



Universiteit  
Leiden  
The Netherlands

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

Yang, Y.

### **Citation**

Yang, Y. (2018, May 30). *The two sides of Wh-indeterminates in Mandarin : a prosodic and processing account*. LOT, Netherlands Graduate School of Linguistics, Utrecht. Retrieved from <https://hdl.handle.net/1887/62454>

Version: Not Applicable (or Unknown)

License: [Licence agreement concerning inclusion of doctoral thesis in the Institutional Repository of the University of Leiden](#)

Downloaded from: <https://hdl.handle.net/1887/62454>

**Note:** To cite this publication please use the final published version (if applicable).

Cover Page



Universiteit Leiden



The handle <http://hdl.handle.net/1887/62454> holds various files of this Leiden University dissertation

**Author:** Yang, Yang

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

**Date:** 2018-05-30

## 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.<sup>9</sup> Hence the clause types cannot be marked as in a *wh*-

---

<sup>9</sup> 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 HL%, while in declaratives the NPA can vary among a H\*, H\*+L or L+H\* and the BT is LL%. 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<sup>10</sup>; when syntactically unmarked, they are also string identical with their declarative counterparts and the clause types are also prosodically encoded. Prosodic markings in Mandarin<sup>11</sup> 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).

---

<sup>10</sup> 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.

<sup>11</sup> 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).<sup>12</sup> 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.

<sup>12</sup> There are more experimental conditions in Hu (2002). Here we only list the ones that are related to the current study.

14 The Two Sides of *Wh*-indeterminates: A Prosodic and Processing Account

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 constructing 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 Lǐu 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 Lǐu 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 Lǐu 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  $\times$  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)  $ST = 12\log_2(\text{Hz}/100)$ , while for male speakers we used formula (ii)  $ST = 12\log_2(\text{Hz}/50)$ .

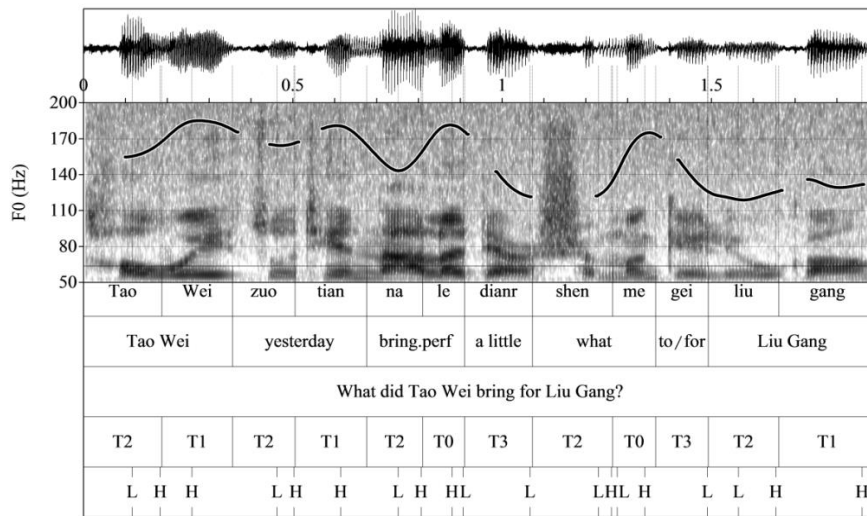


Figure 1. An exemplar waveform and spectrogram from a male participant with superimposed F0-contours, syllables, glosses and F0 measurements obtained based on the specific tones.

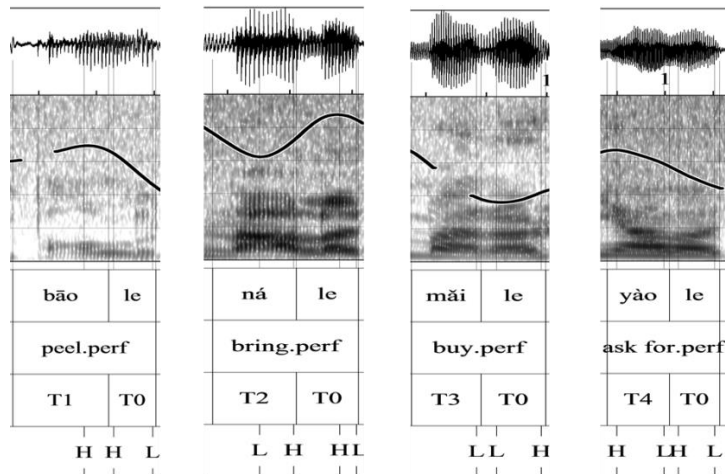


Figure 2. The F0 measurement of le obtained based on the preceding T1, T2, T3 and T4 respectively.

**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 *Liu 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.

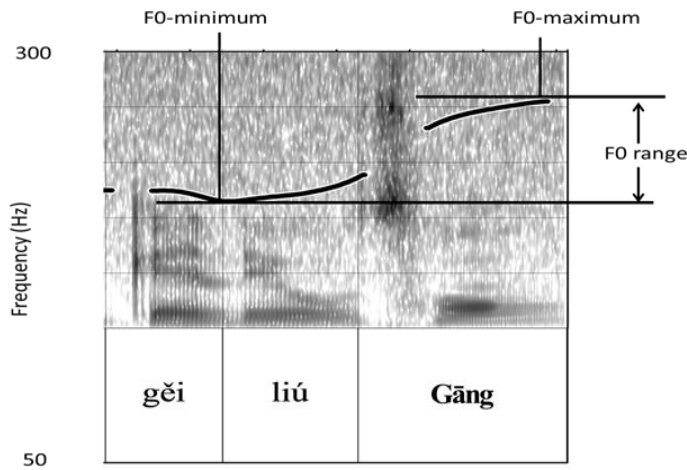


Figure 3. The F0 contours of the preposition phrase *gěi* indirect object and its F0-range obtained.

**Intensity range**

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

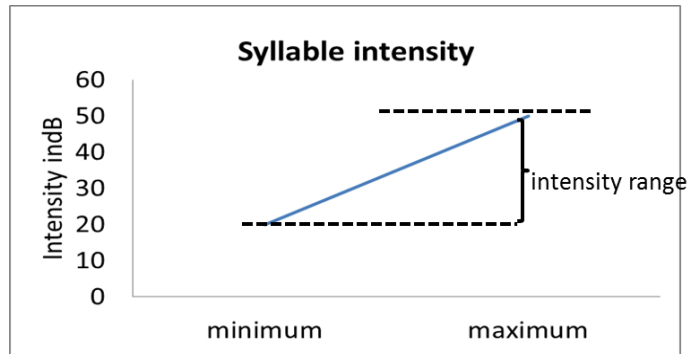


Figure 4. The Maximum intensity and Minimum intensity of a syllable and its intensity range.

### 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), [ $\beta = 28.634$ , S.E. = 9.307,  $t = 3.077$ ,  $p < 0.01$ ].

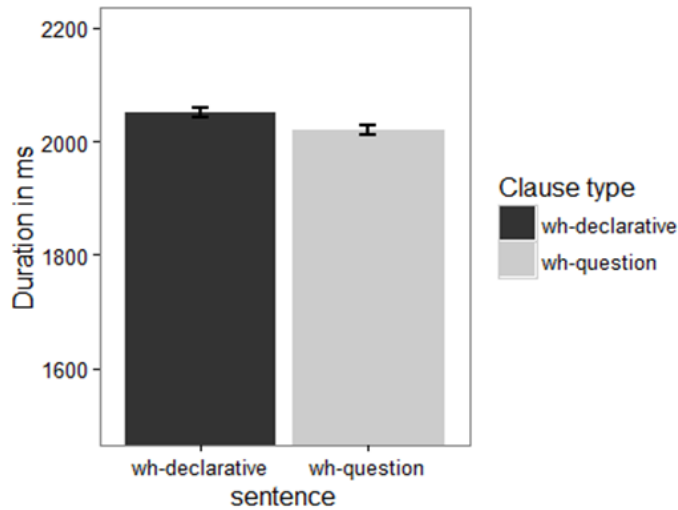


Figure 5. Mean sentence duration in ms with error bars showing standard error across clause types.

**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 *diǎnr* and of the *wh*-word. Specifically, the duration of the Subject, Verb plus *le* and *diǎnr* 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), [ $\beta = 5.626$ ,  $p < 0.01$ ;  $\beta = 27.993$ ,  $p < 0.001$ ;  $\beta = 5.339$ ,  $p < 0.001$ ]. 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.

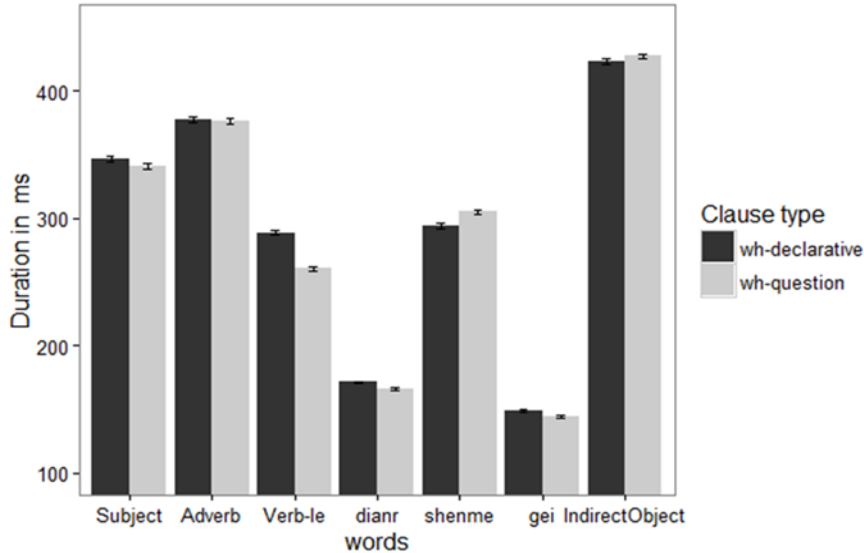


Figure 6. Mean word duration in ms with error bars showing standard error across clause types.

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

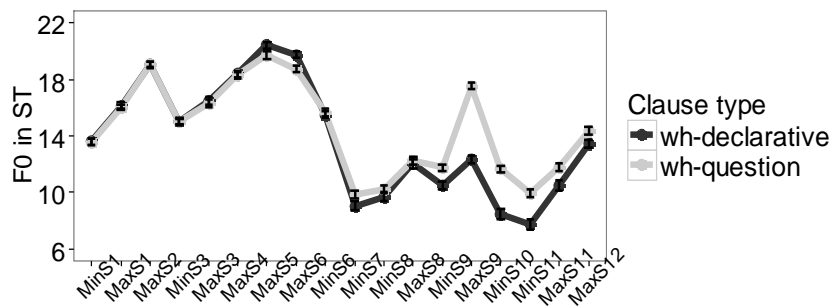
	Estimate $\beta$	Std. Error	$t$ -value	$p$ -value
subject	5.626	1.860	3.024	< 0.01
adverb	1.146	3.039	0.377	> 0.1
verb- <i>le</i>	27.993	2.730	10.253	< 0.001
<i>diǎnr</i>	5.339	1.239	4.310	< 0.001
<i>shénme</i>	-11.065	3.588	-3.084	< 0.01
preposition phrase	-0.397	3.818	-0.104	> 0.1

**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*).

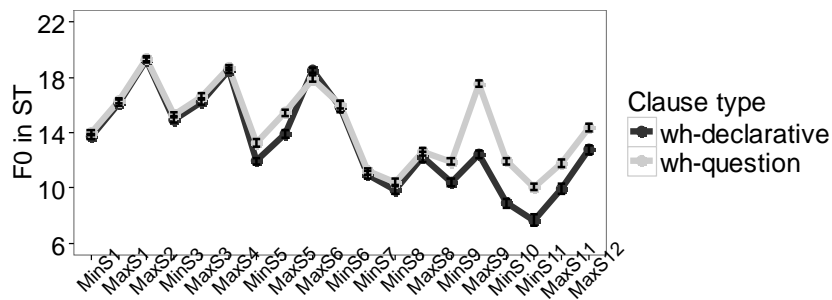
In the pre-*wh*-word region, we also found F0 differences at the verb when it bears T2. The F0-minimum and F0-maximum of the T2 verb in *wh*-declaratives ( $\bar{x} = 11.947$  ST,  $\bar{x} = 13.929$  ST) are lower than that of *wh*-questions ( $\bar{x} = 13.274$  ST,  $\bar{x} = 15.465$  ST) respectively,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



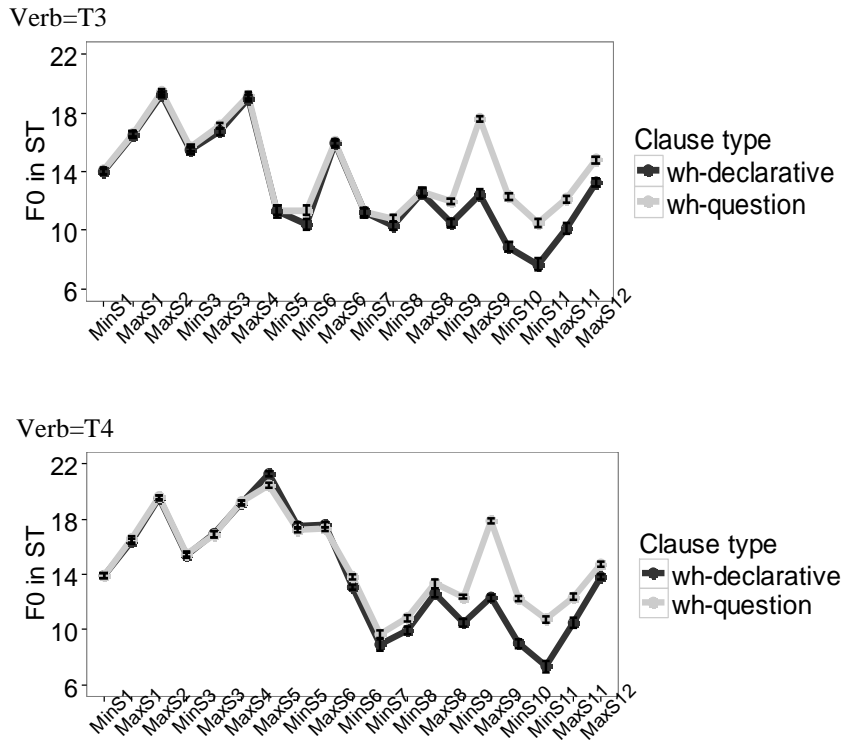


Figure 7. Stylized<sup>13</sup> means of F0 curves in ST across clause types with error bars showing standard errors.

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

	Estimate $\beta$	Std. Error	<i>t</i> -value	<i>p</i> -value
F0-min <i>shén</i>	-0.714	0.171	-4.167	<0.001
F0-min <i>me</i>	-1.630	0.263	-6.199	<0.001
F0-max <i>me</i>	-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 <i>gěi</i>	-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

<sup>13</sup> 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 $\beta$	Std. Error	<i>t</i> -value	<i>p</i> -value
F0 range <i>me</i>	-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.

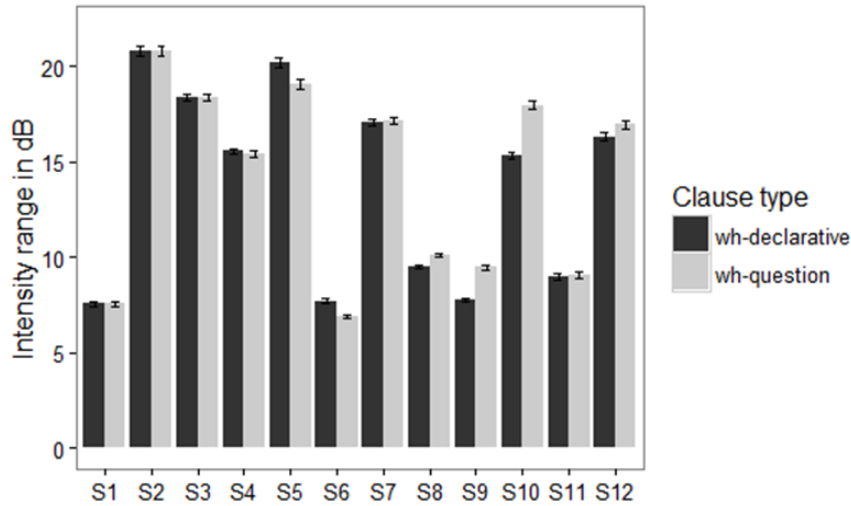


Figure 8. Mean intensity range of each syllable (S) across clause types with error bars showing standard errors.

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

	Estimate $\beta$	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.

