<table>
<thead>
<tr>
<th>Title</th>
<th>Predicting the development of interrogative forms and functions in early years: a corpus-based study of Mandarin-speaking young children</th>
</tr>
</thead>
<tbody>
<tr>
<td>Author(s)</td>
<td>Li, H; Jing, M; Wong, ECM</td>
</tr>
<tr>
<td>Citation</td>
<td>Journal of Child Language, 2017, v. 44 n. 1, p. 216-238</td>
</tr>
<tr>
<td>Issued Date</td>
<td>2017</td>
</tr>
<tr>
<td>URL</td>
<td><a href="http://hdl.handle.net/10722/229522">http://hdl.handle.net/10722/229522</a></td>
</tr>
<tr>
<td>Rights</td>
<td>Journal of Child Language. Copyright © Cambridge University Press.; This work is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License.</td>
</tr>
</tbody>
</table>
Predicting the development of interrogative forms and functions in early years: a corpus-based study of Mandarin-speaking young children*

HUI LI, MENGGUO JING AND EILEEN CHIN MEI WONG

The University of Hong Kong, PRC

(Received 21 May 2015 – Revised 10 October 2015 – Accepted 19 January 2016)

ABSTRACT

This study examined the development of and possible predictors of interrogative forms and functions in early childhood Mandarin. All the interrogatives drawn from the Early Child Mandarin Corpus (168 children 2;6, 3;6, 4;6, and 5;6) were analyzed. The main results indicated that (i) there were significant age effects in interrogative forms and functions, with the periods between the ages of 3;6 and 4;6 and between the ages of 2;6 and 3;6 being critical in the early acquisition of interrogative forms and functions, respectively; (ii) the form–function preference was verified, with wh-questions being primarily used to seek information (RfI), and intonation/echo and rhetorical questions being used to request action (RfA); (iii) more than half (59.5%) of the Mandarin interrogatives were used for RfI, whereas only 38.9% of them were used for RfA; and (iv) age, TV viewing time, and parent–child conversation time were the significant predictors of interrogative development.

INTRODUCTION

A large volume of research has been conducted over the last century to understand fully the development of interrogative forms and functions in English-speaking children (see Chouinard, 2007, for a review). However, Mandarin Chinese, the most widely spoken language in the world (Graddol, 2004), has not been equally empirically explored. Recently, there have been Chinese studies examining the development of interrogative forms in L1 children (Fahn, 2003; Zhou, 2015), and forms


[*] This paper is based on a project (RGC Ref No. 747109) funded to the first author by the Research Grants Council of the Government of the Hong Kong SAR. Thanks are given to all the participating children and their parents, Professor A. H. Liu, Dr Eva Chen, and the student research assistants. Address for correspondence: Dr Hui Li, Men Wah Complex Room 509, Faculty of Education, The University of Hong Kong, Hong Kong. tel: (852) 39175248; fax: (852) 2858-5649; e-mail: huili@hku.hk
(Ding, 2006; Gao, 2009; Yuan, 2007) and functions (Wang, 2010; Yuan, 2013; Zhu & Wu, 2011) in L2 learners. However, very few were conducted with representative samples to establish the developmental norms or patterns of interrogative form and function in Mandarin Chinese. We have already conducted a corpus-based study on early question acquisition in Cantonese-speaking young children in Hong Kong (Li, Tse, Wong, Wong & Leung, 2013). Mandarin Chinese, however, differs greatly from Cantonese in terms of lexicon, syntax, and pronunciation (Tse & Li, 2011) and therefore deserves a similar study. In addition, the Mandarin-speaking children in the present study have a monolingual language environment in Beijing, whereas their counterparts in Hong Kong thrive in a ‘trilingual and biliterate’ society (Li & Rao, 2000). Comparing their acquisition of interrogative forms and functions can provide some cross-linguistic evidence for the effects of language environments. This study is thus aimed at filling the gap in the literature.

Another objective of this study is to examine further the cross-linguistic universality of the Information Requesting Mechanism (IRM) proposed by Chouinard (2007). Having reviewed all the studies on the early acquisition of English questions, Chouinard hypothesized that request for information (RfI) might be the universal and dominant function of children’s question. This hypothesis, however, was challenged by our previous study on early child Cantonese in 492 children aged 36, 48, and 60 months (Li et al., 2013), which found that request for action (RfA) rather than RfI might be the dominant interrogative function. This inconsistent finding was regarded as the result of the cultural difference between Cantonese- and English-speaking young children (Li et al., 2013). But it might also be caused by the age differences in samples or the linguistic differences between English and Cantonese. To verify this cultural influence, we need to conduct a similar study on Mandarin Chinese using samples that share the same Chinese culture and communication task but speak different languages. Therefore, in this study, we thoroughly examine the interrogatives elicited from a similar early child Mandarin corpus and empirically explore the factors influencing the early acquisition of interrogatives in order to determine whether the IRM is applicable to Mandarin Chinese.

**Development of interrogative forms in Mandarin Chinese**

Mandarin Chinese has a very complex interrogative system, which is composed of a large number of members: question words (i.e. ‘what’, ‘who’), question forms (i.e. the ‘A-not-A’ construction), question mood particle (i.e. *ma*), and intonation (Gao, 2007). The various combinations of these members could make a question ambiguous and even controversial.
for Chinese speakers, and even linguists, in terms of the scope of questioning, the degree of question mood, and the degree of grammaticalization of various question forms. This could easily lead to misunderstanding and miscommunication (Gao, 2000) and make the acquisition of interrogative forms a really challenging task (Gao, 2007) for young Mandarin-speaking children.

The Chinese interrogative form is challenging not only for young children but also for linguists, who have not achieved any consensus on a typology (Gao, 2000; Matthews & Yip, 1994; Ross & Ma, 2006; Tse & Li, 2011). Liing (2014) identified four major interrogative forms: particle questions, ‘háishì’ questions, ‘A-not-A’ questions, and wh-questions. Also, Chinese linguists have explored A-not-A questions in different syntactic contexts (Cheng, 1984), and their modularity and forms (Huang, 1991; Hsieh, 2014). In this study, we adopted the most comprehensive and systematic typology, which was proposed by Matthews and Yip (1994), as it has been verified by our previous corpus-based studies on early child Cantonese (Tse & Li, 2011; Li et al., 2013). As Cantonese has many more interrogative forms than Mandarin Chinese (Li et al., 2013; Li et al., 2015), we used the eight forms commonly shared by Cantonese and Mandarin interrogatives to allow a cross-linguistic comparison between Cantonese and Mandarin interrogatives (for details, please see ‘Coding of the interrogatives’ below).

Currently, the most common topic for early interrogative studies concerns wh-questions in Mandarin Chinese. Fahn (2003) studied the acquisition of monoclausal wh-questions and found that subject wh-questions were much easier to learn than object wh-questions for Chinese children aged between four and six years. Zhou (2015) found that young children exhibited an adultlike sensitivity to the licensing environments for the non-interrogative use of wh-words. Su, Jin, Wan, Zhang, and Su (2014) found that young children with autism spectrum disorder (ASD) could understand the question readings rather than the statement readings of wh-words. Also, there have been other wh-question studies on L1–L2 transfer and comparison among Japanese speakers (Yuan, 2007) and English speakers (Yuan, 2013), and on different strategies for the acquisition of different interrogative forms (Gao, 2009).

Another common topic is the sequence of question acquisition in Chinese children. Ding (2006) found that the sequence of yes/no question acquisition was ma question > tag question and ba question > rhetorical question, and this sequence was roughly similar for L1 and L2 Chinese learners. Wong and Ingram (2003) found that the sequence of Cantonese question acquisition between one and three years of age was particle/intonation questions > wh-questions > A-not-A questions. They tended to believe that syntactic complexity, cognitive complexity, and personal/social factors, as
well as their interactions, could have determined the sequence of question acquisition. In addition, our previous Cantonese studies (Li et al., 2013; Tse & Li, 2011) found that girls produced more types of interrogative forms than boys, and that there was a significant form–function preference, with \textit{wh}- and multiple questions being primarily used to seek information. Is the same pattern found in the interrogative development of Mandarin-speaking young children? This study addresses this question by examining the development of interrogative forms with a large-scale sample of Mandarin-speaking young children.

\textit{Development of interrogative functions in Mandarin Chinese}

Interrogative function refers to the pragmatic function of questioning (Li et al., 2013), and it has been widely and repeatedly studied in the past two centuries since Sully (1896) identified two major functions of interrogatives: asking for information (‘what’, ‘how old’, ‘where’, ‘who’) and asking for reasons and causes. Piaget (1926) created a complete typology of interrogative functions: causality, reality and history, action and intention, classification and evaluation, and calculation. Later on, Smith (1933), Holzman (1972), and Sachs and Devin (1976) further categorized child questions into eight categories: external world, internal state, request, imperative, attention getting, instruction, confirmation, and repetition. Since the 1990s, a corpus approach has been employed to investigate the development of interrogative functions (Freed, 1994; Sinclair & Gessel, 1990). It is important to note that half of those functions identified in previous studies (Freed, 1994; Holzman, 1972; Piaget, 1926; Sinclair & Gessel, 1990; Sully, 1896) are asking for information (RfI), indicating that interrogatives often serve the function of information seeking.

Recently, Chouinard (2007) conducted a systematic literature review and four separate studies on early interrogative development in young American children. Together, the results of her four studies supported the existence of the IRM as a way for children to learn about the world. As RfI was the most dominant function (81\%) of American children’s questions in her studies, she hypothesized that RfI might be “a universal function of children’s interrogatives”. Tse and Li (2011) used Chouinard’s framework to test data drawn from their Early Child Cantonese Corpus, and developed a typology of Cantonese interrogative functions: (i) request for information (RfI) about names, locations, descriptions, and so forth; (ii) request for action (RfA), such as to draw attention to themselves, to declare something, to ask permission to do something, to generate conversation, to seek confirmation or verification, to suggest certain actions, and to prohibit some actions that they dislike; (iii) emotional
expression to convey their annoyance (anger) or disappointment (dissatisfaction); and (iv) egocentric speech, which is directed to no one but themselves. Using this typology, our previous study (Li et al., 2013) challenged the generality and applicability of Chouinard’s IRM hypothesis and found that request for action (RfA) (rather than RfI) was the dominant function of early childhood Cantonese. This contrary finding, however, needs to be verified further with empirical evidence from an early Mandarin corpus that shares the same culture and communication task as in the Cantonese study but involves different languages. If Cantonese- and Mandarin-speaking young children share the same pattern, which is different from that of English, we could attribute this inconsistent finding to cultural influences. If not, the factor of language (instead of culture) should be highlighted.

Furthermore, previous studies have found cross-linguistic evidence to support the form–function preference in early childhood interrogatives. James and Seebach (1982) found that yes/no questions were primarily used as directives, whereas wh-questions were employed to initiate and maintain conversations. A similar pattern was found by Shatz (1979) and Celce-Murcia and Larsen-Freeman (1999) with English-speaking children. A previous study on early Cantonese questions (Li et al., 2013) also found similar patterns: wh-questions were primarily used to request information, and intonation/echo and the rhetorical questions were used to request action. The present study further explores whether this preference is cross-linguistic, based on the evidence from a corpus-based study of Mandarin-speaking young children.

Predictors of early language development

Many family, preschool, and environmental factors are known to have an influence on early language acquisition. Family socioeconomic status (SES) has been identified as one of these factors (Hart & Risley, 1992, 1995; McLoyd, 1998). Home language input such as parental support has been found to be another predictor of young children’s acquisition of wh-questions (Gullo, 1982a, 1982b). Other factors such as a supportive home environment with rich reading resources, parent–child shared reading, and maternal book-reading strategies can also make a great contribution to early language development (DeJong & Leseman, 2001; Reese & Fivush, 1996; Roberts, Jurgens & Burchinal, 2005). For example, Rowland, Pine, Lieven, and Theakston (2003) found that the acquisition order of wh-questions could be predicted by the frequency with which the mother used particular wh-words and verbs in conversations with the child.

Another important predictor of early language acquisition is TV viewing time. Recent studies have repeatedly documented a negative association
between early TV exposure and language development (Barr, Lauricella, Zack & Calvert, 2010; Hudon, Fennell & Hoftyzer, 2013; Pempek, Kirkorian & Anderson, 2014; Tomopoulos, Dreyer, Berkule, Fierman, Brockmeyer & Mendelsohn, 2010). Nathanson and Rasmussen (2011) found that mother–child communication was less frequent and less verbally responsive when dyads viewed TV. Duch et al. (2013) found that children who watched more than 2 hours of television a day had lower communication scores. A very recent neuroimaging study (Takeuchi et al., 2013) also confirmed the negative effect of TV viewing on a child’s verbal intelligence quotient (VIQ). However, Takeuchi et al.’s study also found some positive effects of TV viewing on the regional gray/white matter volume of the frontopolar and medial prefrontal areas, visual cortex, hypothalamus/septum, and sensorimotor areas in young children. All these brain areas are associated with cognitive and intellectual abilities. Does TV viewing have any negative or positive effects on early interrogative development? Another aim of the present study is to determine whether TV viewing has any negative or positive effects on early interrogative development in Mandarin-speaking young children. The following research questions were set up to guide this study:

1. What are the interrogative forms produced by a representative sample of Mandarin-speaking young children? Are there any age and gender differences in the development of interrogative forms?

2. What interrogative functions are present in early childhood Mandarin? Are there any age and gender differences in the development of interrogative functions?

3. What are the predictors of early interrogative development in Mandarin Chinese?

**METHOD**

**Participants**

The Early Childhood Mandarin Corpus used in this study consisted of the utterances produced by 168 Mandarin-speaking preschoolers randomly sampled from eight preschools located in the four major districts of Beijing: Chaoyang, Dongcheng, Xicheng, and Haidian. This sample represented children from four age groups (ages 2;6, 3;6, 4;6, and 5;6), with twenty-one boys and twenty-one girls in each age group. As shown in Table 1, the average age of each group was very close to their target age, except for group 4 whose average age was 5;4. This is because the sampling was carried out very early in the first semester, so most of the kindergarteners had not achieved the target age. All the participants were native speakers of Mandarin, and their parents and teachers also spoke
Mandarin at home and at preschool, respectively. Most of the parents were degree holders, and some children had taken extracurricular activities and private tutoring in the early years (see Table 1). The average TV viewing time ranged between 0.8 to 1.1 hours per day, and the parent–child conversation time ranged between 1.5 to 1.9 hours per day.

**Communication task**

Participants of the same age were paired randomly (boy/girl, boy/boy, or girl/girl) and were left in the play corner to play for 30 minutes. A set of regular toys, including cooking materials, faux food and fruit, furniture and electrical appliances, hospital materials, and vehicles, was set up in the corner (Li et al., 2013). The participants were free to play with any toys and engaged in any type(s) of play they wished. They were encouraged to talk while they were playing, and their conversation was videotaped with a digital video camera. The assistant researcher observed but did not intervene during the sessions, and there were no other children in the corner.

**Transcription**

All the videotaped conversations were transcribed by trained research assistants to a level of detail that captured all words and word fragments.
(including overlapping speech) audible to the ear, and non-lexical fillers (e.g. *ah*), as well as other vocalizations (e.g. laughter). Next, other research assistants double-checked the transcripts against the video clips to ensure accuracy of transcription. Then, the Chinese script was segmented into utterances by the trained research assistants.

**Coding of the interrogatives**

Questions and interrogatives were judged according to the standards given in Li *et al.*’s (2013) early childhood Cantonese study. A question was defined as a linguistic expression of enquiry that invites or calls for a reply. An interrogative was defined as the linguistic form of a question and could be a word, a phrase, or a sentence. All the interrogatives were analyzed and coded based on the coding systems as follows. Two authors of this paper coded 10% of the interrogatives elicited from the corpus, and their mutual agreement was more than 93%, indicating excellent inter-coder reliability. Then, one of the co-authors coded the rest of the data.

1. **Intonation and echo questions:** raising the intonation of a declarative or echoing others’ declaratives to construct an interrogative form:
   
   "皮球不见了?"
   \(\text{Piqiu bujian le?}\) 
   Ball disappeared SFP
   ‘The ball disappeared?’

2. **Yes/no questions:** adding a particle like *ma* (吗), *ba* (吧), *aa* (呀) to a declarative sentence to ask for yes or no answers:
   
   "你 想 喝 水 吗?"
   \(\text{Ni xiang heshui ma?}\) 
   You want drink water SFP
   ‘Do you want to drink water’

3. **Wh- questions:** using *wh*-words such as *shenme* (what), *shui* (who), *wei shenme* (why) (为什么), and so on to produce an interrogative:
   
   "这是什么呀?"
   \(\text{Zheshi shenmo ya?}\) 
   This is what SFP
   ‘What is this?’

4. **A-Not-A questions:** repeating (or reduplicating) the verb or adjective with the negative marker *bu* in between to deliver an interrogative:
   
   "你认不认识他?"
   \(\text{Ni ren bu renshi ta?}\) 
   You know not know him
   ‘Do you know him or not?’
(5) Alternative question: conjoining two phrases with the conjunction
haishi (还是) or huozhe (或者) to offer a choice between two alternatives:

今天是星期一还是星期二?

Jintian shi xinqiyi haishi xinqier?

‘Is today Monday or Tuesday?’

(6) Rhetorical question: an interrogative expressing negative meaning with
positive form or vice versa, which is not supposed to be answered:

我怎么知道呢？

Wo zenmo zhidao ne?

‘How should I know?’ (I don’t know, indeed)

(7) Indirect questions: using a clause beginning with wh-words to report a
question, which is normally called a reported question:

你知到谁要这本书吗？

Ni zhidao shui yao zheben shu ma?

You know who want this book?

‘Do you know who wants to have this book?’

(8) Multiple questions: two or more question words appearing in a
statement to request more information, as in the English “Where and
when are we going to meet?”:

我们几点钟在哪里见面?

Women jidianzhong zai nani jianmian?

‘When and where should we meet up?’

Parental report

One parent of each participating child was invited to complete the Home
Language Environment Survey (HLES), which was developed from the
Home Literacy Environment Index (HLEI) (Li & Rao, 2000). The HLES
consists of twenty-nine questions that tap parental beliefs and practices
about the home language environment, language input, parental
involvement, language resources, and learning strategies. The HLEI has
been developed and validated for use in Beijing, Hong Kong, and
Singapore (Li & Rao, 2000).

RESULTS

Altogether, 3,850 instances of Mandarin interrogative were elicited from
the corpus, which included all the utterances produced by the 168
children. On average, each participant uttered 22·92 interrogative types
during the half-hour period of communication. Some 2,171 (56·4%) of the
interrogatives were produced by the boys, and the remaining 1,679 (43.6%) by the girls. The interrogative types produced by the 2:6, 3:6, 4:6, and 5:6 age groups were 734 (19.1%), 451 (11.7%), 936 (24.3%), and 1729 (44.9%), respectively. The following analyses were executed to explore the developmental trends in interrogative form and function, possible form—function preferences, and the predictors of interrogative development.

The development of interrogative forms

First, a descriptive analysis was conducted to examine the occurrence of all the interrogative forms proposed in this study. As shown in Table 2, the four age groups produced five types of Mandarin interrogative forms: ‘intonation/echo’, ‘yes/no question’, ‘wh-question’, ‘A-not-A’, and ‘rhetorical question’. However, ‘alternative question’, ‘indirect question’, and ‘multiple question’ were not found or were very rarely produced. In particular, wh-questions (57.5%) were very frequently produced, followed by intonation/echo (25.7%) and yes/no question (12.8%).

Next, a set of ANOVAs with ages by sex as the independent variables was applied to the three dependent variables: the total number of interrogative types, forms, and functions. Significant Age effects were found in the total number of interrogative types (F(3,165) = 54.21, p < .001, η² = .247), forms (F(3,165) = 25.92, p < .001, η² = .323), and functions (F(3,165) = 28.98, p < .001, η² = .348). There was no significant Gender effect in the total number of interrogative types (F(1,166) = 2.69, p > .05, η² = .015), forms (F(1,166) = 2.97, p > .05, η² = .018), and functions (F(1,166) = 3.26, p > .05, η² = .020). In addition, there was no Age × Gender effect in the total number of interrogative types, forms, and functions (ps > .05). A set of post-hoc Tukey (HSD) tests found no significant differences in interrogative forms between the 2:6 and 3:6 age groups (ps > .05), whereas significant differences were found between the 3:6 age group and the 4:6/5:6 age groups (ps < .05). The older children produced significantly more types of interrogative form than the children under age 3:6. This indicated that between the ages of 3:6 and 4:6 might be the critical period in the early development of interrogative forms.

Third, follow-up one-way ANOVAs were conducted to explore age differences in the five interrogative forms found in this study. Significant age differences were found in the distribution of wh-questions (F(1,4) = 5.84, p < .001), with post-hoc Tukey tests showing a statistically significant difference between the 3:6 and 4:6 age groups (p < .001), and between the 3:6 and 5:6 age groups (p < .01). The older children produced significantly more types of wh-questions than the children under age 3:6. In addition, significant age differences were also found in the distribution of rhetorical questions (F(1,4) = 2.97, p < .05), with post-hoc Tukey tests showing a
EARLY INTERROGATIVE DEVELOPMENT IN MANDARIN

### Table 2. Age differences in interrogative form and function

<table>
<thead>
<tr>
<th>Function</th>
<th>Age 2.5 (n = 734)</th>
<th>Age 3.5 (n = 451)</th>
<th>Age 4.5 (n = 936)</th>
<th>Age 5.5 (n = 1729)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean (%)</td>
<td>Mean (%)</td>
<td>Mean (%)</td>
<td>Mean (%)</td>
</tr>
<tr>
<td>Intonation / Echo</td>
<td>180 (25.7%)</td>
<td>129 (28.6%)</td>
<td>278 (29.7%)</td>
<td>450 (26%)</td>
</tr>
<tr>
<td>Yes/No Question</td>
<td>94 (12.8%)</td>
<td>74 (16.4%)</td>
<td>175 (18.7%)</td>
<td>308 (17.8%)</td>
</tr>
<tr>
<td>Wh-Question</td>
<td>424 (57.5%)</td>
<td>237 (52.5%)</td>
<td>438 (46.5%)</td>
<td>841 (48.6%)</td>
</tr>
<tr>
<td>A-Not-A</td>
<td>19 (2.6%)</td>
<td>5 (1.1%)</td>
<td>35 (3.7%)</td>
<td>82 (4.7%)</td>
</tr>
<tr>
<td>Alternative question</td>
<td>0 (0%)</td>
<td>0 (0%)</td>
<td>2 (0.2%)</td>
<td>2 (0.1%)</td>
</tr>
<tr>
<td>Rhetorical question</td>
<td>6 (0.8%)</td>
<td>5 (1.1%)</td>
<td>6 (0.6%)</td>
<td>46 (2.7%)</td>
</tr>
<tr>
<td>Indirect question</td>
<td>0 (0%)</td>
<td>0 (0%)</td>
<td>0 (0%)</td>
<td>0 (0%)</td>
</tr>
<tr>
<td>Multiple questions</td>
<td>0 (0%)</td>
<td>0 (0%)</td>
<td>0 (0%)</td>
<td>0 (0%)</td>
</tr>
<tr>
<td>Wrong use / Unclassified</td>
<td>2 (0.3%)</td>
<td>1 (0.2%)</td>
<td>2 (0.2%)</td>
<td>0 (0%)</td>
</tr>
<tr>
<td>Request for information</td>
<td>507 (68.8%)</td>
<td>253 (56.1%)</td>
<td>510 (54.9%)</td>
<td>1008 (58.3%)</td>
</tr>
<tr>
<td>Request for action</td>
<td>214 (29%)</td>
<td>186 (41.2%)</td>
<td>409 (44%)</td>
<td>715 (41.4%)</td>
</tr>
<tr>
<td>Emotional expression</td>
<td>0 (0%)</td>
<td>0 (0%)</td>
<td>2 (0.2%)</td>
<td>1 (0.1%)</td>
</tr>
<tr>
<td>Egocentric speech</td>
<td>10 (1.4%)</td>
<td>1 (0.2%)</td>
<td>3 (0.3%)</td>
<td>4 (0.2%)</td>
</tr>
<tr>
<td>Wrong use / Unclassified</td>
<td>1 (0.1%)</td>
<td>1 (0.2%)</td>
<td>5 (0.5%)</td>
<td>1 (0.1%)</td>
</tr>
</tbody>
</table>

Note: The percentages in brackets were calculated just for each age group.

Statistically significant differences were found between the 3:6 and 5:6 age groups \( p < .05 \). The 5:6 age group produced significantly more types of rhetorical questions than the 3:6 age group. These results further supported the conclusion that there is a dramatic growth in interrogative forms between the ages of 3:6 and 4:6 in Mandarin-speaking children.

Last, gender differences were explored with a set of one-way ANOVA. All the interrogative forms appeared to be quite evenly distributed between girls and boys, except for the *wh*-questions. A significant gender difference was found in the distribution of *wh*-questions \( F(1, 4) = 5.05, p < .05 \), with the girls (52%) tending to ask a greater proportion of *wh*-questions than the boys (49.3%) (see Table 3). In summary, a significant age effect was found in interrogative forms, and the period between the ages of 3:6 and 4:6 was identified as critical in the early acquisition of interrogative forms.

### The development of interrogative functions

As can be seen in Table 2, all the interrogative functions produced by the participants could be placed into the three major categories suggested by Chouinard (2007): (i) information-seeking (request for information, *RfI*), (ii) action-beseeching (request for action, *RfA*), and (iii) non-information-seeking. It was found that the interrogatives in early childhood Mandarin served the function of soliciting information (59.5%) more often than the
Wrong use / unclassified by asking questions, but fewer requests for information. In summary, a gender difference was found in interrogative functions, and a significant gender difference was only found in the distribution of RfI. Also, between 418
the ages of 2;6 and 3;6 seemed to be the critical developmental period for interrogative functions.

The form–function preference in Mandarin interrogatives

As shown in Table 4, a form–function preference was found in the production of interrogatives: the wh-questions (86.5%) were primarily used to request information, whereas the intonation/echo (79.8%) and the rhetorical (85.7%) questions were employed to request action. These outcomes, as shown in Table 4, indicated that the Mandarin-speaking young children tended (i) to seek information more frequently with wh-questions (86.5%); (ii) to request action more frequently with rising tone (79.8%) and rhetorical questions (85.7%); (iii) to be less likely to use wh-questions to request action (12.7%); and (iv) to reduce their use of rhetorical questions (14.3%) and a rising tone (19.8%) to elicit information.

Predictors of Chinese interrogative development

To explore the variables associated with early interrogative development, we conducted a correlation analysis on the variables involved. The correlation matrix indicated that the following factors were significantly correlated with individual production of question types, interrogative forms, and functions: preschooling experience, extracurricular experience, private tutoring in Chinese, early Chinese literacy, and TV viewing time (hours).
Next, to determine the relative contribution of these correlated factors, we conducted four-step hierarchical regression analyses with individual production of question types, interrogative forms, and functions as the dependent variables. Results are shown in Table 5. In Step 1, we entered age and gender to control for their effects. In Step 2, preschooling experience and curriculum type were entered. In Step 3, extracurricular experience and private tutoring in Chinese were entered. In Step 4, early Chinese literacy learning, TV viewing time, and parent–child conversation time were entered.

The change in $R^2$ between the four steps indicated that: (i) the age and gender of the children could jointly explain 26.9%, 32%, and 34.7% of the variation in individual production of interrogative types, forms, and functions, respectively. In addition, age turned out to be the most significant predictor of the production of interrogatives; (ii) preschooling experience and curriculum type, extracurricular experience, and private tutoring in Chinese could jointly explain 2.8%, 2.4%, and 2.7% of the individual differences in interrogative types, forms, and functions, respectively. However, none of them was a significant predictor; and (ii) early Chinese literacy learning, TV viewing time, and parent–child conversation time could jointly explain 5%, 2%, and 2.4% of the variation in individual production of interrogative types, forms, and functions, respectively. In particular, TV viewing time turned out to be the significant predictor of question production and interrogative functions: the greater the viewing time, the more the types of questions and functions produced by the children. However, it had no significant effect on the production of interrogative forms. It was very interesting to note that parent–child conversation time became a significant predictor of

<table>
<thead>
<tr>
<th>Function</th>
<th>Request for information</th>
<th>Request for action</th>
<th>Non-information seeking</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intonation / echo</td>
<td>207 (19.8%)</td>
<td>835 (79.8%)</td>
<td>4 (0.4%)</td>
<td>1046</td>
</tr>
<tr>
<td>Yes/No Question</td>
<td>340 (52.2%)</td>
<td>306 (47%)</td>
<td>5 (0.8%)</td>
<td>651</td>
</tr>
<tr>
<td>Wh-Question</td>
<td>1653 (86.5%)</td>
<td>242 (12.7%)</td>
<td>16 (0.8%)</td>
<td>1911</td>
</tr>
<tr>
<td>A-Not-A</td>
<td>59 (41.8%)</td>
<td>82 (58.2%)</td>
<td>0</td>
<td>141</td>
</tr>
<tr>
<td>Alternative question</td>
<td>1 (25%)</td>
<td>3 (75%)</td>
<td>0</td>
<td>4</td>
</tr>
<tr>
<td>Rhetorical question</td>
<td>9 (14.3%)</td>
<td>54 (85.7%)</td>
<td>0</td>
<td>63</td>
</tr>
</tbody>
</table>

NOTE: The percentages in brackets were calculated for each type of interrogative form.
question production, although the correlation between them was negative. This result indicated that the longer the parent–child conversation time, the fewer the questions produced by the child. In summary, this set of hierarchical regression analyses found that age, TV viewing time, and

<table>
<thead>
<tr>
<th>Variable</th>
<th>Beta</th>
<th>$R^2$</th>
<th>$\Delta R^2$</th>
<th>$F$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Predicting question production</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Step 1</td>
<td>Age</td>
<td>-0.503***</td>
<td>-0.269</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>Gender</td>
<td>0.129</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Step 2</td>
<td>Schooling experience</td>
<td>0.204</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Curriculum type</td>
<td>0.043</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Step 3</td>
<td>Extracurricular experience</td>
<td>0.210</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Private tutoring in Chinese</td>
<td>-0.152</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Step 4</td>
<td>Early Chinese literacy learning</td>
<td>0.096</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>TV watching time</td>
<td>0.179**</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Parent–child conversation time</td>
<td>-0.142*</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Predicting form types</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Step 1</td>
<td>Age</td>
<td>0.554***</td>
<td>-0.320</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>Gender</td>
<td>0.119</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Step 2</td>
<td>Schooling experience</td>
<td>0.158</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Curriculum type</td>
<td>0.024</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Step 3</td>
<td>Extracurricular experience</td>
<td>0.261</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Private tutoring in Chinese</td>
<td>-0.184</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Step 4</td>
<td>Early Chinese literacy learning</td>
<td>0.077</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>TV watching time</td>
<td>0.111</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Parent–child conversation time</td>
<td>0.078</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Predicting function types</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Step 1</td>
<td>Age</td>
<td>0.577***</td>
<td>-0.347</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>Gender</td>
<td>-0.123</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Step 2</td>
<td>Schooling experience</td>
<td>0.148</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Curriculum type</td>
<td>0.063</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Step 3</td>
<td>Extracurricular experience</td>
<td>0.311</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Private tutoring in Chinese</td>
<td>-0.265</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Step 4</td>
<td>Early Chinese literacy learning</td>
<td>0.001</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>TV watching time</td>
<td>0.133*</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Parent–child conversation time</td>
<td>-0.058</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

NOTES: * $p < 0.05$; ** $p < 0.005$; *** $p < 0.001$.  

TABLE 5. Summary of hierarchical regression analyses for home variables predicting Chinese interrogative development
DISCUSSION

As the largest investigation ever made so far into early Mandarin interrogatives, this study set out to examine rich corpus data in an attempt to ascertain the trends and predictors with respect to early interrogative development. Significant age differences were found for both interrogative forms and functions, whereas significant gender differences were found only in the distribution of *wh*-questions and RfI. The widely reported form–function preference was verified, and the following predictors were also identified: age, TV viewing time (hours), and parent–child conversation time (hours). This section discusses these findings and their implications for early language learning and teaching.

Age differences in interrogative development

The present study found that the Mandarin-speaking children produced five types of Mandarin interrogative forms, but not the three complicated forms. In addition, the study found a significant age effect in interrogative forms and functions, and identified the periods between the ages of 3;6 and 4;6 and between the ages of 2;6 and 3;6 as critical in the early acquisition of interrogative forms and functions, respectively. All these findings, however, were inconsistent with those of our previous study on early Cantonese interrogatives (Li et al., 2013), in which all eight interrogative forms and functions were produced by the 3;0–6;0 age groups, and no age difference was found. These inconsistent findings reflect the two different developmental trajectories that the young Mandarin-speaking and the young Cantonese-speaking children followed in acquiring interrogative forms: in Beijing, the young children did not acquire all the interrogative forms and functions until age 5;6, and the periods between 3;6 and 4;6 and between 2;6 and 3;6 witnessed a rapid growth in interrogative forms and functions. In Hong Kong, on the other hand, the young children acquired all the interrogative forms and functions by the age of three, so there was no age difference.

This contrast indicated that the Cantonese-speaking young children developed interrogative forms and functions much earlier than their counterparts in Beijing. In addition to the differences in language (Cantonese versus Mandarin), there were significant differences in language environment: Beijing is a monolingual society, whereas Hong Kong is a typical ‘biliterate and trilingual’ society (Li & Rao, 2000). The Hong Kong children were immersed in Cantonese, English, and Mandarin environments (home, school, community, and media) in the early years.
and might have benefited from this experience. However, this inconsistent finding is empirically supported by the findings of Li and Chen’s (1997) longitudinal study on Mandarin-speaking children’s question acquisition. In their case study, a child was followed from age one to age five, and her interrogative utterances were analyzed. The results indicated that the child could produce all the interrogative forms listed in the case study around 4;6. Although their study and our study did not share exactly the same typology of interrogative forms, they both indicated that Mandarin-speaking young children tended to acquire some new interrogative forms after the age of three. This difference between early Cantonese and Mandarin interrogative acquisition might be associated with the differences in languages, samples, and language environments. Also, we tend to believe that early language environments affect question acquisition, which deserves further comparative investigations.

Another notable finding was that the 168 Mandarin-speaking children produced altogether 3,850 types of interrogative during the half-hour communication sessions, generating a questioning rate of 45.83 per child per hour. This is significantly more than the questioning rates (12.76 questions per hour) in Li et al.’s (2013) study with Cantonese-speaking children, and the questioning rates (26 questions per hour) in Tizard and Hughes’s (1984) study with English children. This finding indicated that the young Beijing children might be more talkative with their peers and tend to raise more questions than their Hong Kong counterparts. This difference might also reflect their different early language and learning experiences at home and in school. This, again, deserves further studies.

Gender differences in interrogative development

Very few significant gender differences were found in early Mandarin interrogative development. First, regarding the early interrogative forms, it was found that the girls produced relatively more wh-questions (in percentages) than did the boys. This finding is not very consistent with the findings of previous studies, which found that boys utilized more wh-questions (Davis, 1932; Li et al., 2013; Smith, 1933). As our studies on early childhood Cantonese (Li et al., 2013) and Mandarin shared the same research design and communication task, this inconsistent finding, again, might have been caused by the differences in samples (different age groups), languages (Cantonese versus Mandarin), and language environments between Beijing and Hong Kong. This study, for instance, found that the 2;6 age group produced many more wh-questions (57.5%) than the other age groups (ranging between 46.5% and 52.6%). More studies, however, are also needed to identify the real causes.
Second, the present study found that girls tended to make more requests for information (RfI) than the boys. This finding is inconsistent with that of our previous study, which found no significant gender difference in the interrogative functions produced by the Cantonese-speaking preschoolers (Li et al., 2013). As the two studies share the same culture, design, and task, this inconsistent finding indicated that the gender difference in interrogative functions might differ from Mandarin to Cantonese. Although Chouinard (2007) did not report any significant gender difference in her four studies, other studies on English-speaking children (Davis, 1932; Smith, 1933) have found that boys tend to ask more causal questions using wh-questions than girls. So far, three types of results have been reported concerning the early development of RfI functions: boys > girls (Davis, 1932; Smith, 1933), boys = girls (Chouinard, 2007; Li et al., 2013), and boys < girls (this study). In this connection, the only conclusion might be that gender difference in early interrogative functions could vary across languages, communication tasks, and social contexts. This casts some doubt on the cross-linguistic universality of gender differences in interrogative functions.

Universality of the IRM hypothesis

The second objective of this study was to examine the cross-linguistic universality of the Information Requesting Mechanism (IRM) proposed by Chouinard (2007), which has been challenged by our previous study on early Cantonese interrogatives (Li et al., 2013). The present study, however, found that more than half (59.5%) of the Mandarin interrogatives were requests for information (RfI), whereas only 38.9% of them were requests for action (RfA). This large-scale study has provided very solid cross-linguistic evidence to support Chouinard’s IRM, indicating that the IRM mechanism might be applicable across languages and circumstances. This finding, however, is inconsistent with that of our previous study on early Cantonese interrogatives (Li et al., 2013). The two studies were conducted by the same team and with the same research design and communication task, and the only difference is that the current study included a 2;6 age group. Also, the present study found that the 2;6 age group produced relatively more information-seeking interrogatives (70%) and significantly boosted the overall percentage of RfI interrogatives. This finding indicated that the young Mandarin-speaking children tended to make more requests for information (RfI) before the age of three, and that the dramatic changes occurred during the period between the ages of 2;6 and 3;6.

Coincidently, Chouinard’s four studies (2007) mainly focused on younger children, under age three or four, and her samples were relatively younger
than the samples in our early Cantonese corpus. Cantonese-speaking children
under age three might also produce more RfI in their interrogative functions.
Therefore, we tend to integrate our studies on early Cantonese and Mandarin
with Chouinard’s study on early English, and modify the IRM hypothesis as
follows: RfI predominates before age 3;6, while RfA predominates in the
post-3;6 years, as the dominance of RfI versus RfA might change over the
course of a child’s development. However, further empirical studies are
needed to verify this modified hypothesis.

Last, the present study has further verified the form–function preference
in early Mandarin interrogatives. The young children tended to use
wh-questions to request information, and intonation/echo questions to seek
action. In contrast, intonation/echo questions were less frequently used to
elicit information, and wh-questions were rarely used to request action.
This has verified the patterns found in our previous study on
Cantonese-speaking children (Li et al., 2013) and in Celce-Murcia and

The predictors of Mandarin interrogative development
The present study found that, in addition to the age of the children—the
most significant contributor—TV viewing time was a significant predictor
of interrogative types and functions. This finding implied that the greater
the TV viewing time, the more interrogative types and functions produced
by the children. This finding is not very consistent with recent studies,
which have consistently documented a negative association between early
TV exposure and language development (Barr et al., 2010; Hudon et al.,
2013; Pempek et al., 2014; Tomopoulos et al., 2010). However, none of
these studies used interrogatives as the language performance indicator. A
possible explanation for these ‘inconsistent’ findings is that TV viewing
could have triggered the young children’s curiosity about the outside
world, and accordingly generated more questions. Furthermore, as soon as
the young children had learnt more about the world through watching
TV, they would ask more questions with more pragmatic functions. And
further studies will be done to understand which interrogative functions
are directly affected by TV viewing time. Questions, however, do not
necessarily correlate with other language performance indicators. For
instance, we did not find any correlations between MLU and interrogative
development in our previous study on Cantonese-speaking children (Li
et al., 2013). Therefore, we would not regard this finding as inconsistent
with others; instead, we tend to conclude that TV viewing has both
positive and negative effects on various domains of early language
development. Thus, this study tends to lend empirical support to a
balanced view of TV viewing effects.
Second, the present study found that the more the parent–child conversation time, the fewer the interrogative types produced by the child. In early childhood years, parents are the primary ‘important others’ of children, and family is the most important environment. In parent–child shared book reading and communication, children respond to the story or adults with comments and questions, and parents maintain their interaction with children by proposing and answering questions. In this mutual process, children acquire the ability to make questions. In Chinese culture, however, parents do not converse on equal grounds with children so routinely, and the latter are not encouraged to question authoritative adults, such as parents and teachers (Li et al., 2013). Moreover, Chinese children are expected to follow their parents’ or teachers’ instructions to the letter. Therefore, the more the parent–child conversation, the fewer the questions raised by the child. Nevertheless, another possible interpretation might be that parent–child conversation could have satisfied children’s curiosity to a large extent, so the latter would have fewer questions. This assumption also needs further examination. This is the first corpus-based psycholinguistic study on the development of interrogative forms and functions and their predictors among Beijing children. Using a corpus as a primary source of evidence allowed us to establish the norms and to test the cross-linguistic generality of Choinard’s IRM hypothesis. However, a major limitation of this corpus-based research is that the data only represented a sample of the entire language, and not the target language in its totality (Tse & Li, 2011). Therefore, it should be borne in mind that this corpus might not contain all the information needed to explore this issue. To minimize this eventuality, the sample size needs to be increased, and data needs to be collected from more typical, everyday contexts. This might be precisely the primary limitation of this study. In addition, the sample was cross-sectional rather than longitudinal, and the age range examined should have extended downwards in order to elicit a more complete picture of the development of interrogatives in the very early years, such as from age one to age two. Nevertheless, the typology and pattern established in the present study contributes to furthering the understanding of the development of interrogative forms and functions in childhood Mandarin Chinese and to broadening the database of interrogative development in different languages. Furthermore, this study has provided some cross-linguistic evidence to support the IRM hypothesis and has cast new light on the importance of cross-cultural and cross-linguistic comparison of early child interrogatives. In this connection, this study has raised more questions than answers about early child interrogatives.
REFERENCES


Smith, M. E. (1933). The influence of age, sex, and situation on the frequency, form and function of questions asked by preschool children. Child Development 4, 201–3.


