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

Yang, Y.

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.

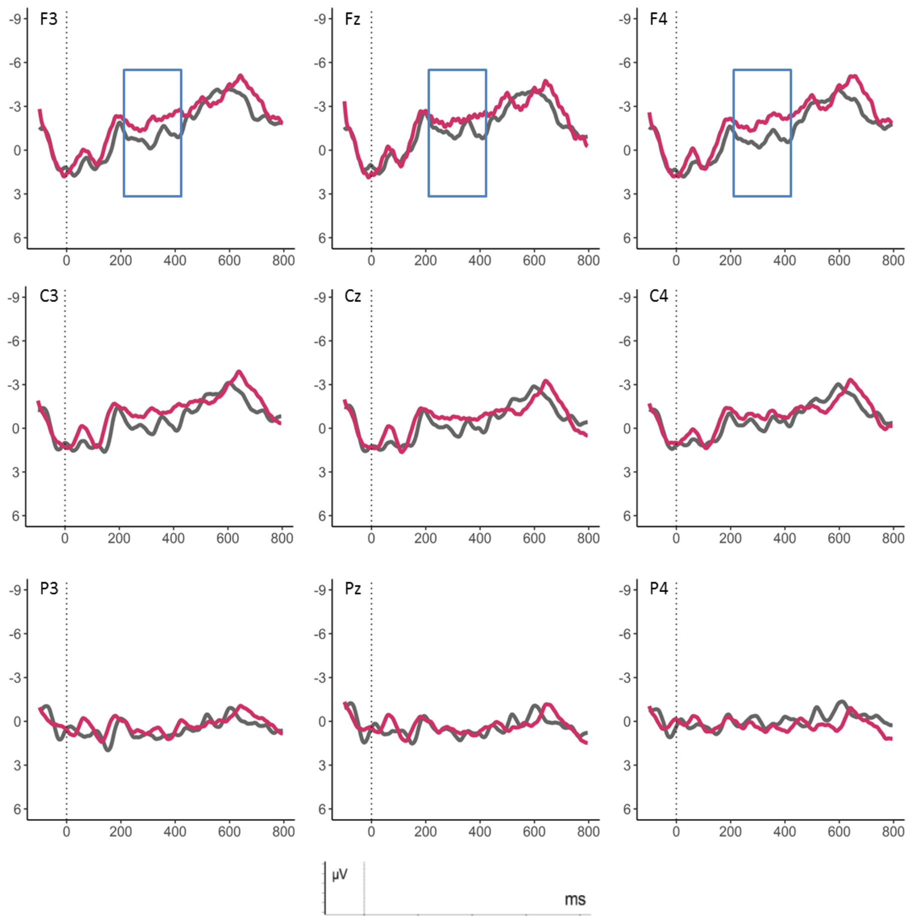

Figure 6. Grand average ERPs at 9 exemplar electrodes for each ROI time-locked to the onset of the critical word *shénme* (condition (9a) *wh*-declarative is in black, and condition (9c) (D-Q) in red).

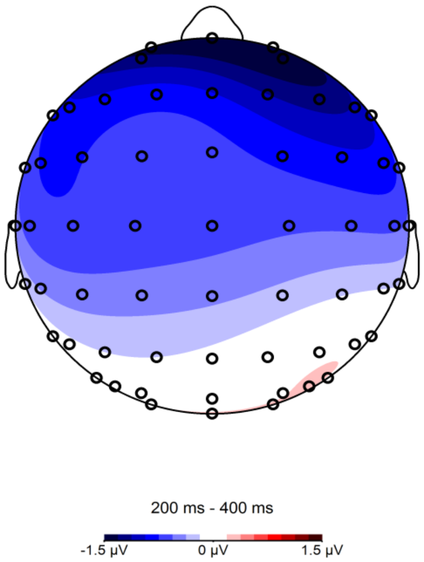

Figure 7. Topographic distributions of the mean ERP differences at the 200-400 ms time window. Condition (9c) (D-Q) was compared with condition (9a) (*wh*-declarative).

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.

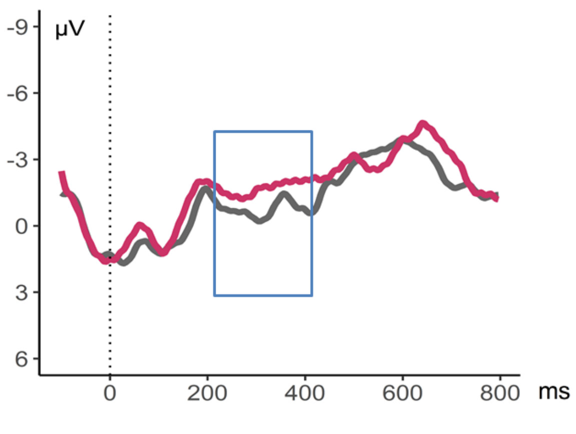

Figure 8. Grand average ERPs at anterior position time-locked to the onset of the critical word *shénme* (condition (9a) *wh*-declarative in black and condition (9c) (D-Q) in red).

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.

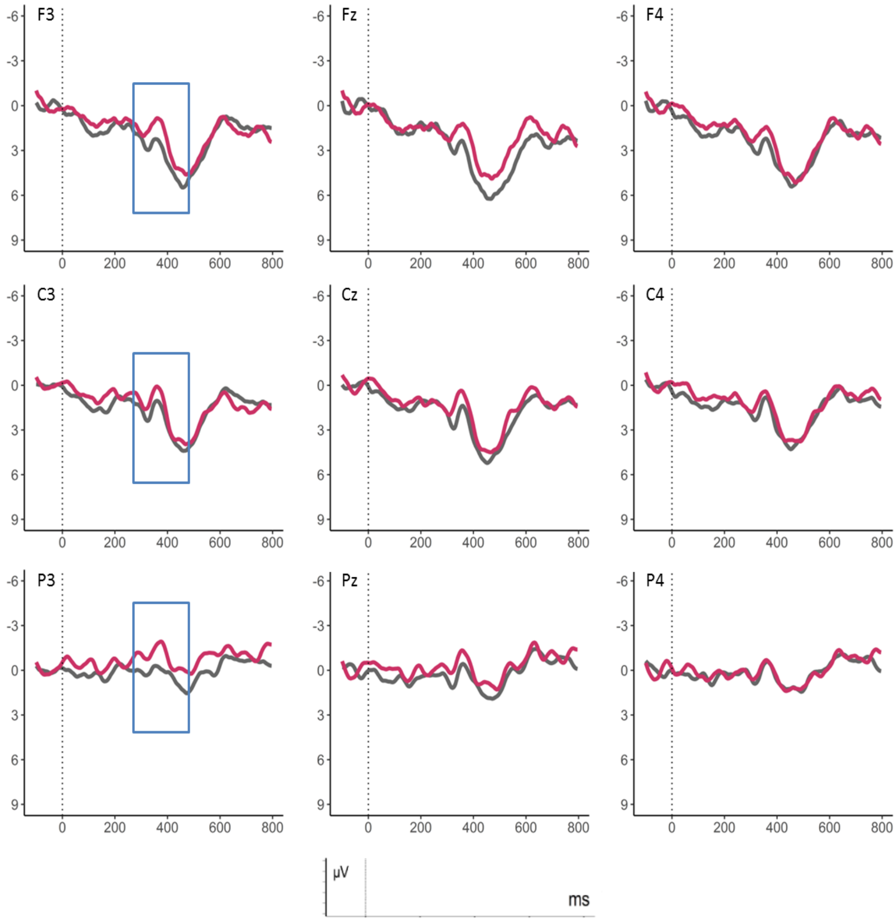

Figure 9. Grand average ERPs at 9 exemplar electrodes time-locked to the onset of the subject between condition (10a) D-D in black, and condition (10c) D-Q in red.

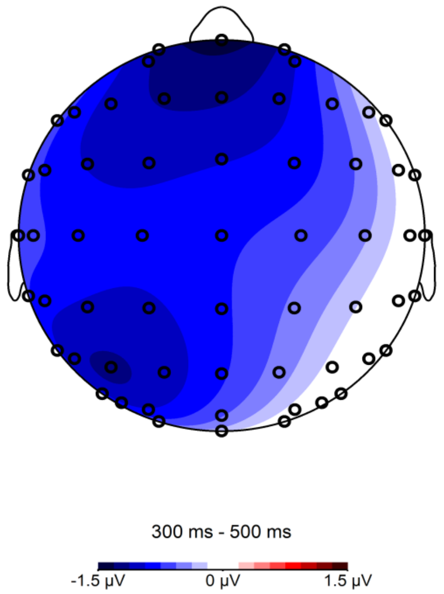

Figure 10. Topographic distributions of the mean ERP differences at the 300-500 ms time window. Condition (10c) (D-Q) was compared with condition (10a) (D-D).

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.

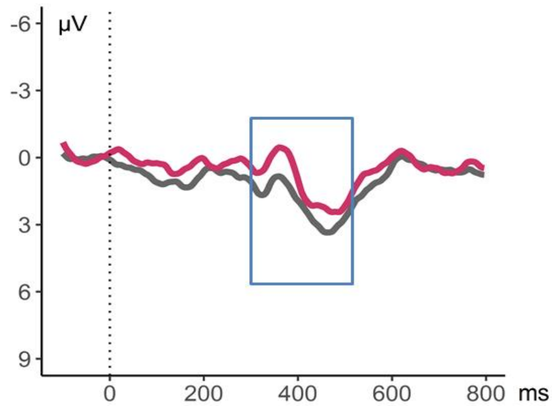

Figure 11. Grand average ERPs for left hemisphere electrode sites time-locked to the onset of the subject with condition (10a) D-D in black and condition (10c) D-Q in red.

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.

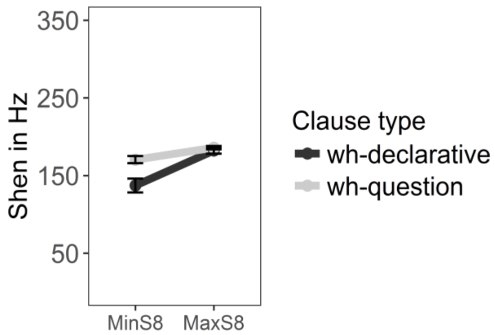

Figure 12. Stylized mean F0 at *shén* between *wh*-declaratives and *wh*-questions with error bars showing standard errors.

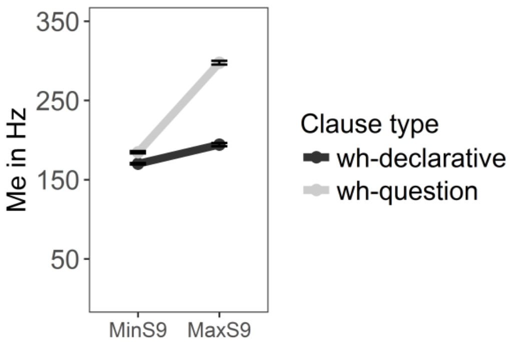

Figure 13. Stylized mean F0 at *me* between *wh*-declaratives and *wh*-questions with error bars showing standard errors.

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.

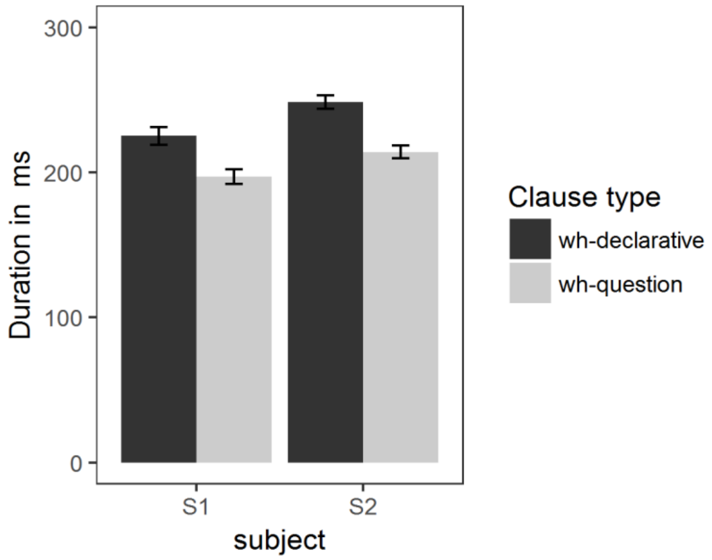

Figure 14. Mean syllable duration of the subject in ms with error bars showing standard errors between *wh*-declaratives and *wh*-questions.

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

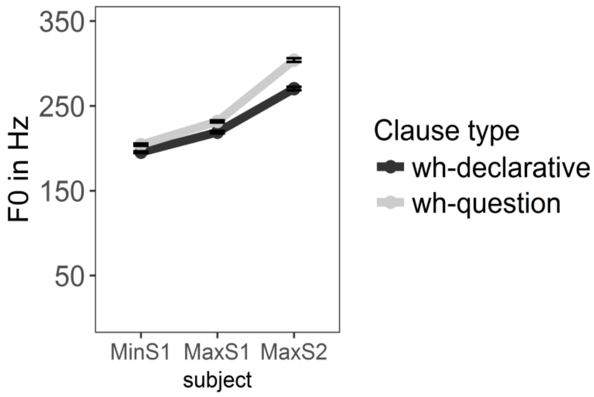

Figure 15. Stylized mean F0 at the subject between *wh*-declaratives and *wh*-questions with error bars showing standard errors.

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).