Developmental Changes in Early Child Lexicon in Mandarin Chinese

Abstract

In this paper we report a large-scale developmental study of early productive vocabulary acquisition by 928 Chinese-speaking children aged between 12 and 30 months, using the Early Vocabulary Inventory for Mandarin Chinese (Hao, Shu, Xing, & Li, 2008). The results show that: (a) social words, especially words for people, are the predominate type of words in Chinese-speaking children's earliest productive vocabulary; (b) overall Chinese-speaking children's vocabulary contains greater proportions of nouns than other word categories, especially at the earliest vocabulary stage; and (c) verbs tend to appear earlier for Chinese-speaking children as compared with English-speaking children at the same levels of vocabulary development. In addition, our study has identified the underlying variables that influence the age of acquisition of words, specifically, the interplay between conceptual (imageability) and linguistic properties (word frequency, word length, and grammatical category) that jointly shape the development of Mandarin-speaking children's early vocabulary.

Keywords: early lexical development; age of acquisition (AoA) of lexical categories; nouns and verbs; Mandarin Chinese

Introduction

Uncovering universal principles versus language-specific patterns has been an important goal in language acquisition research. In the last several decades, researchers have been concerned with whether children learning different languages would have the same kinds of first words in their early productive vocabulary. A controversial issue in this regard is whether children display the same developmental profile across different languages or show different developmental patterns shaped by the specific properties of their target language. In this study, we address this issue by undertaking a large-scale study of Mandarin Chinese-speaking children, specifically to answer the following questions: (1) Does lexical development in early child Mandarin follow a universal pattern? In this regard we will examine the age of acquisition of lexical categories, with particular focus on the 'noun bias' versus 'verb bias' debate in the literature. (2) What underlying variables, including conceptual and linguistic properties of words, might impact the age at which lexical items are acquired in early child Mandarin? To address this question we will examine factors such as conceptual properties of words, word frequency, and word length that may modulate the observed empirical patterns in early lexical development.

Universal vs. language-specific pattern in the development of lexical categories

One well-studied domain with regard to universal vs. language-specific pattern in the development of lexical categories is the relative order of acquisition of nouns and verbs. A number of studies have indicated that the noun bias, a predominance of nominal items,

including common nouns and proper nouns, characterize children's early productive speech in many languages, and predicate items such as verbs come in quantity only afterwards (Gentner, 1982; Gentner & Boroditsky, 2001; Gillette, Gleitman, Gleitman, & Lederer, 1999; Gleitman & Gleitman, 1994; Golinkoff, Mervis, & Hirsh-Pasek, 1994; Markman, 1987; Waxman & Booth 2003). The predominance of nouns is a universal pattern, according to these researchers, because nouns are conceptually and perceptually more accessible than other lexical categories in language. Nouns referring to people and objects are also more stable across languages, while words representing actions or relations exhibit more variability and language-specific characteristics in the way in which actions and relations are encoded.

More recent studies, however, have challenged this view, with cross-linguistic data particularly from Asian languages. For example, Choi and Gopnik (1995) and Tardif (1996) have argued that the noun bias is weak or non-existent in languages such as Mandarin Chinese and Korean. Choi and Gopnik (1995) compared Korean-speaking and English-speaking children's productive vocabulary from the mean age of 14 months to 22 months, and found that Korean children produced equal proportions of nouns and verbs when their vocabulary reached about 50 words, while English-speaking children produced more nouns than verbs. Furthermore, Tardif (1996) and Tardif, Shatz, and Naigles (1997) have shown that Mandarin-speaking children produced more verbs than nouns, as compared with English-speaking children. These researchers argued against the universal pattern of noun bias and pointed out that there are language-specific differences not only in the encoding of nouns and verbs across languages, and also in the way early lexical items appear in child-directed parental speech. They proposed that some languages such as Chinese may be more

verb-friendly, in that ellipses of subjects or objects or both are allowed, making verbs more salient in the input language. Additionally, verbs in these languages may have higher imageability than nouns (Ma, Golinkoff, Hirsh-Pasek, McDonough, & Tardif, 2009), making them more accessible to children.

One line of evidence against a universal noun bias has also come from studies on the very earliest stages of lexical development, when children have acquired less than 50 words. Bornstein and Cote (2004) compared lexical development in children from seven different languages and found that children showed noun advantage across vocabulary levels only after 50 words. Bates, Marchman, Thal, Fenson, Dale, Reznick, Reilly, and Hartung (1994) had similar findings. They showed that routines or social words, which are not necessarily always nouns (e.g., *bye-bye*), appear earlier than common nouns in children's first words (Bloom, Tinker, & Margulis, 1993; Caselli, Bates, Casadio, Fenson, Fenson, Sanderl et al., 1995). For example, Caselli et al. (1995) analyzed the vocabulary composition of English- and Italian-speaking children aged from 8 to 16 months using the MacArthur-Bates Communicative Developmental Inventory (CDI) (Fenson, Dale, Reznick, Thal, Bates, Hartung et al., 1993). The results showed that the proportion of social words combined with words for people, onomatopoeic and routines was higher (over 60%) than that of common nouns (roughly 29%) within the first 50 words, regardless of languages. Kauschke and Hofimeister (2002) also found that although nouns were among the earliest vocabulary items in German-speaking children, the percentage of nouns was rather small (5.83%) and limited to persons and basic-level objects. Finally, Tardif, Fletcher, Liang, Zhang, Kaciroti and Marchman (2008) compared the vocabulary composition of children speaking English,

Mandarin Chinese, and Cantonese, using the CDI and its Mandarin and Cantonese versions. Their results showed striking similarities in children's first 10 words, which were predominantly words for people in all three languages. Sound effect words were the second largest category children could produce, though it was quite small in Mandarin (8.7%) compared to person terms (77.7%). The major differences existed in the relative percentage of common nouns and verbs: English-speaking children produced more common nouns (19.4%) than verbs (0.7%), while Mandarin-speaking children produced more verbs (7%) than common nouns (3.2%), and Cantonese-speaking children produced roughly equal numbers of verbs (4.8%) and nouns (5.7%).

Although the above studies all reported that the earliest acquired words were not necessarily object names, the specific categories reported as the earliest vocabulary were different. In Kauschke and Hofimeister's (2002) study, relational words (e.g., *up, again*) and personal-social words (e.g., *hi, no*) were the dominant categories (over 80%) within the earliest sampling points (13 and 15 months). In the studies of Caselli et al. (1995) and Tardif et al. (2008), words referring to people and sound effects were the two largest categories for English-speaking children, and in Tardif et al. (2008) words for people were the largest word category for Mandarin-speaking children. Other than these discrepancies, different criteria were used to count the word categories in different studies; for example, nouns not only included concrete nouns and abstract nouns, but also words for people in Kauschke and Hofimeister (2002), and words referring to outside places were excluded from common nouns in Caselli et al. (1995), but were included in Tardif et al. (2008).

Additional mixed findings were also obtained from several studies that compared Asian

languages with Western languages. For example, although Choi and Gopnik (1995) found equal proportion of nouns and verbs in early child Korean lexicon, Kim, McGregor and Thompson (2000) failed to replicate these finding using similar data collection method - a combination of maternal diaries and checklists. Kim et al. found that both English-speaking and Korean-speaking children produced more nouns than verbs when their vocabulary was close to the 50-word mark, and children from both languages acquired roughly the same number of nouns, except that the Korean-speaking children produced significantly more verbs than their English-speaking peers. Au, Dapretto and Song (1994) also reported that nouns outnumbered verbs in Korean-speaking children's early vocabulary, and there was no significant difference between these two groups of children in their study. Tardif and colleagues reported that Chinese-speaking children produced more verbs than nouns when measured by spontaneous speech in their earlier studies (Tardif, 1996; Tardif, Shatz, & Naigles, 1997), but in their later studies, nouns and verbs were found to be more evenly distributed in Chinese-speaking children earliest vocabulary (Tardif, Gelman, & Xu, 1999). A clear noun bias also occurred after children acquired more than 20 words when measured by an adapted Mandarin CDI, although this bias was not as strong as that seen with English-speaking children at comparable age or vocabulary level (Tardif, 2006; Tardif, Gelman, & Xu, 1999). Finally, Liu, Zhao, and Li (2008) conducted a corpus-based study comparing English, Mandarin, and Cantonese corpora in the CHILDES database (MacWhinney, 2000), and their analyses showed an even distribution of nouns and verbs in Mandarin and Cantonese but with increasing diversity and complexity as a function of age in all three languages examined.

In sum, while a significant amount of research has been devoted to the study of early lexical development in different languages, inconsistencies and controversies with regard to detailed developmental profiles remain, especially concerning what the predominant lexical categories are in children's early vocabulary, how vocabulary composition changes over time, and how differences might occur as a function of language-specific or culture-specific properties. In this study, we aim at providing some insights into these questions by examining Mandarin-speaking children's productive vocabulary across the 12-30 months range with a larger sample size (928 children), and on a more fine-grained level using our revised Early Vocabulary Inventory for Mandarin Chinese (Hao, Shu, Xing, & Li, 2008). This parental report method was modeled closely after the original CDI that has been proven a valid and powerful tool to assess children's early vocabulary (Dale, 1991; Dale, Bates, Reznick, & Morisset, 1989). A more comprehensive knowledge is needed for understanding Mandarin-speaking children's vocabulary acquisition compared to that for English-speaking children. To that end, our study emulates the study of Bates et al. (1994) in an attempt to chart a detailed developmental profile of vocabulary growth for different lexical categories at varying vocabulary levels. Given our literature review above, to address the first issue posed at the beginning regarding whether Chinese-speaking children follow a universal pattern in lexical development, we have to further ask three specific questions: (a) What lexical category or categories will first appear in Mandarin-speaking children's early vocabulary? (b) Do nouns dominate the vocabulary, and if so at which specific vocabulary stage of development? (c) How do lexical categories in Chinese-speaking children's productive vocabulary compare with those in the productive vocabulary of English-speaking children at different

developmental stages?

Variables underlying early lexical development

Another goal of this study, posed as the second question at the beginning, is to identify the underlying variables that modulate the age of acquisition (AoA) of early vocabulary, in addition to our efforts to describe the developmental trajectories. In the extant literature, two major types of factors have been suggested to play significant roles: conceptual variables such as the imageability of object-referring nouns as compared with action-referring verbs, and linguistic, input variables such as the frequency and length of words in the child's early vocabulary.

Early studies advocating the noun bias have explored the conceptual variables in particular, and suggested that nouns are acquired early because they are usually conceptually and perceptually more accessible than verbs (Gentner, 1982; Gentner & Boroditsky, 2001; Gillette et al., 1999; Gleitman & Gleitman, 1994). More recent studies have further attempted to demonstrate the role of specific conceptual variables underlying vocabulary development. In particular, the SICI theory (Maguire, Hirsh-Pasek, & Golinkoff, 2006) highlights the important roles of four variables of objects in the child's learning environment: Shape, Individuation, Concreteness and Imageability (SICI). 'Shape' refers to a persistent, tangible object contour for an object and overall configuration of the action. 'Individuation' refers to the ease with which the referent can be distinguished from the surrounding of the environment. For example, the referent of the noun *cat* can be easily observed in the world, while the referent of the noun *idea* cannot be observed. 'Concreteness' refers to the degree to which the

object encoded by the word is manipulable (e.g., the learner can see, hear, and touch the object), while 'imageability' refers to the ease with which a word can arouse a mental image.

According to the SICI theory, the above four factors commonly determine the ease with which a novel word is learned, regardless of grammatical category. These factors are distributed on a continuum spanning word classes, and they also allow for overlap between different grammatical classes. Common nouns label objects that are located at the more concrete, easily individuated end of the continuum with higher imageability and consistent shapes, whereas verbs fall at the less concrete, difficult to individuate end with lower imageability and variable shape. Such differences between common nouns and verbs may explain why nouns tend to be learned earlier than verbs in general. Thus, rather than just looking at the relative order of acquisition of nouns and verbs as a whole, the SICI theory examines the differences of nominal and verb categories along a continuum.

It should be noted that although the SICI theory highlights four different variables as the acronym suggests, the proponents of the theory did not intend the variables to be taken at their face value (see Maguire et al. 2006, p. 375). Rather, the authors suggest that SICI is intended to include many factors that scale the difficulty of learning a particular word. In addition, the literature so far is unclear about the relative or unique role that each of the four SICI variables plays in determining the AoA of early vocabulary. This is perhaps partly due to the fact that these four variables are highly correlated with each other, and it is not easy to isolate the unique contribution of each. To make the matter worse, the literature has also used some of the terms loosely or interchangeably (e.g., concreteness has often been used interchangeably with imageability). Given this situation, we have not designed the current study as a test of the

SICI theory, but rather, use the SICI theory as a starting point for thinking about the role of conceptual properties of words, as opposed to linguistic properties (e.g., frequency and length, discussed below). Nevertheless, we highlight and test one specific variable in the SICI theory, imageability, on the basis of the empirical evidence that this variable is a reliable predictor of AoA of lexical categories.

A number of studies have reported that imageability significantly impacts the acquisition and processing of nouns and verbs not only in adult rating (Bird, Franklin, & Howard, 2001), but also in children based on AoA scores obtained from parental reports such as CDI (Ma et al., 2009; McDonough, Song, Hirsh-Pasek, Golinkoff, & Lannon, 2011). For example, McDonough et al. (2011) selected 120 words (76 nouns and 44 verbs) that had both published AoA data and rated imageability data. Their hierarchical multiple regression analysis with AoA as an independent variable showed that both word grammatical category (noun or verb) and imageability had unique contributions, but that imageability was more powerful. Ma et al. (2009) asked adults to rate the imageability of nouns and verbs that appear in the Chinese and English CDIs. The results showed that Chinese verbs' imageability value was higher than that of English verbs. Compared with English verbs, Chinese verbs that are acquired earlier by children had the properties of semantic specificity and higher imageability. Thus, imageability has been found to be related to the verb advantage for Chinese children.

In addition to the conceptual properties of words such as imageability, the current study also examines non-conceptual, linguistic factors that are relevant to early lexical development, such as word length, and word frequency. Zhao and Li's (2008) computational simulations of vocabulary development in English and Chinese showed that the mean length of words (in

phonemes) gradually increased as a function of vocabulary size in both languages, and the average phonemic length of Chinese verbs is not only shorter than that of Chinese nouns, but also shorter than that of English verbs. This word length advantage may make Chinese verbs easier to acquire than nouns for Chinese learning children. Based on the findings of Zhao and Li (2008), we would expect that word length would also play a unique role for the AoA of words.

Input frequency may also contribute to the noun and verb disparity. Goodman, Dale and Li (2008) investigated the relationship between input frequency measured by child-directed speech and AoA of the CDI vocabulary in English. They found that within each grammatical category (nouns, verbs, adjectives, and closed-class words), the more frequent the word is in parental speech, the earlier the child can produce the word. Interestingly, the correlation between input frequency and AoA disappeared when words of all grammatical categories were calculated together (i.e., it plays a less important role across categories). Furthermore, McDonough et al. (2011) found a unique role of input frequency on the AoA of the English CDI vocabulary when common nouns and verbs were included, and Ma et al. (2009) similarly identified the impact of input frequency on AoA for both Chinese and English CDI words. Finally, in computational models Li, Zhao, and MacWhinney (2007) showed that both input frequency and word length significantly modulate the lexical growth patterns for early English vocabulary, especially with regard to the shape of vocabulary spurt. To what extent input frequency plays a unique role in the AoA of early child Mandarin lexicon remains an open issue, one that we will examine in the current study alongside the effects of other conceptual and linguistic variables.

In what follows, we first discuss the methodology of the research and analyze the developmental trajectory of the early child Mandarin lexicon, in order to address the first question regarding universal vs. language-specific patterns raised at the beginning. We then test the relative influences of the afore-mentioned conceptual and linguistic variables on the AoA of early child Mandarin vocabulary, in order to address the second question regarding the underlying variables that govern developmental patterns.

Methods

Participants

Two groups of participants were recruited for the purpose of this study: one group consisting of parents and their children who provided the AoA data of the early vocabulary for our study, and another group consisting of college students who provided rating data on imageability of words.

Parents and children: 928 Mandarin-speaking children aged between 12 and 30 months were recruited from Beijing 301 Hospital, Beijing Haidian Hospital, and Tieying Hospital, where they received regular physical checkups. Each age group had at least 40 children (range: 40-67). Parents (71% are mothers on average) were asked to indicate whether their child speaks the word on the designated questionnaire (the checklist, see below) depending on the age of their children and to provide basic demographic information of their educational level, occupation and income (see Table 1 for more details about the children and their family information for each age group). Most families (about 82%) belong to the middle class as assessed by our questionnaire based on annual family income. Over 80% mothers in this

study had graduated from colleges or vocational schools and were white-collar workers, such as editors, teachers, accountants, engineers, lawyers, salesmen, or technicians in factories etc. A trained nurse in the hospital helped to collect the questionnaires. Among the 928 parents, 278 of them had children aged between 12 to 16 months and completed the Infant Checklist, and 650 had children aged between 17 to 30 months and completed the Toddler Checklist (for more information see Hao et al., 2008). A child was not counted in Table 1 and the child's data were excluded from our analysis if the child had any the following conditions: premature birth, exposure to a second language, or insufficient birth information. Parents were provided written consent to the data collection procedure, understanding that they were free to withdraw from the process at anytime during the completion of the inventory forms.

College students: 40 undergraduates from Beijing Language and Culture University participated in the rating study. They were all native speakers of Mandarin Chinese. They were asked to rate words from the Toddler Checklist with respect to Imageability of words (see details below). They received token gifts for their participation.

[Insert Table 1 about here]

Materials and procedure

The instrument used in the present study was the Early Vocabulary Inventory for Mandarin Chinese, a checklist that followed the format of the original CDI for English (see Hao et al., 2008 for a complete list of the words from the two checklists, downloadable at http://blclab.org/early-vocabulary-inventory-for-mandarin-chinese/). The construction of the checklist considered problems associated with previous checklists for Mandarin Chinese and

reflected the language-specific properties in capturing early vocabulary development. A number of steps were taken to ensure the high validity of the checklist for Chinese children. First, among all the possible Chinese equivalents of the lexical items that appeared in the English CDI, we excluded items such as *lawn mower*, which are uncommon in Chinese daily life, but included items that are more culturally frequent, such as *xiong2mao1* 'panda' and *xi1fan4* 'porridge'. Second, we also asked nine mothers whose children aged between 17 and 30 months to revise and refine the checklist based on their experience with their own children (see Hao et al., 2008, for a more detailed description of the construction of the instrument).

The final checklist includes two parts, following the example of the original English CDI. Part 1 is the Infant Checklist, used for infants 12 to 16 months of age, which contains 232 words. It consists of the following 14 categories: *action words, animals, body parts, clothing, descriptive words, food and drink, furniture and rooms, games and routines, outside things and places to go, people, pronouns, small household items, toys,* and *vehicles*. The Infant Checklist has 160 items (69%) in common with the English CDI Infant Checklist; these items are direct translation equivalents in the two languages. Part 2 is the Toddler Checklist, used for children 17 to 30 months of age, which contains 710 words and thus a more comprehensive list of words than the Infant Checklist. It consists of the following 20 categories: *action words, animals, body parts, clothing, connecting words, descriptive words, food and drink, furniture and rooms, games and routines, helping verbs, outside things, people, places to go, pronouns, quantifiers and articles, question words, small household items, toys, vehicles,* and *words about time.* The Toddler Checklist has 407 items (57%) in common with the English CDI Toddler Checklist.

The 710 words from the Early Vocabulary Inventory for Mandarin Chinese were split into two sets of 355 words for student ratings. Each college student rated one set of words on imageability on a 7-point scale. Imageability was rated with regard to how easily each word elicited a mental image, where 1 = least imageable, such as *idea* and 7 = most imageable, such as *apple* (see also Barca, Burani, & Arduino, 2002; Liu, Shu & Li, 2007).

To identify the role of input on AoA of words, we used word frequency estimates from the SUBTLEX-CH of Cai and Brysbaert (2010), a Chinese word frequency database based on film subtitles, as approximate estimates of children's language input. The word frequency norms based on the subtitles of movies and TV series were recently developed in several languages including English, Dutch, German, Spanish, Greek, and Chinese. In several studies they were shown to consistently outperform the traditional norms based on written texts like books, newspapers or internet pages during word recognition tasks, and were thus considered to provide good estimates of daily language exposure (Brysbaert, Buchmeier, Conrad, Jacobs, Bölte, & Böhl; 2011; Cai & Brysbaert, 2010; Cuetos, Glez-Nosti, Barbon, & Brysbaert, 2011; Keuleers, Brysbaert, & New, 2010; New, Brysbaert, Veronis & Pallier, 2007).

Data analyses

Two steps of data analysis procedures were applied: First, the percentage of each category of vocabulary was reported to describe the trajectories of vocabulary development for children, as has been done in previous studies (Bates et al., 1994; Tardif, 1996). Second, in order to explore which factors could be the predictors of the AoA for a specific word, a set of other analyses were conducted: (1) we calculated the Pearson correlations between imageability, word frequency, word length, grammatical category membership (noun or verb) ,

and the AoA of words; and (2) we ran multiple regression analyses treating AoA as the dependent variable, and the imageability of words, word frequency, word length, and word grammatical category as independent variables.

Because of our focus on nouns and verbs in this study, for the word grammatical category variable we only included the noun-verb contrast. This decision was also based on the consideration that nouns and verbs made up the majority of the early vocabulary, while other lexical categories (e.g., adjectives and closed-class words) were significantly fewer in number. The word length variable here is mainly focused on the difference between monosyllables vs. multi-syllables. Chinese words consist of one or more syllables, each of which is represented by a character in written script. Among the 450 words we analyzed in the second part, 136 were monosyllables, 276 disyllables, and only 38 tri-syllables.

When calculating correlations and conducting regression analyses, the averaged value of imageability over raters was calculated for each word and entered for analyses, the word frequency data was log10 transformed, and the variables of word length and word grammatical category were dummy coded (word length: 1=monosyllabic, 0=multisyllabic; grammatical category: 1=noun, 0=verb).

The final words used in the second part of our data analyses (i.e., correlation analysis and multiple regression analysis) only included 348 common nouns and 102 action verbs (i.e., 450 out of the 710 words). The other 260 words were not included in this part of the analyses because they were: (1) words (64 items; 9.01%) that are not acquired until 30 months old by the 50% criteria of Goodman et al. (2008), for example, *cha1zuo4* ‘plug socket’; (2) other types of words like social words (57 items; 8.02%), such as *ye2ye* ‘grandpa’, descriptive

words (69 items; 9.72%), such as *kong1de* 'empty', closed-class words (38 items; 5.35%) and others (10 items; 1.4%); and (3) words (5 common nouns and 17 action verbs; 3.1%) that cannot be found in SUBTLEX-CH database, for example, *gong1gong4qi4che1* 'bus'.

Results

Developmental changes in vocabulary composition among Chinese infants (12-16 months)

Bates et al. (1994) argued that it is important to examine the developmental changes in vocabulary composition by vocabulary level, rather than by chronological age (see also Pine & Lieven, 1990). Following their reasoning, we divided the infant participants into the following six groups based on their total production vocabulary: (1) 0 words (N = 30, 10.79% of the participant sample); (2) 1-5 words (N = 103, 37.05%); (3) 6-10 words (N = 58, 20.86%); (4) 11-20 words (N = 23, 8.27%); (5) 21-50 words (N = 44, 15.83%); (6) more than 50 words (N = 20, 7.19%). We also followed Bates et al. (1994) to define the various lexical categories: (1) *Common nouns* included items from the following 10 categories: animal names, body parts, clothing, food and drink, furniture, small household items, toys, vehicles, places to go except *cheng2li, hai3bian1* (which are locative nouns) outside things except *shang4ban1, tian1shang, wai4mian* (which are not common nouns). In total, there were 125 common nouns, making up 53.88% of the Infant Checklist; (2) *Predicates* included items from the verb category and the adjectives category. There were 65 predicates, making up 28.02% of the Infant Checklist; (3) *Closed-class items* included only five pronouns (5 items, 2.2% of the Infant Checklist); and (4) *Social words* included items referring to people and to routines, making up 14.66% of the Infant Checklist (34 items), which have been reported in the

literature as the earliest acquired words by children in English and other languages (Bloom et al., 1993; Caselli et al., 1995; Caselli, Casadio, & Bates, 1999; Kauschke & Hofimeister, 2002).

Figure 1(a) presents percent changes of the above four major categories as a function of the various vocabulary levels for Chinese infants. In order to clearly reveal the similarities and differences between Chinese-speaking children and English-speaking children, we present the data from Table 2 in Caselli et al. (1995) as Figure 1(b), which has comparable data points as Figure 1(a).

[Insert Figures 1(a)-(b) about here]

Figures 1(a) indicate that for young Chinese-speaking infants at the earliest stages of vocabulary development, social words accounted for 80% of their total production vocabulary, while common nouns represented only 10.56%. When the production vocabulary expanded to a mean of 7.33 words (i.e., between 6 to 10 words), social words still predominated the vocabulary but common nouns increased slightly from 10.56% to 16.24%. Common nouns increased dramatically when children acquire an average of 15 words (between 11 and 20), while the use of social words rapidly decreased. From there on, common nouns increased steadily to a peak of 48.52% when children acquired more than 50 words. A comparison of Figure 1(a) and Figure 1(b) shows that the complementary or contrasting developmental patterns for social words versus common nouns are strikingly similar for Chinese-speaking children and English-speaking children, in that increases in common nouns are accompanied

by corresponding decreases in the proportion of social words.

It is clear from Figure 1 that common nouns and social words are the two largest categories, and there were very few closed-class items, for both English-speaking and Chinese-speaking children. However, while there was virtually no difference between predicates and closed-class items in terms of the number of words produced by English-speaking children, as reported by Bates et al. (1994) and Caselli et al. (1995), predicates appeared as early as common nouns did in Mandarin Chinese, even though the speed of increase did not match that of the common nouns within the first 50-word mark. This major difference between Chinese and English provides some support to the claim that verbs tend to appear earlier for Chinese-speaking children than for English-speaking children (Tardif et al., 1999; Tardif, 1996, 2006). Still, within the first 50 words, the percentage of common nouns exceeds that of the predicates in Chinese, and this was verified by paired-samples t-tests applied to the vocabulary levels between common nouns and predicates (all ps < .05).

Developmental changes in vocabulary composition among Chinese toddlers (17-30 months)

In this section, we further focus on the developmental trajectory of common nouns, social words, predicates, and closed-class words for older Chinese-speaking toddlers (17-30 months). Figure 2 presents the percentages of common nouns, social words, predicates, and closed-class words at each vocabulary level. Given that social words are not the core vocabulary for toddlers and the proportion of social words decreases dramatically across vocabulary level as shown in Figure 2, we will not further highlight the social words in this section.

Similar to the analyses of the infant data, the toddler participants were divided into the following eight groups based on their total production vocabulary: (1) 0-50 words (N = 70, 10.77% of the participant sample); (2) 51-100 words (N = 72, 11.08%); (3) 101-200 words (N = 103, 15.85%); (4) 201-300 words (N = 113, 17.38%); (5) 301-400 words (N = 80, 12.31%); (6) 401-500 words (N = 46, 7.08%); (7) 501-600 words (N = 69, 10.62%); (8) 601-710 words (N = 97, 14.92%). We used the same criteria in defining the lexical categories and in calculating the percentages of each category at each vocabulary level. Among the 710 words in the Toddler Checklist, there were 382 common nouns (53.8% of the checklist), 199 predicates (28%), 49 closed-class words (6.9%) from the following five categories: connecting words, helping verbs, pronouns, quantifiers and articles, and question words, 62 social words (8.7%) and 18 others (2.5%).

[Insert Figure 2 about here]

Figure 2 shows that common nouns have complete dominance over other categories across all vocabulary levels beyond the 50-word mark. Within the first 50 words, the percentage of common nouns was similar to that of social words (37.11% and 32.43% respectively), but it rapidly increased at the 51-100 word stage. Consistent with data from English, as reported in Bates et al. (1994), common nouns made up more than 50% of the total vocabulary for Chinese-speaking children with vocabularies between 101 and 200 words, and continued to increase slowly but then leveled off.

In contrast to common nouns, the proportion of predicates stayed roughly around 30%

for the Chinese toddlers, without significant change across vocabulary levels. This percentage, however, is higher compared with children learning other languages at the same developmental vocabulary levels. For example, Caselli et al. (1999) observed a developmental pattern that the proportion of predicates increased from below 10% when vocabulary size was less than 50 to about 25% when vocabulary size reached 400-500 for both English and Italian children (aged from 1;6 to 2;6). Behrens (1998) used a longitudinal approach to study the production of verbs in ten children who acquired German, English and Dutch, and found that many verbs did not appear in the production vocabulary until the children entered their third and fourth year of life.

With respect to the percentage of closed-class words, Chinese toddlers did not show substantial increases during this period. This pattern indicates that closed-class words are difficult to learn. Bates et al. (1994) found English-speaking children's closed-class words do not show substantial development until they have acquired more than 400 words. In an input analysis Goodman et al. (2008) also showed that the input frequency of closed-class words, though generally high, has the least influence on the order of acquisition in vocabulary development, as compared with frequency effects of other categories such as nouns and verbs. One longitudinal study by Kong, Hu, Ou'yang, Chen, Ding, Wang et al. (2004) also provided some evidence of the late acquisition of closed-class words by Chinese-speaking children. They found quantifiers and articles would develop rapidly only after 4 years of age, and pronouns except *wo3* 'I' also appeared late.

First 50 words: A detailed analysis

According to Caselli et al. (1995), 'first acquired' is defined in terms of the percentage of children in the sample who are reported to produce the word. Table 2 presents the first 50 words in the production vocabulary of Mandarin-speaking infants by their decreasing order of appearance.

[Insert Table 2 about here]

Table 2 shows that the first five words produced by Chinese-speaking children are all kinship terms. Consistent with the results of English-speaking and Italian-speaking children (Caselli et al., 1995; Gentner, 1982; Nelson, 1973), Chinese-speaking children's two earliest acquired words were "daddy" (91.8%) and "mommy" (90.7%). The next two kinship terms were "grandma" (47.6%) and "grandpa" (35.2%). We attributed the results of these early kinship terms to the characteristics of the Chinese culture, where the immediate family members in the household often include both parents and grandparents. Before children can say anything, they are encouraged to address these members at homes, especially *ba4ba* 'daddy', *ma1ma* 'mommy', *ye2ye* 'grandpa' and *nai3nai* 'grandma'. The fifth earliest item was *bao3bao* 'baby', used to refer to the infant itself.

The semantic category of the next five words expanded from solely kinship terms to common nouns and words for routines. There were still two kinship terms among this set: *a1yi2* 'auntie', which can refer to mother's sister and any female similar to the child's mother in age, and *ge1ge* 'elder brother', for a male child older than the child. The two common nouns were words for animals, *mao1* 'cat' and *gou3* 'dog' and one routine word is *bu4* 'no'.

These results are comparable to Italian-speaking children's earliest vocabulary composition (Caselli et al., 1995) and to the findings of Tardif et al. (2008). While Italian-speaking children acquire an average of 4 kinship terms in the earliest 10 words, Chinese-speaking children use 7 kinship items. As we discussed earlier, terms for people and routine were combined to form a large category - social words, and this is clearly reflected in the first 10 words used by Chinese children. Kauschke and Hofmeister (2002) reported that personal-social words occupied 25.19% of German-speaking children's earliest vocabulary. D'Odorico, Carubbi, Salerni and Calvo (2001) found that half of Italian-speaking children's first 50 words are onomatopoeic words, routines or names of people (see also Caselli et al., 1995). These findings demonstrate that the importance of social words, appearing as the first category in children' expressive vocabulary regardless of the language the child is learning.

Consistent with the findings of Caselli et al. (1995) for English-speaking and Italian-speaking children, verbs did not appear in the first ten words produced by Chinese-speaking children. In Caselli et al. (1995), English-speaking children failed to produce any verb among the first 50 words, and then produced 4 verbs in the second 50 words; Italian-speaking children produced only one verb *DARE* 'to give' among the first 50 words and added 4 more verbs when the vocabulary expanded to 100. In our study, Chinese-speaking children produced a total of six verbs among the first 50 words, *bao4* 'hug', *da3* 'hit', *na2* 'bring' or 'take', *yao4* 'want', *chi1* 'eat', and *zou3* 'walk' or 'go'. In the next 50 words, predicates (verbs and descriptive words) increased gradually: children produced 10 verbs and 10 adjectives among the second 50 words. In total, predicates constituted 29% of the first 100 words in child Mandarin, a quite large number compared to the percentages of

predicates for English and Italian-speaking children.

We now look at the 20 common nouns in the first 50 words. Among the first acquired common nouns, six were names of animals, such as *mao1* 'cat', *gou3* 'dog', *niao3* 'bird'; four were words for food and drink, such as *fan4* 'food' and *rou4* 'meat'; three were words for outside things, such as *hua1* 'flower' and *shu4* 'tree'; two were words for toy, one for body part, one for clothing and one for household items. These words bear some common characteristics: (1) all of them are monosyllabic words, except *wa2wa* 'doll', with a syllabic reduplication; (2) most of them are for basic-level concepts, for example, *yu2* 'fish', (3) most of them are related to early life experience of the child, about objects or items in the child's immediate environment. These data are consistent with the findings of Gentner (1982) and Nelson (1973), in that the earliest common nouns are those for basic-level or socially relevant concepts.

Comparing the individual items first acquired by Chinese-speaking and English-speaking children, we can find both similarities and differences. As can be seen from Table 2, there are 19 items in total among the earliest acquired vocabulary by children from both languages: five items for people (*grandma* has two equivalents in Chinese, *nai3nai* and *lao3lao*, and *grandpa* also has two equivalents, *ye2ye* and *lao3ye*), five items for games and routines, one descriptive word, and the remaining items common nouns. However, in English-speaking children's first 50 words, about 50% can be regarded as social words, and more than 40% are common nouns, whereas in Chinese-speaking children's first 50 words, the percentages of social words and common nouns were similar (roughly around 40%), while the remaining 20% of words were mainly predicates (i.e., more predicates in child Mandarin than in child

English).

The predictors of age of acquisition of early vocabulary

To understand the mechanisms underlying the similarities and differences between children learning Mandarin and other languages, we describe in this section our analyses of conceptual (imageability) and linguistic (frequency, length, and grammatical category) variables as predictors of the age of acquisition (AoA) of early vocabulary. The imageability rating data from college students and other variables on common nouns and action verbs are presented in Table 3. On average, the rating results for nouns on imageability were higher than for verbs ($F(1, 448) = 268.26$, $p < .001$), which indicates that the referents to which the nouns correspond can readily generate a mental image. These rating results are consistent with the predictions of the SICI theory (Maguire et al., 2006). On the other hand, word frequency was higher for Chinese verbs than for Chinese nouns ($F(1, 448) = 128.82$, $p < .001$), which is consistent with some recent statistical analyses that show that Chinese-speaking children's language input may contain more verbs of higher frequency (Sandhofer, Smith, & Luo, 2000; Zhao & Li, 2008). In addition, these verbs were found to be significantly shorter in length than the nouns ($F(1, 448) = 89.54$, $p < .001$).

[Insert Table 3 here]

Pearson correlations between imageability, word frequency, word length, grammatical category membership, and the AoA of words are presented in Table 4. Both word frequency

and word length were negatively correlated with the words' AoA, r = -.451, p< .001 (word frequency with AoA) and r = -.428, p < .001 (word length with AoA). This makes sense in that words that are shorter and higher in frequency tend to be acquired earlier. These two variables seem to have the most impact on AoA of child Mandarin lexicon (see regression analysis below). They also showed significant correlations with the other two variables (imageability and grammatical category) (all ps < .001). Imageability also significantly correlated with AoA, r = -.099, p < .05. However, grammatical category membership did not correlate with AoA, r = .023, p > .1.

[Insert Table 4 here]

To explore the relative contributions of imageability, word frequency, word length (monosyllabic or multisyllabic word), and grammatical category (noun or verb) to the AoA of vocabulary, we ran multiple regression analyses with the AoA of words as the dependent variable and the four variables as predictors. The results of the final regression analysis were displayed in Table 5. Together the four predicting variables accounted for 37.2% of the AoA variance (F(4,445) = 67.38, p < .001). Their unique contributions were evaluated through the interpretations of squared semi-partial coefficients (sr^2) (Cohen, Cohen, West, & Aiken, 2003). Specifically, word frequency uniquely accounted for 14.1% of the AoA variance, word length 6.7%, imageability 5.2%, and grammatical category 1.3% (see Table 5). These data suggested that word frequency, word length, and imageability played predominant roles on predicting the AoA of words, with grammatical category playing a relatively smaller role.

[Insert Table 5 here]

[Insert Figure 3 here]

Finally, based on the word length difference between nouns and verbs we observed and on similar findings by Zhao and Li (2008), we further compared the percentages of monosyllabic and multisyllabic words in nouns and verbs at different vocabulary levels. The data were presented in Figure 3. Two observations should be noted here. First, the percentage of multisyllabic words increases rapidly as the vocabulary grows. Second, different developmental patterns exist for nouns versus verbs with regard to this word length increase. While there were in general more monosyllabic than multisyllabic verbs, there were more multisyllabic nouns once children have reached the 100 to 200 vocabulary mark. The contrasting patterns of nouns and verbs indicate that syllable length may play differential roles on the development of different grammatical categories (see General Discussion below).

General Discussion

This study had two important goals. First, we wanted to chart the developmental trajectory of early vocabulary acquisition in Mandarin Chinese based on a large-scale parental report study, as has been previously done in English by Bates and colleagues (Bates et al., 1994; Caselli et al., 1995). In this respect we wanted to see if Mandarin learning children follow a universal pattern of lexical development as children learning other languages, or, if they show

language-specific patterns that differ from those in English and other European languages. Second, we wanted to understand the important contributing factors that are responsible for the age of acquisition of lexical categories across the early stages of lexical development. With respect to the first goal, our results show that (a) social words, especially words for people, are the predominate type of words in Chinese-speaking children's earliest productive vocabulary; (b) overall Chinese-speaking children's vocabulary contains greater proportions of nouns than other word categories, especially at the earliest vocabulary stage; and (c) verbs tend to appear earlier for Chinese-speaking children as compared with English-speaking children at the same levels of vocabulary development. These results address the three specific questions we raised in the Introduction with regard to the developmental trajectories of child Mandarin lexicon. With respect to the second goal, the correlation and multiple regression analyses show that it is the interplay between conceptual (imageability) and linguistic variables (word frequency, word length, and grammatical category) that jointly determine the AoA of Mandarin-speaking children's early vocabulary.

First words and kinship terms in early child Mandarin

The data in the present study provide the first clear evidence that social words predominate Chinese-speaking children's earliest vocabulary, and the majority of social words in our study are words for people. This statement is verified by two findings: (1) as illustrated in Figure 1(a), words for people occupy more than 60% of the vocabulary when the lexicon contains less than 10 words (i.e., word types), and (2) there are seven words representing kinship terms among the ten earliest acquired words, as Table 2 shows. These

findings are consistent with the results from children speaking German (Kauschke & Hofimeister, 2002), Italian (D'Odorico et al., 2001; Caselli et al., 1995), and English (Caselli et al., 1995). Our results, together with the findings from other studies suggest that social words make up the first category in children' expressive vocabulary regardless of language.

The predominance of kinship terms in early child Mandarin vocabulary may be specially related to the richness of kinship terms in the Chinese language and the emphasis on kinship relations in the Chinese family culture. In Chinese, close relatives have specific names depending on whether they are from the maternal or paternal side, their gender, and relative age (see also Tardif et al., 2008). A somewhat extreme example is the word *cousin* in English that translates to eight terms in Chinese: *tang2xiong1* 'son of father's brother, older than the speaker', *tang2di4* 'son of father's brother, younger than the speaker', *biao3ge1* 'son of mother's brother, older than the speaker', *biao3di4* 'son of mother's brother, younger than the speaker', *tang2jie3* 'daughter of father's brother, older than the speaker', *tang2mei4* 'daughter of father's brother, younger than the speaker', *biao3jie3* 'daughter of mother's brother, older than the speaker', *biao3mei4* 'daughter of mother's brother, younger than the speaker'. On the other hand, Chinese also has some kinship terms that serve more general purposes, instead of labeling direct kinship relationship. For example, *ye2ye* refers to grandpa literally, but also may be used to refer to any male with roughly the same age of as speaker' own grandpa. Unlike American mothers favoring object naming games (Bornstein, Haynes, & Painter, 1998; Fernald & Morikawa, 1993; Goldfield, 1993; Tamis-LeMonda, Bornstein, Cyphers, Toda, & Ogino, 1992; Tardif et al., 1997), Chinese families encourage children to practice name calling in appropriate ways with various kinship terms for addressing relatives and

acquaintances. Given these unique characteristics in the Chinese language and the culture, it is only natural that Chinese children's earliest vocabulary contains a large number of kinship terms. Tardif et al. (2008) also showed that there are eight to nine kinship terms among the top 20 words produced by Mandarin- and Cantonese-speaking children, while only two in comparable English-speaking children's samples.

This precedence and predominance of kinship words seem to contradict the predictions of some existent theories regarding the order of acquisition of words. The 'Division of Dominance' theory (Gentner & Boroditsky, 2001) and the 'SICI Continuum' theory (Maguire et al., 2006) are two hypotheses that make explicit predications on the acquisition sequence of different word categories. According to these hypotheses, relational nouns such as *uncle, grandmother* should be acquired after the proper nouns (e.g., *Mary*, *Sue*) and concrete nouns (e.g., *ball*, *spoon*), because the former denote the relations between people and is harder for children to acquire than the latter. However, both the present study and Tardif et al. (2008) find a large number of relational kinship terms in Mandarin- and Cantonese-speaking children's earliest vocabulary. Of course, one could argue that, despite the appearance of the words in children's early vocabulary, very young children may not understand the exact meanings of the kinship terms, and only later on they can start to appreciate the complex relations that hold among the different people as denoted by these terms. This is certainly a hypothesis that needs to be tested empirically in future research.

Developmental changes in vocabulary composition across stages

Overall our data and analyses show that the vocabulary development trajectory for

Chinese-speaking children fits the four-stage model of lexical development proposed by Caselli et al. (1999). At the very first stage of lexical development, especially when children's vocabulary is within 10-20 words, onomatopoeic terms, social routines, and names for favorite people are predominate. Some social routine words such as *bu4* 'no', *xie4xie* 'thanks', also appear in the earliest vocabulary.

The second stage, according to Caselli et al., is *Reference*, when children's vocabulary accumulates to 50 to 200 words. During this stage, in English and many other Indo-European languages, the vast majority of words are nominals, common nouns and proper nouns referring to objects, things, and people. Chinese-speaking children's vocabulary development is also characterized by the rapid growth of nominals. When the size of vocabulary goes beyond 20 words, common nouns quickly outnumber social words and become the largest category in children's lexicon.

The third stage is the so-called *Predication*, according to Bates and colleagues (Bates et al., 1994; Caselli et al., 1995; 1999), where verbs and adjectives show substantial increase after children have acquired 100 or more words. Here Chinese-speaking children demonstrate some different developmental patterns in the acquisition of predicates when compared with their English and Italian peers: (1) the percentage of predicates is higher throughout the early vocabulary in Mandarin-speaking children than in Italian- or English-speaking children, and (2) the percentage of predicates at each vocabulary level is constant, roughly around 30% out of the total vocabulary (see Figure 2), whereas the percentage of predicates increases from less than 10% up to 25% for Italian or English children. It has been argued that increases in predicates are closely related to the development of ability to understand relational meanings

(Caselli et al., 1999). The overall higher percentage of predicates in Chinese seems to contradict this argument, but upon detailed examination of the actual predicate words produced by Chinese-speaking children, we can see that the Chinese verbs are mostly words of higher imageability or concrete actions: for example, *fei1* 'fly', *ku1* 'cry', and *chi1* 'eat' when their total vocabulary is between 101 and 200. Only a small percentage of children can produce some verbs that encode relational meaning: for example, 21% of the children produced *zhui1* 'chase' and 17% of the children produced *tui1* 'push' at this time. The overall higher percentage of verbs in Chinese children's speech as compared with other languages is probably due to the higher imageability and shorter syllable lengths that characterize many Chinese verbs (see Liu, Zhao, & Li, 2008 and more discussion below).

Finally, the fourth stage is *Grammar* according to Caselli et al. (1999), at which time grammatical function words develop when the size of the vocabulary goes between 300 and 500 words. However, the Chinese-speaking children in the present study do not display a substantial growth in grammatical function words, even at the relatively late stages. Examining the closed-class words when children's overall vocabulary exceeds 400 words, we find that more than half of the children could not produce half of the closed-class words. Some highly frequent closed-class words in adult spoken language are rarely used by children: for example, only 10 percent of the children are reported to produce connecting words such as *ru2guo3* 'if', and only 30% of children are reported to produce helping verbs such as *bi4xu1* 'must'. Children at this stage also have trouble producing plural pronouns, such as *wo3men* 'we', *ni3men* 'you', *zhe4xie1* 'these', *na4xie1* 'those'. One speculation about the lack of progress with closed-class words in Chinese-speaking children is that the Chinese language

relies heavily on semantic-pragmatic-contextual cues (Li, 1998), and as such function words are less important especially in child-directed speech and hence less likely picked up by children during early stages of lexical development.

When describing the developmental changes in Chinese-speaking children's lexical composition, we should highlight that despite the language-specific differences between Chinese and other languages, *nouns are learned before verbs, even in Chinese*. We quantify this claim based on several considerations. First, Gentner's original noun bias proposal includes common nouns and proper nouns (Gentner, 1982). Our analysis above shows that although the vast majority of the words are social words when children acquire less than 20 words, a large number of these words are kinship terms in Chinese, which are often used by children as proper nouns to name particular listeners in the speech environment. Moreover, the percentage of common nouns quickly exceeds that of social words after 20 words. Second, after children have acquired 50 words, common nouns occupy almost half of the lexicon at each vocabulary level in child Mandarin. These results are consistent with the claims that nominals (including proper nouns and common nouns) are first acquired by children and that common nouns are the predominant category overall. Third, although the first verbs appear as early as common nouns do for Chinese-speaking children, the percentage of common nouns is significantly higher than that of verbs at every vocabulary level (see Figure 1a and 2). This pattern is consistent with several recent studies of Chinese-speaking children's vocabulary development (Liu, 2007; Liu et al., 2008; Tardif, 2006).

While it appears indeed a universal pattern that nouns are learned before verbs, our analyses above also show significant differences between Chinese and English and other

languages. In fact, based on our analyses above of the developmental stages of early vocabulary according to Caselli et al. (1999), we can see that in three of the four stages Chinese-speaking children are influenced by the language- or culture-specific characteristics (more social words, more predicates, and fewer closed-class words as compared with the numbers in English), especially with respect to the early verb advantage. Compared with other languages, verbs appear at the earliest stage of lexical development in child Mandarin, as has been reported in other verb-friendly languages (Brown, 1998; Gopnik & Choi, 1995; Tardif, 1996).

Tse, Chan and Li (2005) also reported a much stronger verb advantage in Cantonese-speaking children. However, it seems that the differences between the current Mandarin data and Tse et al.'s Cantonese data may be due to the age of samples and the data collection methods, in addition to possible cultural and linguistic differences between Cantonese and Mandarin: First, the children in Tse et al.'s study were preschoolers and much older than the children in our study (aged between 36-70 months). These children may be at a stage to produce sophisticated sentences in which verbs serve as the pivot, and therefore may have experienced more accelerated growth of verbs than children at earlier stages. An interesting possibility that follows from here is that one might observe stronger verb advantage as the children become older; Second, with regard to methods of data collection, the data of Tse et al. (2005) were based on recordings of 30-minute free play sections while our study was based on cross-sectional parental reports. Spontaneous speech within a limited time session may depend heavily on the free play context in which children tend to produce more verbs than nouns (Tardif et al., 1999).

In sum, we must conclude that both language-specific properties and universal mechanisms are responsible for the patterns we find in Chinese-speaking children's vocabulary development.

Factors affecting the age of acquisition of vocabulary

In addition to describing the developmental trajectories of early vocabulary, we have identified a number of conceptual and linguistic factors that influence the age of acquisition of words based on our analyses.

Input frequency. Input frequency has been explored extensively in child language and in general it is accepted that higher frequency words tend to be learned earlier than lower frequency words by children. The close relationship between word frequency and AoA (mostly based on adults' subjective ratings) has also been carefully investigated in the adult word reading literature, in which a strong negative correlation (from -0.40 to -0.71) is usually reported (see a review Juhasz, 2005 and Hernandez & Li, 2007). Our multiple regression analysis also confirms that word frequency is a crucial factor in modulating the AoA of words. Our results are also consistent with some previous studies of the relationship between language input and vocabulary acquisition in Chinese-speaking children. Specifically, Tardif et al. (1997; 1999) analyzed parental language input and found that Chinese caregivers tended to produce more verb types and verb tokens than English-speaking caregivers, which the authors interpreted as a significant contributor to the early verb advantage in child Chinese. Zhao and Li's (2008) simulation results also showed that early acquired words in their model had higher frequency, and the number of high-frequency verbs was larger in Chinese than in English.

One could argue that the word frequency norms used here (based on subtitles) are not as good estimates of language input to infants and toddlers as estimates based on child-directed parental speech such as those available in the CHILDES database (MacWhinney, 2000). Although child-directed parental speech from CHILDES has been successfully used as estimates of input in previous studies (e.g., Goodman et al., 2008), we did not use frequency counts based on the CHILDES parental speech in this study due to the very small sample of corpora available for Mandarin Chinese in CHILDES. Additionally, one could take a more broad view of the language environment in which the child grows up, especially with regard to overhead adult-adult conversations, and speech interactions from the TV, etc (e.g., Akhtar, Jipson, & Callanan, 2001), and in this view, perhaps the subtitle frequency norms can indeed serve as good approximations of the linguistic input to children.

Word length. Word length is another important factor, especially at the earliest stage of vocabulary development. In Figure 3, we showed that the percentage of monosyllabic verbs is notably higher than that of multisyllabic verbs, whereas for nouns, the reverse pattern is observed: children produce more multisyllabic nouns than monosyllabic nouns after they reach 22 months of age (or 100-200 vocabulary mark). In our checklist, multiple syllabic verbs occupy only 29% of the total vocabulary. We argue that word length here would provide more advantage to verbs than to nouns for child learning.

There have been some earlier studies indicating that monosyllabic items predominate early stages (from 1;5 to 2;2), and longer words gradually appear at later stages of testing intervals (Maekawa & Storkel, 2006; Storkel, 2004; Waterson, 1971; 1978). Storkel and colleague also found strong influence of word length on English-speaking children's early

word learning, and suggested that word length constrains children's word learning by affecting working memory processes necessary to hold the phonological form at the outset of word learning. Furthermore, Zhao and Li's (2008) simulated vocabulary development in English and Chinese with a computational model and also found that early acquired words tended to be words with shorter phonemic length in both languages. Together, these studies and ours show that specific linguistic properties of words significantly impact early vocabulary development.

Imageability. The SICI theory (Maguire et al., 2006), as discussed in the Introduction, highlights the contribution of four factors to early vocabulary development: Shape, Individuation, Concreteness, and Imageability of the objects to which words refer. As such, the SICI theory provide useful dimensions for us to think about the role of nonlinguistic conceptual properties of words, as opposed to linguistic properties of words. In this study, we tested role of the imageability of objects encoded by words, and found that: (1) Chinese nouns are more imageable than verbs, (2) imageability is significantly correlated with the AoA of words, and (3) imageability can uniquely account for the variances in the AoA of words in early Mandarin lexicon.

Our findings on imageability were consistent with previous studies in English (Ma et al., 2009; McDonough et al., 2011) and in Chinese (Ma et al., 2009). As compared with Ma et al.'s and McDonough et al.'s studies, our study had a larger sample of words and included more variables. Nevertheless, in all studies, including ours, imageability was confirmed as having a significant unique contribution in predicting the AoA of early words. The question of why high-imageability words are acquired earlier than low-imageability words has been

addressed by several researchers (Golinkoff & Hirsh-Pasek, 2006; Maguire et al., 2006; Tulving & Thomson, 1973). For example, words with high imageability denote objects or actions that tend to be perceptually salient and accessible, easier to represent or retrieve, may appear earlier in life, and hence be acquired easier. However, detailed studies of the correlations of imageability with concreteness, as well as the interaction between imageability and shape, individuation, and concreteness, should be conducted further across a larger range of lexical development.

Grammatical category. Grammatical category membership (noun or verb) is also found to account for a small but significant amount of variance in the AoA of early words in our regression analysis. Chinese-speaking children acquired fewer verbs than nouns across the various stages of vocabulary development. In the literature, researchers have provided several reasons to account for why nouns and verbs may differ on conceptual difficulty (Gentner, 2006; Parish-Morris, Pruden, Ma, Hirsh-Pasek, & Golinkoff, 2010). Specifically, verb learning requires children to detect which semantic components should be packaged into the meaning of a verb (Tomasello, 1995). Semantic components encoded into the concepts of verbs differ depending on the instruments (e.g., *comb*, *vacuum*), manner of the motion (e.g., *jump*, *float*), path of the motion (e.g., *approach*, *ascend*), etc. Moreover, languages differ in what semantic components are packaged into their verbs (Gentner & Boroditsky, 2001), and such complexities require an extended time for children to understand how their language typically encode verbs. However, such conceptual difficulties may be compensated by the ease of verbs in linguistic properties in Chinese. For example, the observed verb advantage in Chinese-speaking children may benefit in part from the short word length and higher input

frequencies as discussed earlier, and in part from the easy accessibility of the concepts that Chinese verbs label (e.g., English verb 'carry' can be expressed by several different Chinese verbs like *bao4*, *bei1*, *kang2*, *ti2* etc.; see a detailed discussion in Saji, Imai, Saalbach, Zhang, Shu, & Okada, 2011).

In summary, our study provides a first detailed, large-scale analysis of the developmental patterns underlying early vocabulary acquisition in Mandarin Chinese. We have identified four key factors relevant to the age or order of acquisition of words, specifically conceptual property (imageability), word frequency, word length, and grammatical category of the input language. In doing so, we hope to provide insights into current debates regarding cross-linguistic similarities and differences in the acquisition of nouns, verbs, adjectives, and closed-class words during the early years of child language development.

References

Akhtar, N., Jipson, J., & Callanan, M. A. (2001). Learning words through overhearing. *Child Development, 72*(2) : 416-430.

Au, T., Dapretto, M., & Song, Y.-K. (1994). Input vs. Constraints: early word acquisition in Korean and English. *Journal of Memory and Language, 33,* 567-582. doi:10.1006/jmla.1994.1027

Barca, L., Burani, C., & Arduino, L. S. (2002). Word naming times and psycholinguistic norms for Italian nouns. *Behavior Research Methods, Instruments, & Computers, 34*, 424-434.

Bates, E., Marchman, V., Thal, D., Fenson, L., Dale, P., Reznick, J. S., Reilly, J., & Hartung, J. (1994). Developmental and stylistic variation in the composition of early vocabulary. *Journal of Child Language, 21,* 85–123. doi: 10.1017/S0305000900008680

Behrens, H. (1998). How difficult are complex verbs? Evidence from German, Dutch, and English. *Linguistics, 36,* 679-713. doi: 10.1515/ling.1998.36.4.679

Bird, H., Franklin, S., & Howard, D. (2001). Age of acquisition and imageability ratings for a large set of words, including verbs and function words. *Behavior Research Methods, Instruments, & Computers, 33*, 73-79.

Bloom, L., Tinker, E., & Margulis, C. (1993). The words children learn: Evidence against a noun bias in early vocabularies. *Cognitive Development, 8,* 431-450. doi:10.1016/S0885-2014(05)80003-6

Bornstein, M., & Cote, L. R. (2004). Cross-linguistic analysis of vocabulary in young children: Spanish, Dutch, French, Hebrew, Italian, Korean, and American English. *Child*

Development, 75, 1115–1139. doi:10.1111/j.1467-8624.2004.00729.x

Bornstein, M. H., Haynes, O. M., & Painter, K. M. (1998). Sources of child vocabulary competence: A multivariate model. *Journal of Child Language, 25*, 367-393. doi: 10.1017/S0305000998003456

Brown, P. (1998). Children's first verbs in Tzeltal: Evidence for an early verb category. Linguistics, 36(4), 715-753.

Brysbaert, M., Buchmeier, M., Conrad, M., Jacobs, A.M., Bölte, J., & Böhl, A. (2011). The word frequency effect: A review of recent developments and implications for the choice of frequency estimates in German. *Experimental Psychology, 58*, 412-424.

Cai, Q. & Brysbaert, M. (2010). SUBTLEX-CH: Chinese word and character frequencies based on film subtitles. *PLoS ONE,* 5, e10729. doi:10.1371/journal.pone.0010729

Caselli, M. C., Bates, E., Casadio, P., Fenson, J., Fenson, L., Sanderl, L., & Weir, J. (1995). A cross-linguistic study of early lexical development. *Cognitive Development, 10,* 159–199. doi:10.1016/0885-2014(95)90008-X

Caselli, M. C., Casadio, P., & Bates, E. (1999). A comparison of the transition from first words to grammar in English and Italian. *Journal of Child Language, 26*, 69–111. doi: 10.1017/S0305000998003687

Choi, S., & Gopnik, A. (1995). Early acquisition of verbs in Korean: a cross-linguistic study. *Journal of Child Language, 22,* 497–529. doi: 10.1017/S0305000900009934

Cohen, J., Cohen, P., West, S. G., & Aiken, L. S. (2003). *Applied multiple regression/correlation analysis for the behavioral sciences* (3rd ed.). Mahwah, N.J.: L. Erlbaum Associates.

Cuetos, F., Glez-Nosti, M., Barbon, A., & Brysbaert, M. (2011). SUBTLEX-ESP: Spanish word frequencies based on film subtitles. *Psicologica, 32*, 133-143.

Dale, P. (1991). The validity of a parent report measure of vocabulary and syntax at 24 months. *Journal of Speech and Hearing Sciences, 34*, 565–571.

Dale, P., Bates, E., Reznick, S., & Morisset, C. (1989). The validity of a parent report instrument of child language at twenty months. *Journal of Child Language, 16*, 239-250.

D'Odorico, L., Carubbi, S., Salerni, N., & Calvo, V. (2001). Vocabulary development in Italian children: a longitudinal evaluation of quantitative and qualitative aspects. *Journal of Child Language, 28*, 351–372. doi: 10.1017/S0305000901004676

Fenson, L., Dale, P., Reznick, J. S., Thal, D., Bates, E., Hartung, J., et al. (1993). *MacArthur Communicative Development Inventories: User's guide and technical manual.* San Diego, CA: Singular.

Fernald, A., & Morikawa, H. (1993). Common themes and cultural variations in Japanese and American mothers' speech to infants. *Child Development, 64*, 637–656. doi: 10.1111/j.1467-8624.1993.tb02933.x

Gentner, D. (1982). Why nouns are learned before verbs: linguistic relativity versus natural partitioning. In S. A. Kuczaj (ed.), *Language development: Vol. 2. Language, thought and culture* (pp. 301–334). Hillsdale, NJ: Erlbaum.

Gentner, D. (2006). Why verbs are hard to learn, In K. Hirsh-Pasek, & R. Golinkoff (Eds.), *Action meets word: How children learn verbs* (pp. 544-564). Oxford University Press.

Gentner, D., & Boroditsky, L. (2001). Individuation, relativity, and early word learning. In M. Bowerman & S. C. Levinson (Eds.), *Language, culture, & cognition: Vol. 3. Language*

acquisition and conceptual development (pp. 215-256). New York: Cambridge University Press.

Gillette, J., Gleitman, H., Gleitman, L., & Lederer, A. (1999). Human simulations of vocabulary learning. *Cognition, 73,* 135–176. doi: S0010-0277(99)00036-0 [pii]

Gleitman, L. R., & Gleitman, H. (1994). A picture is worth a thousand words, but that's the problem. In B. Lust, M. Suner, & J. Whitman (Eds.), *Syntactic theory and first language acquisition: Cross-linguistic perspectives. Vol. 1: Heads, projections, and learnability* (pp. 291–299). Hillsdale, NJ: Erlbaum.

Goldfield, B. A. (1993). Noun bias in maternal speech to one year olds. *Journal of Child Language, 20*, 85–99. doi: 10.1017/S0305000900009132

Golinkoff, R.M., & Hirsh-Pasek, K. (2006). Baby wordsmith: from associationist to social sophisticate. *Current Directions in Psychological Science, 15*, 30–33.

Golinkoff, R. M., Mervis, C. V., & Hirsh-Pasek, K. (1994). Early object labels: The case for a developmental lexical principles framework. *Journal of Child Language, 21,* 125-155. doi: 10.1017/S0305000900008692

Goodman, J. C., Dale, P. S. & Li, P. (2008). Does frequency count? Parental input and the acquisition of vocabulary. *Journal of Child Language, 35,* 521-559. doi: 10.1017/S0305000907008641

Gopnik, A., & Choi, S. (1995). Names, relational words, and cognitive development in English- and Korean-speakers: nouns are not always learned before verbs. In M. Tomasello & W. Merriman (Eds.), *Beyond names for things: young children's acquisition of verbs* (pp. 63–80). Hillsdale, NJ: Lawrence Erlbaum.

Hao, M., Shu, H., Xing, A., & Li, P. (2008). Early vocabulary inventory for Mandarin Chinese. *Behavior Research Methods, 40,* 728-733. doi: 10.3758/BRM.40.3.728

Hernandez, A., & Li, P. (2007). Age of acquisition: Its neural and computational mechanisms. *Psychological Bulletin*, *133*, 638-650.

Juhasz, B. (2005). Age-of-acquisition effects in word and picture identification. *Psychological Bulletin, 131*(5), 684-712.

Kauschke, C., & Hofmeister, C. (2002). Early lexical development in German: a study on vocabulary growth and vocabulary composition during the second and third year of life. *Journal of Child Language, 29,* 735–757. doi: 10.1017/S0305000902005330

Keuleers, E., Brysbaert, M. & New, B. (2010). SUBTLEX-NL: A new frequency measure for Dutch words based on film subtitles. *Behavior Research Methods, 42*(3), 643-650.

Kim, M., McGregor, K. K., & Thompson, C. K. (2000). Early lexical development in English- and Korean-speaking children: language-general and language-specific patterns. *Journal of Child Language, 27,* 225-254. doi: 10.1017/S0305000900004104

Kong, L., Hu, D., Ouyang, J., Chen, C., Ding, L., Wang, X., et al. (2004). 汉族儿童实词习得研究[Research on the acquisition of content words in Chinese-speaking children]，合肥：安徽大学出版社[Hefei: Anhui University Press].

Li, P. (1998). Crosslinguistic variation and sentence processing: The case of Chinese. In D. Hillert (Ed.), *Sentence processing: A crosslinguistic perspective*. (Vol. 31 of *Syntax and Semantics*; pp. 33-51). San Diego, CA: Academic Press.

Li, P., Zhao, X., & MacWhinney, B. (2007). Dynamic self-organization and early lexical development in children. *Cognitive science*, *31*(4), 581–612.

Liu, S. X. (2007). *Early lexical development in English, Mandarin, and Cantonese: A cross-linguistic study based on CHILDES*. Unpublished M.A. thesis, University of Richmond.

Liu, S. X., Zhao, X. W., & Li, P. (2008). Early lexical development: A corpus-based study of three languages. In B. C. Love, K. McRae, & V. M. Sloutsky (Eds.), *Proceedings of the 30th Annual Conference of the Cognitive Science Society* (pp. 1343-1348). Austin, TX: Cognitive Science Society.

Liu, Y., Shu, H., & Li, P. (2007). Word naming and psycholinguistic norms: Chinese. *Behavior Research Methods, 39,* 192-198.

Ma, W., Golinkoff, R. M., Hirsh-Pasek, K., McDonough, C., & Tardif, T. (2009). Imageability predicts the age of acquisition of verbs in Chinese children. *Journal of Child Language, 36,* 405-423. doi: 10.1017/S0305000908009008

MacWhinney, B. (2000). The CHILDES Project: Tools for Analyzing Talk. Mahwah, NJ: Lawrence Erlbaum Associates.

Maekawa, J., & Storkel, H. L. (2006). Individual differences in the influence of phonological characteristics on expressive vocabulary development by young children. *Journal of Child Language, 33,* 439-459. doi: 10.1017/S0305000906007458

Maguire, M. J., Hirsh-Pasek, K., & Golinkoff, R. M. (2006). A united theory of word learning: Putting verb acquisition in context. In K. Hirsh-Pasek & R. M. Golinkoff (Eds.), *Action meets word: How children learn verbs* (pp. 364-391). Oxford University Press.

Markman, E. M. (1987). How children constrain the possible meanings of words. In U. Neisser (Ed.), *Concepts and conceptual development: Ecological and intellectual factors*

in categorization (pp. 255–287). New York: Cambridge University Press.

McDonough, C., Song, L., Hirsh-Pasek, K., Golinkoff, R. M., & Lannon, R. (2011). An image is worth a thousand words: Why nouns tend to dominate verbs in early word learning. *Developmental Science, 14,* 181-189. doi: 10.1111/j.1467-7687.2010.00968.x

Nelson, K. (1973). Structure and strategy in learning to talk. *Monograph of the Society for Research in Child Development, 38,* 1–135.

New, B., Brysbaert, M., Veronis, J., & Pallier, C. (2007). The use of film subtitles to estimate word frequencies. *Applied Psycholinguistics, 28*, 661-677.

Parish-Morris, J., Pruden, S. M., Ma, W., Hirsh-Pasek, K., & Golinkoff, R. M. (2010). A world of relations: Relational words. In B. C. Malt, and P. Wolff (Eds.) *Words and the mind: How words capture human experience* (pp. 219-242). Oxford University Press.

Pine, J. M., & Lieven, E. V. (1990). Referential style at thirteen months: why age-defined cross-sectional measures are inappropriate for the study of strategy differences in early language development. *Journal of Child Language, 17,* 625–631. doi: 10.1017/S0305000900010916

Saji, N., Imai, M., Saalbach, H., Zhang, Y., Shu, H., & Okada, H. (2011). Word learning does not end at fast-mapping: evolution of verb meanings through reorganization of an entire semantic domain. *Cognition, 118*(1), 45-61.

Sandhofer, C., Smith, L., & Luo, J. (2000). Counting nouns and verbs in the input: differential frequencies, different kinds of learning? *Journal of Child Language, 27*, 561-585. doi: 10.1017/S0305000900004256

Storkel, H. L. (2004). Methods for minimizing the confounding effects of word length in the

analysis of phonotactic probability and neighborhood density. *Journal of Speech Language and Hearing Research, 47,* 1454-1468. doi:10.1044/1092-4388(2004/108)

Tamis–LeMonda, C. S., Bornstein, M. H., Cyphers, L., Toda, S., & Ogino, M. (1992). Language and play at one year: A comparison of toddlers and mothers in the United States and Japan. *International Journal of Behavioral Development, 15*, 19–42. doi: 10.1177/016502549201500102

Tardif, T. (1996). Nouns are not always learned before verbs: evidence from Mandarin speakers' early vocabularies. *Developmental Psychology, 32,* 492–504.

Tardif, T. (2006). But are they really verbs? Chinese words for action. In K. Hirsh-Pasek, & R. Golinkoff (Eds.), *Action meets word: How children learn verbs* (pp. 477-498). Oxford University Press.

Tardif, T., Fletcher, P., Liang, W., Zhang, Z., Kaciroti, N., & Marchman, V. A. (2008). Baby's first 10 words. *Developmental Psychology, 44,* 929-938. doi:10.1037/0012-1649.44.4.929

Tardif, T., Gelman, S. A., & Xu, F. (1999). Putting the 'noun bias' in context: a comparison of English and Mandarin. *Child Development, 70*, 620–635. doi: 10.1111/1467-8624.00045

Tardif, T., Shatz, M., & Naigles, L. (1997). Caregiver speech and children's use of nouns versus verbs: a comparison of English, Italian, and Mandarin. *Journal of Child Language, 24,* 535–565. doi: 10.1017/S030500099700319X

Tomasello, M. (1995). Pragmatic contexts for early verb learning. In M. Tomasello & W. E. Merriman (Eds.), *Beyond the name for things: Young children's acquisition of verbs* (pp.

115–146). Hillsdale, NJ: Lawrence Erlbaum.

Tse, S. K., Chan, C., & Li, H. (2005). Is the expressive vocabulary of young Cantonese speaker noun or verb dominated? *Early Child Development and Care, 175*, 215-227.

Tulving, E., & Thomson, D.M. (1973). Encoding specificity and retrieval processes in episodic memory. *Psychological Review, 80*, 352–373.

Waterson, N. (1971). Child phonology: a prosodic view. *Journal of Linguistics, 7*, 179-211. doi: 10.1017/S0022226700002917

Waterson, N. (1978). Growth of complexity in phonological development. In N. Waterson & C. Snow (Eds.), *The development of communication*. New York: Wiley.

Waxman, S. R., & Booth, A. E. (2003). The origins and evolution of links between word learning and conceptual organization: new evidence from 11-month-olds. *Developmental Science, 6,* 130-137. doi: 10.1111/1467-7687.00262

Zhao, X. W., & Li, P. (2008). Vocabulary development in English and Chinese: A comparative study with self-organizing neural networks. In B. C. Love, K. McRae, & V. M. Sloutsky (Eds.), *Proceedings of the 30th Annual Conference of the Cognitive Science Society* (pp. 1900-1905). Austin, TX: Cognitive Science Society.

Table 1

Child participants' demographic information for each age group

Group [a]	Number	Male	Mean age	Interviewee [b]	Mother's education [c]	Income [d]
12	45	24/45	12.44	35/44 [e]	36/43	31/42
13	67	34/60	13.54	43/57	47/55	44/51
14	53	26/52	14.42	29/46	39/46	41/47
15	57	23/49	15.39	31/46	36/45	39/46
16	55	28/55	16.53	22/40	34/40	33/37
17	40	19/40	17.61	29/40	33/40	30/37
18	50	24/50	18.56	34/50	40/49	44/49
19	50	26/50	19.38	38/50	42/49	43/49
20	50	26/50	20.47	30/49	42/49	44/50
21	49	19/49	21.60	32/47	40/47	41/45
22	46	25/46	22.49	30/46	36/46	35/45
23	50	23/50	23.54	30/50	39/50	38/49
24	50	24/50	24.61	40/50	43/50	46/50
25	43	24/43	25.66	29/42	35/43	31/39
26	47	23/47	26.66	33/47	40/47	43/46
27	42	15/42	27.66	35/42	31/41	34/41
28	43	27/43	28.59	32/43	35/43	29/38
29	49	22/49	29.61	40/49	39/49	44/49
30	40	23/40	30.55	28/40	30/39	31/39

Note: *a* - groups according to children's age in month; *b* - proportion of mother interviewee; *c* - proportion of mothers' highest education above a college level; *d* - proportion of sum of father's and mother's income more than 4000 CNY per month; *e* - the number (44) is less than the total number (45) of this group because some data are not available. The same situation also applies elsewhere in the table.

Table 2

First 50 words in the production vocabulary of Chinese-speaking children

Rank	Word	Translation	Semantic category	Part-of-speech
1.	爸爸	daddy*	People	Social words
2.	妈妈	mommy*	People	Social words
3.	奶奶	grandma*	People	Social words
4.	爷爷	grandpa*	People	Social words
5.	宝宝	baby*	People	Social words
6.	不	no*	Game and routines	Social words
7.	阿姨	aunt	People	Social words
8.	哥哥	elder brother	People	Social words
9.	猫	cat*	Animal	common nouns
10.	狗	dog*	Animal	common nouns
11.	姐姐	elder sister	People	Social words
12.	大	big	Descriptive words	Predicates
13.	姥姥	mother's mother	People	Social words
14.	鱼	fish*	Animal	common nouns
15.	灯	lamp	Small household item	common nouns
16.	抱	hug	Verb	Predicates
17.	弟弟	younger brother	People	Social words
18.	家	home	Place to go	common nouns
19.	门	door	Furniture and rooms	common nouns
20.	妹妹	younger sister	People	Social words
21.	姥爷	mother's father*	People	Social words
22.	打	hit	Verb	Predicates
23.	水	water	Food and drink	common nouns
24.	花	flower	outside things	common nouns
25.	鸡	chicken	Animal	common nouns
26.	谢谢	thank you*	Game and routines	Social words

27.	饭	food; rice	Food and drink	common nouns
28.	娃娃	doll	Toys	common nouns
29.	拿	bring; take	Verb	Predicates
30.	要	want	Verb	Predicates
31.	再见	bye*	Game and routines	Social words
32.	不要	don't	Game and routines	Social words
33.	鸟	bird*	Animal	Common nouns
34.	我	I	Pronoun	Closed-class
35.	鞋	shoe	clothing	Common nouns
36.	吃	eat	Verb	Predicates
37.	肉	meat	Food and drink	Common nouns
38.	床	bed	Furniture and rooms	Common nouns
39.	书	book*	Toys	Common nouns
40.	好	good	Descriptive word	Predicates
41.	马	horse	Animal	Common nouns
42.	姑姑	father's sister	People	Social words
43.	鱼	fish*	Food and drink	Common nouns
44.	手	hand	Body parts	Common nouns
45.	走	go; walk	Verb	Predicates
46.	是	yes*	Game and routines	Social words
47.	叔叔	father’s younger brother	People	Social words
48.	宝宝自己的名字	baby's name	People	Social words
49.	烫	hot*	Descriptive word	Predicates
50.	树	tree	Outside things	Common nouns

Note: * = words that were listed in English-speaking children’s first acquired 50 words (Adapted from Table 2 in Caselli, Bates, Casadio, Fenson, Fenson, Sanderl, & Weir, 1995)

Table 3

The mean and standard deviant (in parentheses) of variables for common nouns and verbs

	Imageability	Word frequency	Word length	AoA
Verbs (*N=102*)	5.41 (1.01)	2.01 (0.81)	1.34 (0.48)	23.81 (3.32)
Nouns (*N=348*)	6.51 (0.40)	1.03 (0.75)	1.91 (0.55)	24.01 (3.52)

Note: the values for imageability are based on a 7-point scale; the unit of AoA is the child's age in months; word length is counted in syllables; word frequency values are log10 transformed of times of occurrence per million in database.

Table 4

Correlations among the variables and age of acquisition for common nouns and verbs *(N=450)*

Variables	1	2	3	4	5
1.Age of acquisition	---				
2.Imageability	-.099*	---			
3.Word frequency	-.451***	-.428***	---		
4.Word length	-.428***	-.272***	.539***	---	
5.Grammatical category	.023	.612***	-.473***	-.418***	---

Note: *$p < .05$, ***$p < .001$ *(2-tailed)*

Table 5

Multiple regression analyses of imageability, word length, word frequency, and word grammatical category predicting the age of acquisition of words *(N=450)*

Variables	*B*	*S.E.*	*β*	*Squared semi-partial coefficient (sr^2)*
Imageability	-.1.357	.222	-.296**	.052
Word frequency	-1.623	.192	-.481***	.141
Word length	-2.376	.344	-.315***	.067
Grammatical category	-1.279	.421	-.154**	.013

Note: ** $p < .01$, *** $p < .001$. The squared semi-partial coefficient (sr^2) indicates the variances accounted for uniquely by each predictor after the effects of the other predictors were controlled.

Figure Captions

Fig. 1(a). Productive vocabulary size in percentages for common nouns, social words, predicates, and closed-class words as a function of the total vocabulary size from 12 to 16 months for Mandarin Chinese-speaking children

Fig. 1(b). Productive vocabulary size in percentages for common nouns, social words, predicates, and closed-class words as a function of the total vocabulary size from 8 to 16 months for English-speaking children; Adapted from Table 2 in Caselli, Bates, Casadio, Fenson, Fenson, Sanderl, & Weir (1995).

Fig. 2. Productive vocabulary size in percentages for common nouns, social words, predicates, and closed-class words as a function of total vocabulary size from 17 to 30 months for Mandarin Chinese-speaking children

Fig. 3. Percentages of monosyllabic and multisyllabic words within nouns and within verbs as a function of total vocabulary size from 17 to 30 months for Chinese-speaking children

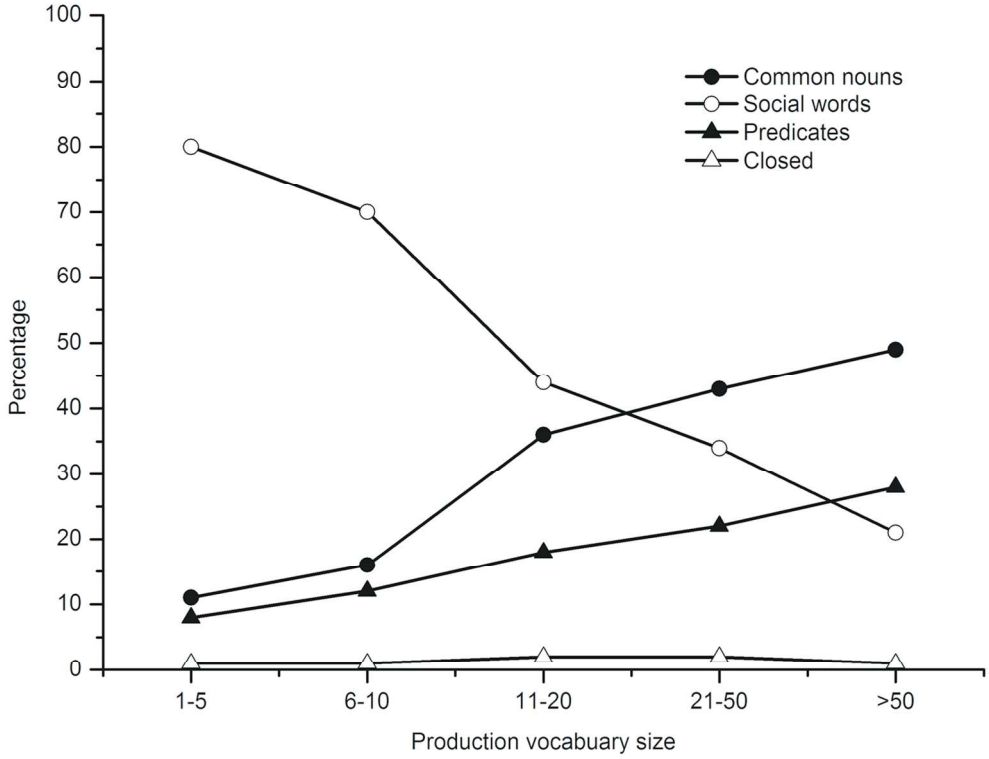

Fig. 1(a). Productive vocabulary size in percentages for common nouns, social words, predicates, and closed-class words as a function of the total vocabulary size from 12 to 16 months for Mandarin Chinese-speaking children
124x94mm (300 x 300 DPI)

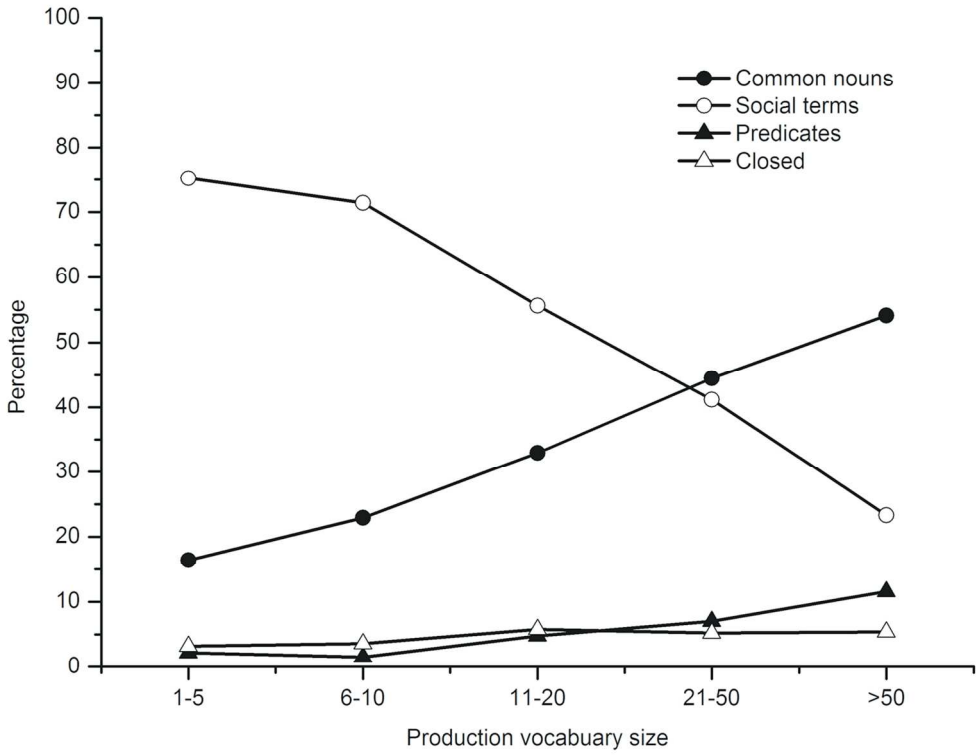

Fig. 1(b). Productive vocabulary size in percentages for common nouns, social words, predicates, and closed-class words as a function of the total vocabulary size from 8 to 16 months for English-speaking children; Adapted from Table 2 in Caselli, Bates, Casadio, Fenson, Fenson, Sanderl, & Weir (1995).
125x96mm (300 x 300 DPI)

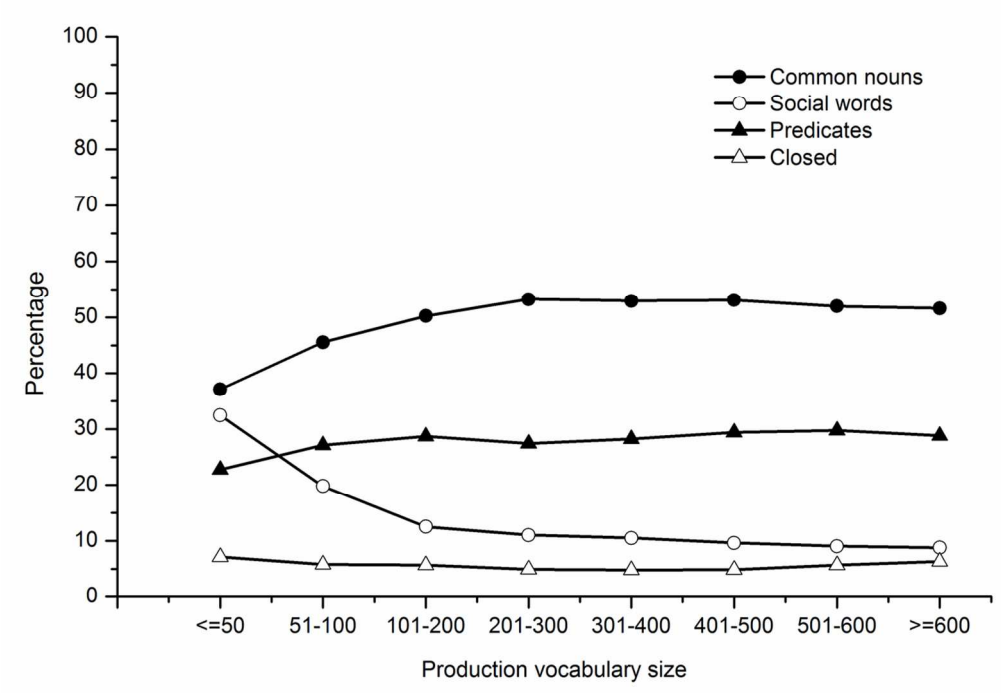

Fig. 2. Productive vocabulary size in percentages for common nouns, social words, predicates, and closed-class words as a function of total vocabulary size from 17 to 30 months for Mandarin Chinese-speaking children
114x79mm (300 x 300 DPI)

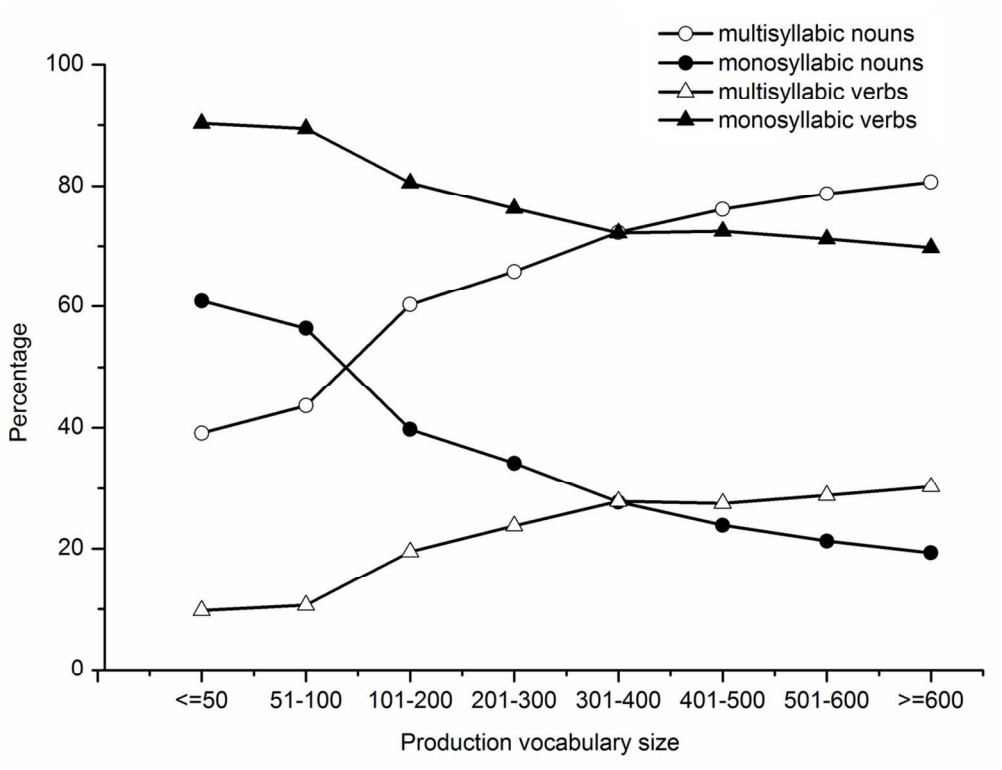

Fig. 3. Percentages of monosyllabic and multisyllabic words within nouns and within verbs as a function of total vocabulary size from 17 to 30 months for Chinese-speaking children.
113x86mm (300 x 300 DPI)