b 那个 老师准备教一个学生学习语言。
是要学习语言吗？Y
#16a 那个客人想要让哪个女士弹奏曲子？
客人是想让人修改曲子吗？N
#16b 那个客人想要让一个女士弹奏曲子。
客人是想让男士弹奏曲子吗？N
#17a 那个对手打算派哪个随从散布谣言？
是对手打算派人造谣吗？Y
#17b 那个对手打算派一个随从散布谣言。
是对手打算派人造谣吗？Y
#18a 那个警官准备带哪个警员执行任务？
是要去执行任务吗？Y
#18b 那个警官准备带一个警员执行任务。
警官准备带的是警员吗？Y
#19a 那个导师打算挑哪个博士接手项目？
是院长打算挑人吗？N
#19b 那个导师打算挑一个博士接手项目。
是硕士要接手项目吗？N
#20a 那个邻居准备雇哪个阿姨打扫院子？
是邻居准备雇人吗？Y
#20b 那个邻居准备雇一个阿姨打扫院子。
要雇的是一个阿姨吗？Y
#21a 那个房东打算留哪个朋友享用晚餐？
是要享用午饭吗？N
#21b 那个房东打算留一个朋友享用晚餐。
是要留一个陌生人吃饭吗？N
#22a 那个主编想要催哪个作者完成专栏？
是社长催人完成专栏吗？N
#22b 那个主编想要催一个作者完成专栏。

被催的是一个编辑吗？N
#23a 那个厨师打算喊哪个同事品尝佳肴？
要品尝的是佳肴吗？Y
#23b 那个厨师打算喊一个同事品尝佳肴。
是要喊同事品尝佳肴吗？Y
#24a 那个教练想要罚哪个学员练习发球？
是教练想要罚人吗？Y
#24b 那个教练想要罚一个学员练习发球。
要被罚的是学员吗？Y

Appendix F

1a wh-declarative with its preceding context biasing it
罗涛家里有很多玩具。前天刘东去罗涛家玩，走的时候包里鼓鼓囊囊的而且特别开心。虽然不知道具体情况，但我们都相信：罗涛前天分了点儿什么给刘东。
1b wh-question with its preceding context biasing it
罗涛有很多玩具而且他乐于与人分享。前天刘东去罗涛家玩，走的时候包里鼓鼓囊囊的而且很开心。我们都好奇地问：罗涛前天分了点儿什么给刘东？
2a wh-declarative with its preceding context biasing it
白薇和罗英从小就是好朋友。每个星期，她们都会去农场的果园采摘。可是，罗英前天生病了，白薇只能自己去。我们都相信：白薇前天摘了点儿什么给罗英。
2b wh-question with its preceding context biasing it
白薇和罗英从小就是好朋友。前天，罗英想去农场的果园采摘，但她生病去不了，好朋友白薇代她去了果园。我们很想问：白薇前天摘了点儿什么给罗英？
3a wh-declarative with its preceding context biasing it
刘波家境不太好，经常做些小偷小摸的事情。昨天，刘波对女友白冰说，有一个礼物要送给她。所以我们几乎可以断定：刘波昨天偷了点儿什么给白冰。
3b wh-question with its preceding context biasing it
刘波家境不太好，经常做些小偷小摸的事情。昨天，刘波为了给女友白冰庆生，又去以身试法了。我们都很好奇地问：刘波昨天偷了点儿什么给白冰？
4a wh-declarative with its preceding context biasing it
昨天，于聪在刘佳的家里做客。于聪肚子饿了，刘佳从冰箱里拿出吃的，然后去厨房找刀子。我们不用问也知道：刘佳昨天切了点儿什么给于聪。
4b wh-question with its preceding context biasing it
昨天，于聪在刘佳的家里做客。于聪肚子饿了，刘佳从冰箱里拿出吃的，然后去厨房找刀子。我们忍不住问：刘佳昨天切了点儿什么给于聪？
5a wh-declarative with its preceding context biasing it
黄威会用柳条编各种东西。前天，黄威在家里忙活了一天，好像是为了给他的好友吴康一个惊喜。 我们不用问也知道：黄威前天编了点儿什么给吴康。

5b wh-question with its preceding context biasing it
黄威会用柳条编各种东西。前天，黄威在家里忙活了一天，好像是为了给他的好友吴康一个惊喜。 我们都好奇地询问：黄威前天编了点儿什么给吴康？
6a wh-declarative with its preceding context biasing it
陈晖捕捞海鲜的技术很高超。因为好友秦东帮过陈晖一个忙，为了感谢秦东，陈晖昨天带着鱼网出海一天。虽不知细节，我们相信：陈晖前天捞了点儿什么给秦东。
6b wh-question with its preceding context biasing it
陈晖捕捞海鲜的技术很高超。因为好友秦东帮过陈晖一个忙，为了感谢秦东，陈晖昨天带着鱼网出海一天。我们都好奇地问：陈晖前天捞了点儿什么给秦东？
7a wh-declarative with its preceding context biasing it
池塘里有很多鱼虾等水产品，很多人下水去捞。陶薇不敢下水，特地让水性好的刘刚帮忙。看到陶薇昨天很开心，我们相信：刘刚昨天捞了点儿什么给陶薇。
7b wh-question with its preceding context biasing it
池塘里有很多鱼虾等水产品，很多人下水去捞。陶薇不敢下水，特地让水性好的刘刚帮忙。看到陶薇昨天很开心，我们忍不住问：刘刚昨天捞了点儿什么给陶薇？
8a wh-declarative with its preceding context biasing it
王超不管吃什么东西，都会掰一半儿分给罗东。昨天，大家打扫班级卫生的时候发现他俩座位下有零食屑，我们都相信：王超昨天掰了点儿什么给罗东。
8b wh-question with its preceding context biasing it
王超不管吃什么东西，都会掰一半儿分给罗东。昨天，大家打扫班级卫生的时候发现他俩座位下有零食屑，我们很好奇地问：王超昨天掰了点儿什么给罗东？
9a wh-declarative with its preceding context biasing it
杨君送给陈芳一个针织品，针脚织得特别细致，据说杨君前天花了一天才搞定。虽然不知道具体是什么，我们可以确定的是：杨君前天织了点儿什么给陈芳。
9b wh-question with its preceding context biasing it
杨君送给陈芳一个针织品，针脚织得特别细致，据说杨君前天花了一天才搞定。我们纷纷好奇地问：杨君前天织了点儿什么给陈芳？
10a wh-declarative with its preceding context biasing it
何超不遵守纪律，昨天幼儿园老师罚他不许吃饭。好朋友徐军看不下去了，正好徐军自己的饭量小，于是，趁老师不在：徐军昨天分了点儿什么给何超。
10b wh-question with its preceding context biasing it
何超不遵守纪律，昨天幼儿园老师罚他不许吃饭。好朋友徐军不忍心他挨饿，于是把自己的饭菜分给他。我们好奇地问：徐军昨天分了点儿什么给何超？
11a wh-declarative with its preceding context biasing it
前天毛嘉带林星去山上玩，看到果树上结满了果实。林星爱吃新鲜水果，毛嘉为了哄她开心就爬到树上去了。我们不用问也知道：毛嘉前天摘了点儿什么给林星。
11b wh-question with its preceding context biasing it
前天毛嘉带林星去山上玩，看到果树上结满了果实。林星爱吃新鲜水果，毛嘉为了哄她开心就爬到树上去了。我们都好奇地问：毛嘉前天摘了点儿什么给林星？

12a wh-declarative with its preceding context biasing it
彭丹因为盗窃被警察带走了。据说昨天在商场，胡忠看上了很贵的东西但没钱买，彭丹就帮他弄到了手。虽不知细节，我们都相信：彭丹昨天偷了点儿什么给胡忠。
12b wh-question with its preceding context biasing it
彭丹因为盗窃被警察带走了。据说昨天在商场，胡忠看上了很贵的东西但又没钱买，彭丹就帮他弄到了手。我们八卦地询问：彭丹昨天偷了点儿什么给胡忠？
13a wh-declarative with its preceding context biasing it
陈希擅长编织，几根绳子在手上就能变化出很多花样。于光昨天去拜访陈希，还拿到了陈希给他现场做的礼物。我们都相信：陈希昨天编了点儿什么给于光。
13b wh-question with its preceding context biasing it
陈希擅长编织。几根绳子在手上就能变化出很多花样。于光昨天去拜访陈希，还拿到了陈希给他现场做的礼物。我们纷纷揣测：陈希昨天编了点儿什么给于光？
14a wh-declarative with its preceding context biasing it
曹辉和程娟家是邻居。程娟这学期住校，生活用品缺这少那，经常要曹辉帮忙。看到前天程娟东西突然变多了，不用问我们也知道：曹辉前天捎了点儿什么给程娟。
14b wh-question with its preceding context biasing it
曹辉和程娟家是邻居。程娟这学期住校，生活用品缺这少那，经常要曹辉帮忙。看到前天程娟东西突然变多了，我们羡慕地问：曹辉前天捎了点儿什么给程娟？
15a wh-declarative with its preceding context biasing it
梁峰来自贫困山区，家里连个像样的电器都没有。于宵知道了以后，前天专程带着物品去探望梁峰。虽不知具体细节，我们都相信：于宵前天捐了点儿什么给梁峰。
15b wh-question with its preceding context biasing it
梁峰来自贫困山区，家里连个像样的电器都没有。于宵知道了以后，前天专程带着物品去探望梁峰。我们纷纷询问：于宵前天捐了点儿什么给梁峰？
16a wh-declarative with its preceding context biasing it
王芬是个心灵手巧的姑娘，尤其擅长针织，听说，为了在陈霜生日的时候给陈霜一个惊喜，王芬昨天忙了一晚上。我们都确信：王芬昨天织了点儿什么给陈霜。
16b wh-question with its preceding context biasing it
王芬是个心灵手巧的姑娘，尤其擅长针织。听说，为了在陈霜生日的时候给陈霜一个惊喜，王芬昨天忙了一晚上。我们忍不住问：王芬昨天织了点儿什么给陈霜？
17a wh-declarative with its preceding context biasing it
昨天何英去林宣家做客，桌上有各种各样的水果。虽然有水果刀，但是何英不会削。林宣是个细心的人，所以我们不用问也知道：林宣昨天削了点儿什么给何英。
17b wh-question with its preceding context biasing it
昨天何英去林宣家做客，桌上有各种各样的水果。虽然有水果刀，但是何英不会削，细心的林宣就过来帮忙。我们好奇地问：林宣昨天削了点儿什么给何英？

18a wh-declarative with its preceding context biasing it
新来的同学曹刚，宿舍里连基本生活设施都没有。前天，学生会号召大家踊跃捐赠，胡昆第一个响应号召。不用问我们也知道：胡昆前天捐了点儿什么给曹刚。
18b wh-question with its preceding context biasing it
新来的同学曹刚，宿舍里连基本生活设施都没有。前天，学生会号召大家踊跃捐赠，胡昆第一个响应号召。我们都很好奇地问：胡昆前天捐了点儿什么给曹刚？
19a wh-declarative with its preceding context biasing it
罗新前天去参加了陶冰的毕业典礼。作为陶冰最好的好朋友，罗新一定不会空着手去。我们可以肯定的是：罗新前天送了点儿什么给陶冰。
19b wh-question with its preceding context biasing it
罗新前天去参加了陶冰的毕业典礼。作为陶冰最好的好朋友，罗新一定不会空着手去。我们好奇地问：罗新前天送了点儿什么给陶冰？
20a wh-declarative with its preceding context biasing it
昨天是开学第一天，吴英忘记带文具盒了。吴英的同学何超知道后帮助了她。虽不知道具体细节，但我们知道：何超昨天借了点儿什么给吴英。
20b wh-question with its preceding context biasing it
昨天是开学第一天，吴英忘记带文具盒了。吴英的同学何超知道后帮助了她。我们好奇地询问：何超昨天借了点儿什么给吴英？
21a wh-declarative with its preceding context biasing it
王新因为思念异地工作的好友徐芳，总想给她一些惊喜。前天，王新带了一包东西去了邮局，虽不知道细节，但我们知道：王新前天寄了点儿什么给徐芳。
21b wh-question with its preceding context biasing it
王新因为思念异地工作的好友徐芳，总想给她一些惊喜。前天，王新带了一包东西到邮局去寄。我们好奇地问：王新前天寄了点儿什么给徐芳？
22a wh-declarative with its preceding context biasing it
吴菲开了一家杂货铺，因为她跟刘康有过节，起初拒绝做刘康的生意。但是刘康昨天还是从吴菲店里满载而归。不用问也知道：吴菲昨天卖了点儿什么给刘康。
22b wh-question with its preceding context biasing it
吴菲开了一家杂货铺，因为她跟刘康有过节，起初拒绝做刘康的生意。但是刘康昨天还是从吴菲店里满载而归。我们忍不住问：吴菲昨天卖了点儿什么给刘康？
23a wh-declarative with its preceding context biasing it
刘青喜欢自己做衣服。王锋问刘青可不可以也帮他做，刘青虽然红着脸没回答，但前天自己在家忙活了一天。所以我们都知道：刘青前天做了点儿什么给王锋。
23b wh-question with its preceding context biasing it
刘青喜欢自己做衣服。王锋问刘青可不可以也帮他做，刘青虽然红着脸没回答，但前天自己在家忙活了一天。我们好奇地问：　刘青前天做了点儿什么给王锋？

24a wh-declarative with its preceding context biasing it
陈刚不太会做饭，但同学王波却非常擅长。昨天王波做了很多好吃的，装在饭盒里去了一趟陈刚家。不用问也知道：王波昨天带了点儿什么给陈刚。
24b wh-question with its preceding context biasing it
陈刚不太会做饭，但同学王波却非常擅长。昨天王波做了很多好吃的，装在饭盒里去了一趟陈刚家。我们忍不住问：王波昨天带了点儿什么给陈刚？
25a wh-declarative with its preceding context biasing it
吴娇是这一带有名的画家，擅长现场创作。刘聪前天一早特地慕名而去，等到下午终于如愿以偿地离开了。我们不用问也知道：吴娇前天画了点儿什么给刘聪。
25b wh-question with its preceding context biasing it
吴娇是这一带有名的画家，擅长现场创作。刘聪前天一早特地慕名而去，等到下午终于如愿以偿地离开了。我们都很好奇地询问：吴娇前天画了点儿什么给刘聪？
26a wh-declarative with its preceding context biasing it
徐涛学会了几样东北炖菜，大锅做的那种。昨天，徐涛邀请杨晶来家里吃饭，恰好杨晶又是东北人。不用问我们也知道：徐涛昨天炖了点儿什么给杨晶。
26b wh-question with its preceding context biasing it
徐涛学会了几样东北炖菜，大锅做的那种。昨天，徐涛邀请杨晶来家里吃饭，恰好杨晶又是东北人。我们都很好奇地问：徐涛昨天炖了点儿什么给杨晶？
27a wh-declarative with its preceding context biasing it
何娟升任单位领导以后，巴结她的人就不胜枚举，胡星就是其中一个。前天，胡星又大包小包地去了何娟家。我们不用猜也知道：胡星前天送了点儿什么给何娟。
27b wh-question with its preceding context biasing it
何娟升任单位领导以后，巴结她的人就不胜枚举，胡星就是其中一个。前天，胡星又大包小包地去了何娟家。我们八卦地问： 胡星前天送了点儿什么给何娟？
28a wh-declarative with its preceding context biasing it
姚辉家里工具齐全，以至于邻居有什么需要的都不去自己买了。前天，邻居徐天往姚辉家跑了好几趟，我们不用问也知道：姚辉前天借了点儿什么给徐天。
28b wh-question with its preceding context biasing it
姚辉家里工具齐全，以至于邻居有什么需要的都不去自己买了。前天，徐天往姚辉家跑了好几趟，我们都好奇地问：姚辉前天借了点儿什么给徐天？
29a wh-declarative with its preceding context biasing it
林珊和男友袁刚是异地恋。林珊经常准备一些礼物给袁刚，正好物流也方便。这不，昨天林珊又叫了快递，我们不用问都知道：林珊昨天寄了点儿什么给袁刚。
29b wh-question with its preceding context biasing it
林珊和男友袁刚是异地恋。林珊经常准备一些礼物给袁刚，正好物流也方便。这不，昨天林珊又叫了快递，我们都八卦地问：林珊昨天寄了点儿什么给袁刚？

30a wh-declarative with its preceding context biasing it
王丹和田英最近几天总是神神秘秘的，前天田英还带了一些现金给王丹。据我们初步了解，事情是这样：王丹前天卖了点儿什么给田英。
30b wh-question with its preceding context biasing it
王丹前天好像以优惠价格出售了什么宝贝给田英，两个人都神神秘秘的，不愿跟大家明说。我们都忍不住问：王丹前天卖了点儿什么给田英？
31a wh-declarative with its preceding context biasing it
彭姗是个美食爱好者，却不喜欢做饭。为了讨好彭姗，男友杨歌昨天主动提出要准备晚饭。虽然不知味道如何，但我们知道：杨歌昨天做了点儿什么给彭姗。
31b wh-question with its preceding context biasing it
彭姗是个美食爱好者，却不喜欢做饭。为了讨好彭姗，男友杨歌昨天主动提出要准备晚饭。我们都很好奇地问：杨歌昨天做了点儿什么给彭姗？
32a wh-declarative with its preceding context biasing it
杨珍前天大老远地赶过来看望生病中的王康。杨珍走后，王康屋里东西变多了，心情也变好了。我们不用问也知道：杨珍前天带了点儿什么给王康。
32b wh-question with its preceding context biasing it
杨珍前天大老远地赶过来看望生病中的王康。杨珍走后，王康屋里东西变多了，心情也变好了。我们忍不住问：杨珍前天带了点儿什么给王康？
33a wh-declarative with its preceding context biasing it
冯嫣性格内向，但是很擅长用画画来表达自己。昨天她和好友秦娟吵架了，晚上她托人给秦娟带去一幅东西，我们不用猜也知道：冯嫣昨天画了点儿什么给秦娟。
33b wh-question with its preceding context biasing it
冯嫣性格内向，但是很擅长用画画来表达自己。昨天她和好友秦娟吵架了，晚上她托人给秦娟带去一幅东西，我们忍不住问： 冯嫣昨天画了点儿什么给秦娟？
34a wh-declarative with its preceding context biasing it
梅涛前天亲自下厨，专门为同事王征做了合乎口味的菜肴。王征是东北人，不爱吃炒菜，只爱吃炖菜。 所以不用说我们也都知道：梅涛前天炖了点儿什么给王征。
34b wh-question with its preceding context biasing it
梅涛前天亲自下厨，专门为同事王征做了合乎口味的菜肴。王征是东北人，不爱吃炒菜，只爱吃炖菜。 我们都忍不住问：梅涛前天炖了点儿什么给王征？
35a wh-declarative with its preceding context biasing it
于帆被派到日本工作，同事陈晶是日本迷，央求于帆帮他买东西。好在国际物流发达，昨天于帆带着包裹去了邮局。不用猜也知道：于帆昨天寄了点儿什么给陈晶。
35b wh-question with its preceding context biasing it
于帆被派到日本工作，同事陈晶是日本迷，央求于帆帮他买东西。好在国际物流发达，昨天于帆带着包裹去了邮局。我们好奇地问：于帆昨天寄了点儿什么给陈晶？

36a wh-declarative with its preceding context biasing it
韩刚和黄佳分手后依然是朋友。韩刚前天逛街，看到黄佳喜欢的东西还是忍不住买了下来，然后约了黄佳见面。我们不用问也知道：韩刚前天送了点儿什么给黄佳。
36b wh-question with its preceding context biasing it
韩刚和黄佳分手后依然是朋友。韩刚前天逛街，看到黄佳喜欢的东西还是忍不住买了下来，然后约了黄佳见面。我们都很八卦地问：韩刚前天送了点儿什么给黄佳？

Curriculum Vitae

Yang Yang was born in 1987 in Shandong province, the People's Republic of China. In 2009 she obtained her Bachelor degree in English Language and Literature from Shanghai International Studies University. In 2010 she started a Master's program at Shanghai International Studies University. In 2012, she spent 10 months as a visiting student in the Psychology Department of Peking University. In 2013 she obtained a Master's degree in General Linguistics and Applied Linguistics from Shanghai International Studies University. She started working at Leiden University Center for Linguistics as a PhD researcher in 2013, supported by China Scholarship Council. This dissertation is the result of her research.