CAMBRIDGE
UNIVERSITY PRESS

ARTICLE

Developing early lexical composition in Mandarin-speaking children: A longitudinal study

Yi-Fang HUNG 🆔 and Chien-Ju CHANG* 🆔

Department of Human Development and Family Studies, National Taiwan Normal University
*Corresponding author: Chien-Ju CHANG, Department of Human Development and Family Studies, National Taiwan Normal University, 162, Section 1, Heping East Road, Taipei 10610, Taiwan.
E-mail: changch2@ntnu.edu.tw

Abstract

This study investigated the developmental pattern of early lexical production and composition in Mandarin-speaking children. Forty Mandarin-speaking children and their parents participated in this one-and-a-half-year longitudinal study, and naturalistic samples of parent-to-child speech in toy play were collected when the children were 1;8, 2;2, and 3;0. The results showed that children's lexical production increased significantly between ages 1;8 and 3;0. The proportion of closed-class words increased significantly with age, whereas the proportion of common nouns showed the inverse pattern, indicating the role of grammatical words increased as the children grew. Furthermore, nouns and verbs were predominant in Mandarin-speaking children between ages 1;8 and 3;0, and Mandarin-speaking children used more verbs than nouns at 2;2 and 3;0 in the toy play context. The longitudinal study clarifies early lexical development in Mandarin-speaking children, which provides a valuable contrast for different language systems.

Introduction

Research conducted in a variety of languages suggests that infants and toddlers experience rapid linguistic and communicative development (Bates, Marchman, Thal, Fenson, Dale, Reznick, Reilly, & Hartung, 1994; Cadime, Moreira, Santos, Silva, Ribeiro, & Viana, 2019; Chi, 2002; Choi & Gopnik, 1995; Goldfield & Reznick, 1990; Hao et al., 2015; Liu & Chen, 2015; Tardif, Fletcher, Liang, & Kaciroti, 2009). Some of these studies indicate that early productive lexicons are related to later stages of cognitive and language development (Lyytinen, Laakso, Poikkeus, & Rita, 1999; Nelson, 1973; Pérez-Pereira & Cruz, 2018; Rescorla, Roberts, & Dahlsgaard, 1997; Rowe, Raudenbush, & Goldin-Meadow, 2012), and could predict future academic achievement (Bleses, Makransky, Dale, Højen, & Ari, 2016; Lee, 2011). Therefore, early lexical development in children has received much research attention over the last several decades.

One of the important issues in early language acquisition research is the developmental pattern of the early lexicon and whether the developmental patterns in children are the same across languages. Numerous studies of children's early lexical development in a variety of languages have reported that children typically produce their first words around the end of their first year, and that after a few months, they experience a rapid growth in vocabulary, the so-called "vocabulary spurt" (Caselli, Bates, Casadio, Fenson, Fenson, Sanderl, & Weir, 1995; Choi & Gopnik, 1995; D'Odorico, Carubbi, Salerni, & Calvo, 2001; Goldfield & Reznick, 1990; Hao, Shu, Xing, & Li, 2008; Nelson, 1973; Tardif et al., 2009). Many researchers have reported noun dominance in early lexical acquisition. For example, Nelson's (1973) study of English-speaking children revealed that nominals were the most common category. Similarly, Goldfield and Reznick's (1990) study of early lexical acquisition of English-speaking children from 1;2 to 1;10 found that almost three-quarters of the vocabulary added in this period comprised nouns. Several studies have reported noun dominance of early vocabulary in the case of English (Bates et al., 1994; Bornstein, Cote, Maital, Painter, Park, Pascual, Pêcheux, Ruel, Venuti, & Vyt, 2004; Caselli et al., 1995; Gentner, 1982; Goldfield & Reznick, 1990), Italian (Bornstein et al., 2004; Caselli et al., 1995), and German (Gentner, 1982). However, more recent studies of Korean (Choi & Gopnik, 1995), Japanese (Ogura, Dale, Yamashita, Murase, & Mahieu, 2006), and Mandarin (Tardif, 1996) have challenged the universality of a noun-bias. For example, Choi and Gopnik's (1995) study investigated lexical development of nine Korean children from 1;2 to 1;10 using maternal reports, and their findings revealed that six of the nine children showed the first verb spurt before the first noun spurt. Tardif's (1996) study of 22-month-old Mandarin-speaking children also demonstrated that most of children in the study produced more verbs than nouns in their spontaneous speech. These studies revealed that nouns were not always predominant over the course of early lexical acquisition, and the use of more verbs than nouns was observed in early lexical development. The question whether children demonstrate a universal noun-bias across languages in early lexical acquisition has aroused researchers' interest, and the possible factors influencing the apparent noun or verb dominance in children's early lexical development have received extensive discussion and investigation.

The developmental changes of early lexical composition

In the last three decades, a growing number of studies have explored the development of early lexical composition, including nouns, verbs, adjectives, adverbs, prepositions, and conjunctions. Studies of lexical composition exploring the distribution of the semantic and syntactic categories of words have provided valuable information about the development of early language acquisition and the relationship between semantic and grammatical development. Previous studies of early lexical development have shown that children's lexical composition changes with an increase in vocabulary or age (Bassano, 2000; Bassano, Eme, & Champaud, 2005; Bates et al., 1994; Caselli et al., 1995; Day & Elison, 2022; D'Odorico et al., 2001; Hao et al., 2015; Kauschke & Hofmeister, 2002; Liu & Chen, 2015; Liu, Zhao, & Li, 2008; Yang, 2015). For example, Bates et al. (1994) conducted a developmental study of early lexical composition using parental reports in English-speaking children aged between 0;8 and 2;6. In their study, they analyzed lexical composition using three variables: the proportions of common nouns, predicates (verbs and adjectives), and closed-class items (pronoun, prepositions, question words, quantifiers, articles, auxiliary verbs, and connectives). Their findings revealed that the proportion of common nouns increased significantly and occupied the highest proportion of total vocabulary in the period from 1 to 100 words; verbs and other predicates increased slowly

at first, but showed the greatest gains between 100 and 400 words; for closed-class words, there was less proportional development between 0 and 400 words, followed by a sharp increase after 400 words. Caselli et al. (1995) analyzed the lexical composition of English- and Italian-speaking children between ages 0;8 and 1;4 using parental reports, and found that in both languages common nouns predominated and grew rapidly when children learned their first 50 to 100 words, while verbs, adjectives, and grammatical function words were extremely rare until children had vocabularies of at least 100 words. Bates et al.'s (1994) and Caselli et al.'s (1995) results indicate that the change in early lexical composition in children revealed a shift in emphasis from common nouns to predicates to grammatical function words as the children's vocabulary increased.

Several longitudinal studies of early lexical development based on naturalistic speech samples have pointed out that nouns are not always predominant over the course of early lexical acquisition (Kauschke & Hofmeister, 2002; Ogura et al., 2006; Tardif, 1996). For instance, Kauschke and Hofmeister (2002) conducted a longitudinal study of early lexical development on German children between ages 1;1 and 3;0 using spontaneous speech, and found that relational words and personal-social words were predominant during the first half of the second year of life. Moreover, their findings showed that children used more verbs than nouns when they were 3;0. Ogura et al. (2006) investigated the use of nouns and verbs by Japanese children between ages 1;0 and 2;0 using spontaneous speech. Their study was conducted on children who were similar in age to the study by Bates et al. (1994), but they further grouped children by grammatical stage. Their findings revealed that the verb types used in the toy play context increased from the single-word stage to the syntactic stage, and more verb types than noun types were used at the syntactic stage according to the narrower definition of nouns, which excluded people and proper nouns. In addition, Ogura et al. (2006) found that there were more children with verb dominance than those with noun dominance in the toy play context at the syntactic stage using narrow definition of nouns. Ogura et al.'s (2006) findings, with some differences from those of Bates et al.'s, showed that noun dominance was not persistent, and children used more verbs than nouns in a particular context and grammatical stage. Both Kauschke and Hofmeister's (2002) and Ogura et al.'s (2006) results indicated that noun dominance was not constant across early lexical development and suggested that children's lexical composition changes with age or the emergence of grammar.

Many studies of early lexical composition using parental reports have indicated that there was an apparent noun dominance and the acquisition of verbs was universally later than that of nouns in children's early production (Bates et al., 1994; Caselli et al., 1995; Hao et al., 2015; Liu & Chen, 2015). However, some studies on early vocabulary development based on naturalistic speech showed that children used more verbs than nouns in their early lexical development (Ogura et al., 2006; Tardif, 1996). Regarding Mandarin-speaking children, Liu and Chen (2015) and Hao et al. (2015) investigated their early lexical development using parental reports and found they have a greater proportion of nouns than other word categories. However, Tardif (1996) reported that 9 out of 10 twenty-two-month-old Mandarin-speaking children produced more verbs than nouns in their naturalistic speech. Previous studies have indicated that although parent-report and observational measures were significantly correlated, there were also systematic quantitative differences between these two measures (Pine, Lieven, & Rowland, 1996; Tardif, Gelman, & Xu, 1999). Compared to parental reports, spontaneous speech samples may provide more fine-grained measures of children's early vocabulary and usage, which concurrently present more semantic, syntactic, and pragmatic properties of children's language production.

Another issue of interest to researchers is the commonality or variability of early lexical development (Bassano et al., 2005; Bates et al., 1994). Bassano et al.'s (2005) study of early lexical development in French children between ages 1;8 and 3;3 using spontaneous speech showed that inter-individual variability in lexical composition was greater at 1;8 than at 2;6 and 3;3, which indicated that individual variation decreased in the course of the third year. However, research has less frequently explored the core lexicons shared by the majority of children and whether they change with age, and these issues remain unclear in Mandarin-speaking children.

Possible influencing factors of lexical composition

As mentioned above, the stage of language development may be an influencing factor of lexical composition. Lexical composition changes with vocabulary growth and the emergence of grammar (Bates et al., 1994; Caselli et al., 1995; Kauschke & Hofmeister, 2002; Ogura et al., 2006; Trudeau & Sutton, 2011). A high interdependence between lexical and grammatical skills across a variety of languages has been reported in previous studies (Caselli, Casadio, & Bates, 1999; Chi, 2002; Dixon & Marchman, 2007; Trudeau & Sutton, 2011; Viana, Pérez-Pereira, Cadime, Silva, Santos, & Ribeiro, 2017). Trudeau and Sutton (2011) investigated expressive vocabulary and early grammar of 16- to 30-month-old children using parental reports. Their results demonstrated that almost 40% of children combined words at 1;8 and nouns occupied the highest proportion of total vocabulary in this period. As the children grew, the frequency of the use of multi-word utterances increased, and the proportions of verbs and grammatical function words used by children also increased. Similarly, Parisse and Le Normand's (2000) study investigated the morphosyntax produced by two-year-old French-speaking children using spontaneous speech. Their findings showed that children primarily use content words in the first word combination, and then they progressively introduce functional words in their utterances to produce more complex structure. The findings of these studies indicate that the proportions of different word categories change from the single-word stage to the syntactic stage. However, to the best of our knowledge, longitudinal studies using spontaneous speech to investigate lexical composition from the single-word stage to the syntactic stage in Mandarin-speaking children are relatively scarce.

Language systems may be the other factor that influences early lexical composition. Choi and Gopnik (1995) investigated children's early lexical development in English and Korean using parental reports. They found that among Korean-speaking children, both verbs and nouns are dominant categories from the single-word stage. Furthermore, six of the nine Korean-speaking children studied showed the first verb spurt before their first noun spurt, but no such early verbs spurt was found in the English-speaking children. Similarly, Tardif et al. (1999) compared the proportions of nouns and verbs in the early vocabularies of English- and Mandarin-speaking toddlers, and found that Mandarin-speaking toddlers used relatively fewer nouns and more verbs than English-speaking toddlers.

Mandarin syntax has some typological characteristics that differ from those of other languages. First, Mandarin is an isolating language. There is very little morphological complexity in Mandarin. Some of the types of morphemes that many languages have are not found in Mandarin. For example, Mandarin pronouns do not have case markers, and Mandarin verbs have no tense markers. Second, Mandarin is a pro-drop language,

permitting omission of the subject or/and the object, and is a topic-comment language with more flexible SVO word order than English (Hsu, 1996; Li & Thompson, 2009). The linguistic features of Mandarin, such as lack of tense, the phenomenon of pronoun-dropping, and topic prominence, may explain why more verbs were used by Mandarin-speaking children than by English-speaking children (Tardif, 1996; Tardif, Shatz, & Naigles, 1997; Tardif et al., 1999).

Past research on the early lexical development of Mandarin-speaking children has mostly focused on children aged one to two years (Tardif, 1996; Tardif et al., 1997, 1999; Tardif, Fletcher, Liang, Zhang, Kaciroti, & Marchman, 2008) and used parental reports (Chi, 2002; Hao et al., 2008, 2015; Liu & Chen, 2015; Tardif et al., 2008, 2009). Although there have been a few corpus-based studies of early lexical development in Mandarin-speaking children (Liu et al., 2008; Yang, 2015), most used a cross-sectional design. Only Hsu (1996) used both cross-sectional and longitudinal data. The sample size of each age group in his study, however, was relatively small (n=12). Longitudinal studies using spontaneous speech and following the same group of Mandarin-speaking children to investigate their lexical development from learning their first words to the production of two-word or even longer sentences are relatively scarce. How does lexical composition change from the single-word stage to the syntactic stage in Mandarin-speaking children? Do the changes in lexical composition follow a universal pattern or a language-specific pattern? Does noun or verb dominance appear during this period? These questions remain unclear and require further longitudinal studies.

Goals of the present study

In this study, we investigate the changes in lexical production and composition in Mandarin-speaking children between ages 1;8 and 3;0 using spontaneous speech samples in order to determine the developmental pattern of early lexical production in Mandarin-speaking children. Three specific goals of this study are: (1) to examine the developmental pattern of early lexical production and composition in Mandarin-speaking children between ages 1;8 and 3;0; (2) to examine whether there is noun or verb dominance in Mandarin-speaking children between ages 1;8 and 3;0; and (3) to identify which words are shared by the majority of the children (core lexicons of children) and which lexical categories the core lexicons belong to.

Method

Participants

Forty 20-month-old children (23 boys, 17 girls) and their parents living in northern Taiwan participated in this one-and-a-half-year longitudinal study. All participating children were first-borns in middle-class families and were typically developing. One child whose mother spoke English with her in the toy play context was excluded from this study. All children in this study spoke Mandarin Chinese as their first language. They were immersed in a Mandarin environment, and bilingual education was not provided during the period of data collection. The mean ages of the mothers and fathers were 32.40 ($SD = 3.54$) and 34.55 ($SD = 3.76$) years, respectively. Most participating mothers (75.0%) and fathers (67.5%) had at least a college-level education.

Procedures and data collection

Parents were briefed about the research purpose and procedures, and signed consent was obtained before participation in the study. All children and their parents were visited at home when the children were 1;8 (Time 1), 2;2 (Time 2), and 3;0 (Time 3). Parents were instructed to play with their children as they usually did at home during each visit, and a set of toys, including puppets, jigsaw puzzles, and Legos, was provided for children and parents to play. Parent-child interactions during play were videotaped, and approximately 15 minutes of videotaped sessions were included in the analyses as per previous studies (Ogura et al., 2006; Southwood & Russell, 2004).

Coding and analyses

Language production in parent-child interactions during play was transcribed and analyzed using the Child Language Data Exchange System (CHILDES; MacWhinney, 2000; MacWhinney & Snow, 1990). All transcripts were coded by a group of graduate research assistants majoring in child development using the CHAT format of the CHILDES (MacWhinney, 2000), and the utterances and words in the transcripts were segmented based on the Taiwan Corpus of Child Mandarin (TCCM) word segmentation rules (Cheung, Chang, Ko, & Tsay, 2011). The TCCM word segmentation standard is mainly based on the Chinese National Standard (CNS) of segmentation principles for Chinese literature processing as the model, but some revisions are made in consideration of the characteristics of children's language development. For example, negative words such as *mei2you3* 'not' and *bu2yao4* 'no' are treated as a segmentation unit. In addition, *dao4* 'arrive', as a postverbal directional verb complement, and *diao4* 'drop', as a post-verbal resultative verb complement, are treated as a segmentation unit and separate from the preceding verb.

After utterance and word segments, the MOR command of the CLAN programs of the CHILDES (MacWhinney, 2000) was used to conduct automatically morphosyntactic analyses and tag grammatical labels (part-of-speech) on each word of the transcripts. Although the MOR command of the CLAN programs of the CHILDES (MacWhinney, 2000) could tag most words correctly in Chinese, there were still some polysemous words belonging to different lexical categories when presented in different contexts that were confused easily and tagged incorrectly. For these situations, we checked and tagged the words manually according to the context in which these words occurred.

The measures of early language production included the number of total utterances, total number of words (word tokens), total number of different words (word types), mean length of utterance (MLU), and frequencies and proportions of word types in 12 word classes and three word categories. The 12 word classes included common nouns (such as *dian4hua4* 'phone', *shou3* 'hand', and *mian4bao1* 'bread'), main verbs (such as *wan2* 'play' and *chi1* 'eat'), adjectives (such as *da4* 'big' and *hong2* 'red'), pronouns (such as *wo3* 'I', *ni3* 'you', and *zhe4* 'this'), prepositions (such as *zai4* 'at', *yong4* 'with', and *gei3* 'for'), conjunctions (such as *hai2you3* 'as well as' and *ran2hou4* 'then'), quantifiers (such as *hen3duo1* 'a lot' and *xie1* 'some'), question words (such as *zen3me* 'how', *shen2me* 'what', and *na3li3* 'where'), adverbs (such as *bu4* 'not', *dou1* 'all', and *hao3* 'very'), numbers (such as *yi1* 'one' and *er4* 'two'), classifiers (such as *ge4* used for people or individual things, *ke1* used for small or spherical objects, and *zhi1* used for long or stick-like objects), and sentence-final particles (such as *le* and *a1*). Of these 12 word classes,

Table 1. Classification of Word Categories and Word Classes

Word category	Word class	Mandarin examples	English translation of examples
Common nouns	common nouns	*dian4hua4, shou3*	phone, hand
Predicates	main verbs	*wan2, chi1*	play, eat
	adjectives	*da4, hong2*	big, red
Closed-class words	pronouns	*wo3, ni3*	I, you
	prepositions	*zai4, yong4*	at, with
	conjunctions	*hai2you3, ran2hou4*	as well as, then
	quantifiers	*hen3duo1, xie1*	a lot, some
	question words	*zen3me, shen2me*	how, what
Other	adverbs	*bu4, dou1*	not, all
	numbers	*yi1, er4*	one, two
	classifiers	*ge4, ke1*	–
	sentence-final particles	*le, a1*	–

8 word classes were further grouped into three word categories: common nouns, predicates (main verbs and adjectives), and closed-class words (pronouns, prepositions, conjunctions, quantifiers, and question words) based on previous studies (Bates et al., 1994; Hao et al., 2015; Liu & Chen, 2015) for subsequent analyses of lexical composition. (see Table 1)

Prior research indicated that definitions of nouns and verbs affected the presence and extent of a noun or verb bias (Tardif, 1996). To permit comparison with previous studies, two methods of counting nouns and verbs were used to examine noun or verb dominance in Mandarin-speaking children in the study. One used a stricter definition, considering only common nouns as nouns and main verbs as verbs. The other used a broader definition; that is, nouns included common nouns and proper names, and verbs included all verbs (main verbs, copular verbs, modal auxiliary verbs, directional verbs, and resultative verbs).

Furthermore, CLAN programs were used to extract the core lexicons of children's language production and examine the word categories in which the core lexicons are distributed. Through these analyses, we sought to determine which words and word categories are shared by the majority of the children and to investigate their developmental patterns. The core lexicons were defined as the lexical items used by over 30 children (75%) in the present study.

Statistical analyses were conducted using IBM SPSS software (version 23.0; IBM Corp., USA). The distribution of data was evaluated using the Kolmogorov-Smirnov test for normality, which showed a non-normal distribution of most of the obtained results. Therefore, the non-parametric Friedman test was used to examine the changes in the total number of utterances, word tokens (total number of words), word types (total number of different words), MLU (mean length of utterance), frequencies, and proportions of word types in 12 word classes and three word categories over time. If there was a statistically significant difference, the Wilcoxon test with Bonferroni correction ($p = .05/3 = .017$) was

used for post hoc analysis. The Wilcoxon test was also used to examine whether there was a significant difference between the proportions of nouns and verbs in the analyses of noun or verb dominance under different counting methods. A chi-square test was used to compare the differences in the numbers of children who demonstrated different patterns of dominance.

Results

Children's early language production and lexical composition

First, we analyzed children's basic language measures during play with their parents. The descriptive statistics and Friedman test results for the total number of utterances, word tokens, word types, and MLU at 1;8, 2;2, and 3;0 are displayed in Table 2. As shown, there were significant age effects on the total number of utterances ($\chi^2(2, n = 40) = 37.665, p < .001$), word tokens ($\chi^2(2, n = 40) = 54.950, p < .001$), word types ($\chi^2(2, n = 40) = 67.200, p < .001$), and MLU ($\chi^2(2, n = 40) = 68.600, p < .001$). Furthermore, post-hoc analyses revealed significant increases in both lexical and grammatical indices, including word tokens, word types, and MLU, between ages 1;8 and 2;2 as well as between ages 2;2 and 3;0.

Second, we analyzed the frequencies and proportions of word types in each of the 12 word classes across age to depict the developmental pattern of lexical composition in Mandarin-speaking children. The proportion is the total number of word types in each word class divided by the total number of all word types (e.g., proportion of verbs = verb types/all word types). Regarding the frequencies of the 12 word classes, the frequencies of all word classes except quantifiers ($\chi^2(2, n = 40) = 4.000, p > .05$) increased significantly with age. In addition, post-hoc analyses revealed that all word classes except common nouns, adjectives, and quantifiers increased significantly between ages 1;8 and 2;2 as well as between ages 2;2 and 3;0, whereas common nouns and adjectives only showed a significant increase between ages 1;8 and 2;2 (common nouns: $Z = -4.288, p < .001$; adjectives: $Z = -4.994, p < .001$), but not between ages 2;2 and 3;0 (common nouns: $Z = -1.494, p > .05$; adjectives: $Z = -1.820, p > .05$) (see Table 3).

Regarding the proportions of the 12 word classes, at 1;8, common nouns and main verbs, which were the two largest classes, accounted for 20.3% and 15.8% of the children's total production vocabulary, respectively. The median proportions of adjectives and adverbs were 4.7% and 2.5%, respectively. As the children grew, the proportion of common nouns decreased significantly ($\chi^2(2, n = 40) = 6.453, p < .05$), while the other word classes except adjectives and quantifiers increased significantly with age. The proportions of quantifiers and adjectives increased and decreased with age, respectively, but the changes did not reach statistical significance (quantifiers: $\chi^2(2, n = 40) = 4.000, p > .05$; adjectives: $\chi^2(2, n = 40) = 2.730, p > .05$) (see Table 4).

Furthermore, the distributions of the proportions of the three word categories were also examined. Common nouns and predicates were the two largest categories at 1;8, accounting for 20.3% and 20.0% of the total productive vocabulary, respectively. Compared to common nouns and predicates, closed-class words were less frequent at 1;8. As the children grew, the proportion of common nouns decreased significantly with age ($\chi^2(2, n = 40) = 6.453, p < .05$), while the proportion of closed-class words increased ($\chi^2(2, n = 40) = 54.650, p < .001$). In contrast to the aforementioned word classes, which changed significantly with age, the median proportion of predicates remained at roughly 20% - 27 %, with less change ($\chi^2(2, n = 40) = 2.450, p > .05$) (see Figure 1).

Table 2. Descriptive Statistics and Friedman Test Results for Early Language Production at 1;8 (Time 1), 2;2 (Time 2), and 3;0 (Time 3)

	1;8 (Time 1)			2;2 (Time 2)			3;0 (Time 3)			Friedman test	
	M (SD)	Mdn	Range	M (SD)	Mdn	Range	M (SD)	Mdn	Range	(2, n = 40)	Post-hoc[a]
Total utterances	37.5 (27.0)	34	1–115	89.0 (45.5)	82.5	15-206	92.3 (50.9)	74.5	25-231	37.665***	T1 < T2, T3
Word tokens	54.3 (44.1)	48	1–192	191.4 (116.7)	162	29-479	286.7 (184.2)	229	63-938	54.950***	T1 < T2 < T3
Word types	24.4 (18.4)	18.5	1–87	70.7 (33.5)	66	20-160	100.9 (42.9)	93.5	39-223	67.200***	T1 < T2 < T3
MLU	1.4 (0.3)	1.4	1.0–2.1	2.1 (0.5)	2.1	1.3-3.2	3.1 (0.8)	3.0	1.8-5.3	68.600***	T1 < T2 < T3

[a]Post-hoc analysis was performed by the Wilcoxon test with Bonferroni correction, with $p < .017$ being considered statistically significant for multiple comparison.

***$p < .001$.

Table 3. Descriptive Statistics and Friedman Test Results for Frequencies of All Word Classes at 1;8 (Time 1), 2;2 (Time 2), and 3;0 (Time 3)

	1;8 (Time 1)			2;2 (Time 2)			3;0 (Time 3)			Friedman test	Post-hoc[a]
	M (SD)	Mdn	Range	M (SD)	Mdn	Range	M (SD)	Mdn	Range	(2, n = 40)	
Common nouns	5.2 (4.8)	3	0–17	11.5 (7.4)	10	0-37	13.9 (7.1)	12.5	1–31	32.726***	T1 < T2, T3
Main verbs	4.5 (4.7)	3	0–18	14.7 (9.4)	14	2–43	22.8 (11.6)	21.5	6–54	55.736***	T1 < T2 < T3
Adjectives	1.3 (1.2)	1	0–4	4.3 (2.6)	3.5	0–11	5.2 (3.3)	5	1–15	38.356***	T1 < T2, T3
Pronouns	0.5 (0.8)	0	0–3	3.2 (1.7)	3	0–7	5.2 (1.7)	5	1–9	63.216***	T1 < T2 < T3
Prepositions	0.2 (0.6)	0	0–3	1.3 (1.6)	1	0–6	2.3 (1.7)	2	0–6	47.230***	T1 < T2 < T3
Conjunctions	0.1 (0.2)	0	0–1	0.6 (0.8)	0	0–2	2.1 (1.5)	2	1–5	46.846***	T1 < T2 < T3
Quantifiers	0.0 (0.0)	0	0–0	0.0 (0.0)	0	0–0	0.1 (0.2)	0	0–1	4.000	–
Question words	0.1 (0.3)	0	0–1	1.2 (1.0)	1	0–3	2.3 (1.2)	2	0–4	50.130***	T1 < T2 < T3
Adverbs	1.1 (1.3)	1	0–5	6.8 (4.0)	6	0–17	11.9 (6.2)	11.5	0–31	61.961***	T1 < T2 < T3
Numbers	0.4 (1.6)	0	0–10	1.7 (2.6)	1	0–15	4.6 (4.1)	3	0–13	49.636***	T1 < T2 < T3
Classifiers	0.6 (0.7)	0	0–3	1.5 (0.9)	1	0–4	2.9 (1.6)	2	1–7	51.179***	T1 < T2 < T3
Sentence-final particles	0.5 (0.9)	0	0–3	2.3 (1.6)	2	0–5	3.2 (1.3)	3	1–6	50.324***	T1 < T2 < T3

Note. Some words classes, such as onomatopoeia, interjections and kinship terms, were not included.
[a]Post-hoc analysis was performed by the Wilcoxon test with Bonferroni correction, with $p < .017$ being considered statistically significant for multiple comparison.
***$p < .001$.

Table 4. Descriptive Statistics and Friedman Test Results for Proportions of All Word Classes at 1;8 (Time 1), 2;2 (Time 2), and 3;0 (Time 3)

	1;8 (Time 1)			2;2 (Time 2)			3;0 (Time 3)			Friedman test	
	M (SD)	Mdn	Range	M (SD)	Mdn	Range	M (SD)	Mdn	Range	(2, n=40)	Post-hoc[a]
Common nouns	20.8 (16.7)	20.3	0–62.5	15.8 (6.5)	15.0	0–35.1	13.4 (4.0)	13.9	2.4–22.2	6.453*	–
Main verbs	17.5 (16.4)	15.8	0–100.0	19.5 (5.2)	19.5	9.1–29.0	22.3 (4.8)	22.6	12.7–34.6	8.150*	T1, T2 < T3
Adjectives	7.2 (15.6)	4.7	0–100.0	5.7 (2.7)	5.1	0–12.5	5.0 (2.0)	4.9	1.1–11.0	2.730	–
Pronouns	1.4 (2.6)	0.0	0–11.8	4.6 (2.4)	4.4	0–15.0	5.6 (2.2)	5.1	2.2–13.7	40.850***	T1 < T2, T3
Prepositions	0.3 (0.9)	0.0	0–3.4	1.5 (1.6)	1.4	0–7.1	2.2 (1.4)	2.1	0–4.8	27.924***	T1 < T2 < T3
Conjunctions	0.2 (0.9)	0.0	0–5.3	0.7 (0.9)	0	0–2.9	2.0 (1.1)	2.2	0–4.9	38.262***	T1 < T2 < T3
Quantifiers	0.0 (0.0)	0.0	0–0	0.0 (0.0)	0.0	0–0	0.0 (0.1)	0.0	0–1.1	4.000	–
Question words	0.4 (1.4)	0.0	0–5.9	1.6 (1.4)	1.5	0–5.0	2.5 (1.5)	2.3	0–7.8	30.662***	T1 < T2, T3
Adverbs	3.3 (3.7)	2.5	0–10.7	9.5 (3.5)	9.0	0–16.7	11.6 (3.0)	11.5	5.5–20.0	46.350***	T1 < T2 < T3
Numbers	0.8 (3.1)	0.0	0–19.2	2.8 (6.3)	1.4	0–39.5	4.2 (3.7)	3.3	0–14.7	32.723***	T1 < T2 < T3
Classifiers	2.0 (2.7)	0.0	0–8.2	2.3 (1.8)	1.9	0–9.1	2.9 (1.2)	2.6	0.6–5.5	6.453*	T1, T2 < T3
Sentence-final particles	1.8 (3.0)	0.0	0–11.1	3.3 (2.1)	3.3	0–8.3	3.4 (1.4)	3.2	1.0–7.7	15.241***	T1 < T2, T3

Note. Some words classes, such as onomatopoeia, interjections and kinship terms, were not included. [a]Post-hoc analysis was performed by the Wilcoxon test with Bonferroni correction, with $p < .017$ being considered statistically significant for multiple comparison.
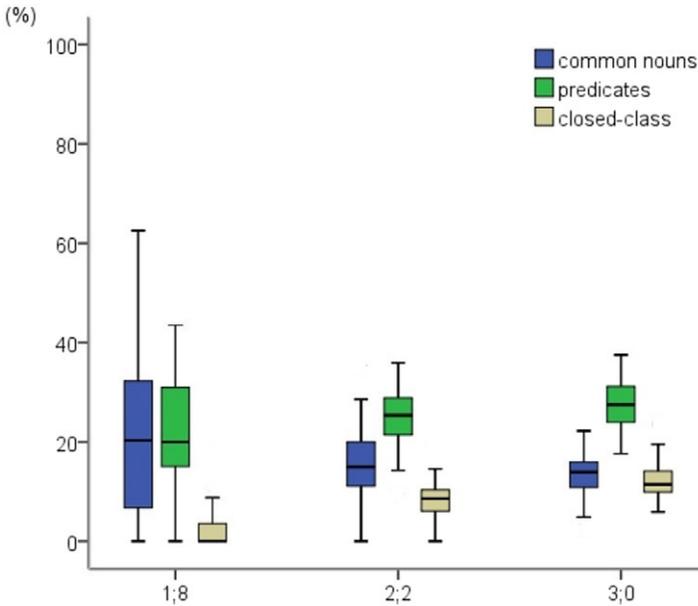*$p < .05$,

***$p < .001$.

Figure 1. Proportions of three word categories at 1;8 (Time 1), 2;2 (Time 2), and 3;0 (Time 3).

Noun and verb dominance in Mandarin-speaking children across age

We examined whether there was noun or verb dominance in Mandarin-speaking children between ages 1;8 and 3;0 in the toy play context. As shown in Table 5, regardless of whether one uses a stricter or a broader definition of nouns and verbs, Mandarin-speaking children used slightly more nouns than verbs at 1;8, but this did not reach statistical significance (stricter definition: $Z = - 1.394$, $p > .05$; broader definition: $Z = - 0.248$, $p > .05$). However, as the children grew, the proportion of verbs was significantly higher than that of nouns at 2;2 (stricter definition: $Z = - 3.047$, $p < .01$; broader definition: $Z = - 4.731$, $p < .001$) and 3;0 (stricter definition: $Z = - 5.135$, $p < .001$; broader definition: $Z = - 5.511$, $p < .001$) (see Figure 2).

Furthermore, we examined whether there were significant differences in the numbers of noun-dominant, verb-dominant, and balanced children. Under the stricter definition of nouns and verbs, including only common nouns and main verbs, the number of children who produced more nouns than verbs was significantly larger than of those who produced more verbs than nouns and of those who produced equal nouns and verbs at 1;8 ($\chi^2(2, n = 40) = 6.350$, $p < .05$). However, there was an inverse pattern at 2;2 and 3;0, such that the number of children who produced more verbs than nouns was significantly larger than of those who produced more nouns than verbs and of those who produced equal numbers of nouns and verbs at 2;2 and 3;0 (at 2;2: $\chi^2(2, n = 40) = 21.350$, $p < .001$; at 3;0: $\chi^2(2, n = 40) = 25.600$, $p < .001$). Similarly, the number of children who produced more verbs than nouns was significantly larger than of those who produced more nouns than verbs and of those who produced an equal number of nouns and verbs at 1;8, 2;2, and 3;0 under the broader definition of nouns and verbs (at 1;8: $\chi^2(2, n = 40) = 6.650$, $p < .05$; at 2;2: $\chi^2(2, n = 40) = 22.500$, $p < .001$) (see Figure 3).

Table 5. Proportions of Nouns and Verbs in Different Counting Methods at 1;8 (Time 1), 2;2 (Time 2), and 3;0 (Time 3)

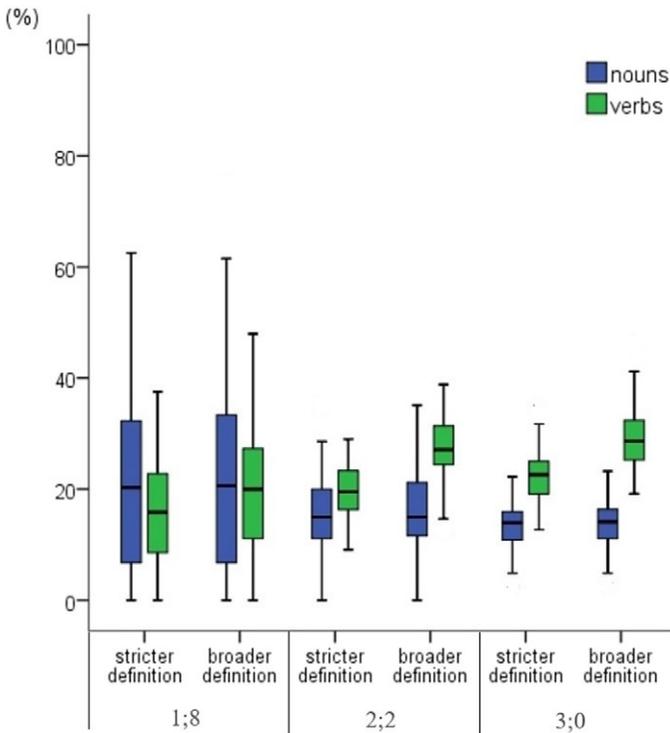| Counting method		1;8 (Time 1)			2;2 (Time 2)			3;0 (Time 3)		
		M	Mdn	Z	M	Mdn	Z	M	Mdn	Z
		(SD)	(Range)		(SD)	(Range)		(SD)	(Range)	
Stricter definition	Common nouns	20.8	20.3	−1.394	15.8	15.0	−3.047**	13.4	13.9	−5.135***
		(16.7)	(0–62.5)		(6.5)	(0–35.1)		(4.0)	(2.4–22.2)	
	Main verbs	17.5	15.8		19.5	19.5		22.3	22.6	
		(16.4)	(0–100)		(5.2)	(9.1–29)		(4.8)	(12.7–34.6)	
Broader definition	Common nouns + proper names	21.7	20.6	−0.248	16.4	15.0	−4.731***	13.9	14.1	−5.511***
		(17.7)	(0–75)		(6.7)	(0–35.1)		(4.2)	(2.4–23.2)	
	Total verbs	20.9	20.0		27.1	27.1		29.9	28.6	
		(17.2)	(0–100)		(6.3)	(12.7–38.8)		(7.1)	(19.2–46.3)	

**p < .01,
***p < .001

Figure 2. Proportions of nouns and verbs under different counting methods at 1;8 (Time 1), 2;2 (Time 2), and 3;0 (Time 3).

Core lexicons of children's language production in the toy play context

We extracted the core lexicons of children's language production and examined which words and lexical categories were shared by the majority of the children in the toy play context. As shown in Table 6, at 1;8, there were no words used by more than 30 children. As the children grew, their core lexicons mainly comprised predicates and closed-class words. In addition, some words that did not belong to the three main lexical categories, such as *bu2yao4* 'no' and *bu4* 'not' among negative words, *ge4* among classifiers, and *le* and *a1* among sentence-final particles, were found to be shared by the majority of the children in the toy play context at 2;2 and 3;0.

Discussion

The developmental changes of early lexical composition

The first goal of the present study was to examine the developmental pattern of early lexical composition in Mandarin-speaking children by analyzing how words were distributed across the three word categories and across the 12 word classes. The results showed that at 1;8, the majority of children's lexicons comprised common nouns and predicates, with closed-class words being less frequent. As the children grew, the proportion of predicates predominated over that of common nouns when the children were
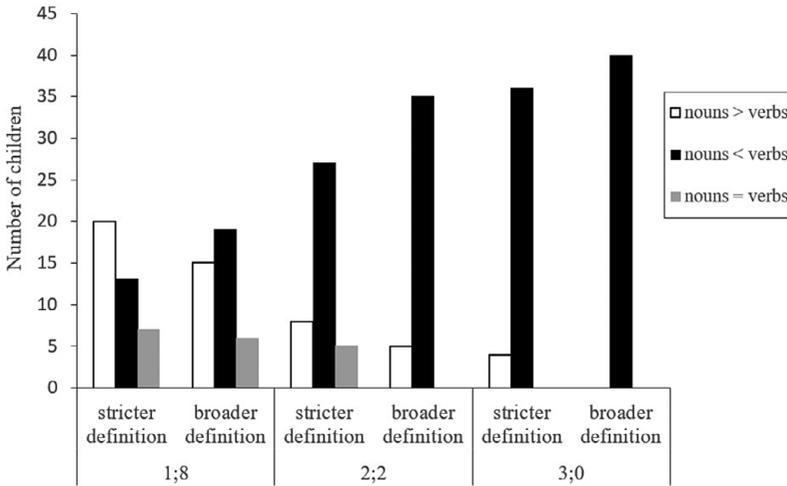
Figure 3. Numbers of noun-dominant, verb-dominant, and balanced children in lexical composition under different counting methods at 1;8 (Time 1), 2;2 (Time 2), and 3;0 (Time 3).

Table 6. Core Lexicons of Children's Language Production at 1;8 (Time 1), 2;2 (Time 2), and 3;0 (Time 3)

	1;8 (Time 1)	2;2 (Time 2)	3;0 (Time 3)
Common nouns	–	–	–
Predicates	–	*hao3* (good)	*hao3* (good)
			wan2 (play)
			yao4 (want)
Closed-class words	–	*zhe4* (this)	*zhe4* (this)
		wo3 (I)	*wo3* (I)
			ni3 (you)
			shen2me (what)
Other	–	*ge4* (classifier)	*ge4* (classifier)
		le (particle)	*le* (particle)
		bu2yao4 (no)	*bu2yao4* (no)
		en1 (okay)	*en1* (okay)
		shi4 (is)	*shi4* (is)
		bu4 (not)	*bu4* (not)
			de (particle)
			mei2you3 (not)
			a1 (particle)
			yi1 (one)

2;2 and 3;0. Previous studies of lexical composition using parental reports demonstrated that the proportion of common nouns is higher than that of predicates, which is inconsistent with our findings (Bates et al., 1994; Caselli et al., 1999; Hao et al., 2015; Liu & Chen, 2015).

The discrepancy between findings of this study and those of prior research might lie in three factors, i.e., data collection methods, activity contexts, and language systems. Studies based on observational and parent-report measures of lexical composition tend to generate different results. For example, Bates et al. (1994) analyzed the lexical composition of English-speaking children between ages 1;4 and 2;6 using parental reports and found that common nouns predominated over predicates. Similar findings were also found in studies of the lexical composition of Mandarin-speaking children based on parental reports (Hao et al., 2015; Liu & Chen, 2015).

However, past studies of early lexical development based on spontaneous speech demonstrated a more even distribution between nouns and verbs, and some of them even found that children produce more verbs than nouns at an early stage (Bassano, 2000; Ogura et al., 2006; Tardif, 1996; Tardif et al., 1999). Pine et al. (1996) investigated the relationship between observational and parent-report measures of lexical composition. They found in parent-report studies that common nouns make up over 50% of the words in the lexicons of 80% of children; however, only around 20% of children showed common nouns making up over 50% of the words in observational measures, which supports the differences between observational and parent-report measures of lexical composition. Tardif et al. (1999) examined the consistency and relative biases of observational and parent-report measures in English- and Mandarin-speaking children. Their findings revealed that for both the English- and Mandarin-speaking children, the proportion of nouns that were produced by children and reported by their parents was higher than that of verbs; in addition, the proportion of verbs that were produced by children but not reported by parents was higher than nouns. The differences in memory accuracy for different parts-of-speech may be a reason for the discrepancy between observational and parent-report measures of lexical composition. Furthermore, compared to parental reports, spontaneous speech samples concurrently present semantic, syntactic, and pragmatic properties of children's language production, meaning that lexical composition based on spontaneous speech is more strongly influenced by the structure of sentences used by children. The properties of spontaneous speech may also create the differences between observational and parent-report measures of lexical composition.

The second possible explanation for this discrepancy may be the activity contexts. Previous studies have suggested that activity contexts affect several aspects of children's speech (Hoff-Ginsberg, 1991; Lucariello & Nelson, 1986; Ogura et al., 2006; Tardif et al., 1999). Tardif et al. (1999) and Ogura et al. (2006) indicated that children use more verbs in the toy play context than in the book-reading context. The present study collected spontaneous speech samples in the toy play context, which may be the reason for the higher proportion of predicates.

The third possible explanation for this discrepancy may be the language systems. Mandarin is a pro-drop language; therefore, nouns are frequently omitted. Hsu (1996) investigated early language development in Mandarin-speaking children and found that they produce combined words and simple sentences at around 1;5 to 1;10. In his study, most of the children at the age of 1;5 were able to combine words to form phrases and simple sentences, and most of the structures were imperatives or verb phrases with the subject deleted. Moreover, Mandarin verbs do not carry tense markings and do not have subject-verb agreement. These morphosyntactic properties of Mandarin verbs make

them easier for Mandarin-speaking children to acquire or use. Tardif et al. (1999) compared the proportions of nouns and verbs in the early lexicons of English- and Mandarin-speaking toddlers and found that Mandarin-speaking children used fewer nouns and more verbs than English-speaking children, which confirmed that language systems may be an influencing factor of lexical composition.

Regarding developmental changes in lexical composition, the results demonstrated that the proportion of common nouns decreased significantly with age, and the proportions of predicates and closed-class words increased, which was similar to the developmental trajectories found in other languages (Bassano et al., 2005; Bates et al., 1994; Caselli et al., 1999; Kauschke & Hofmeister, 2002). For example, Bates et al. (1994) analyzed the lexical composition of English-speaking children using parental reports and found that the proportion of common nouns increased from 0 to 200 words, followed by a proportional decrease after 200 words. In Bates et al.'s study, the median number of words attained by children at 1;8 is around 170, and our study on children aged between 1;8 and 3;0 demonstrates that the proportion of common nouns decreased with age, which is consistent with Bates et al.'s (1994) findings. In addition, Bates et al. (1994) revealed that predicates showed a steady increase, while closed-class words grew markedly after 400 words, which indicates that the development of closed-class words may depend on certain critical masses of nouns, verbs, and other content words. Similarly, previous studies based on spontaneous speech have also demonstrated that the proportion of nouns decreases, while the proportions of predicates and grammatical function words increase in early lexical acquisition (Bassano et al., 2005; Kauschke & Hofmeister, 2002).

Furthermore, the present study found that Mandarin-speaking children used prepositions, conjunctions, question words, classifiers, and sentence-final particles as early as 1;8, and that the use of these word classes increased significantly between ages 1;8 and 3;0. According to past research on grammatical development in Mandarin-speaking children, they first appear to use two-word combinations around one and a half to two years of age, followed by combining three or more words to compose short sentences, and then progressing to complex or compound sentences (Cheng, 1986; Chi, 2002; Hsu, 1996; Liu & Tsao, 2010). Thus, the use of grammatical function words, such as conjunctions and prepositions, may increase as the length and complexity of sentences increase. The present study showed that the median MLU of children's language production increased from 1.4 to 3.0 between ages 1;8 and 3;0, which indicates that the children used longer and more complex sentences with age. This developmental process from the single-word stage to the syntactic stage may increase the use of grammatical words and induce changes in lexical composition. Children's early language development is not only at the level of vocabulary but also at that of grammar, which advances with age. Lexical and grammatical development are two important components of early childhood language development, and they are closely related and not easily separable. Naturalistic speech samples may provide more information about the links between lexicon and grammar.

The use of nouns and verbs in the toy play context across age

The second goal of the present study was to investigate whether noun or verb dominance exists in Mandarin-speaking children during the period between 1;8 and 3;0. The results indicated that children tended to use more nouns than verbs at 1;8. However, as the children grew, they used significantly more verbs than nouns in toy play at 2;2 and 3;0.

The stage of grammatical development may influence the use of nouns and verbs. Bassano (2000) investigated the development of noun and verb word classes in the free speech of a French child between the ages of 1;2 and 2;6, and found that nouns significantly predominated over verbs in the child's productions both in types and tokens until 1;8. But, from around 2;0 onward, verb types equaled noun types and verb tokens surpassed noun tokens in the child's productions. Similar results were found by Ogura et al. (2006), who analyzed children's use of nouns and verbs in the spontaneous speech of Japanese-speaking children and their caregivers, and found that there was a shift from noun dominance to verb dominance in the toy play context from the single-word stage to the syntactic stage. The present study is consistent with Bassano's (2000) and Ogura et al.'s (2006) studies, suggesting that noun or verb dominance may change with the stage of grammatical development.

Core lexicons of children's language production in the toy play context

The third goal of the present study was to examine which words and lexical categories are shared by the majority of children. The results showed that there were no words used by more than 30 in the toy play context at 1;8. Children's core lexicons were mainly composed of predicates and closed-class words when the children were 2;2 and 3;0. A previous study investigated communicative acts during toy play between mothers and children using the Inventory of Communicative Acts-Abridged (INCA-A), and found that statements and responses as well as questions and responses were the two most common speech acts used by Mandarin-speaking toddlers during toy play (Lee, 2019). In parent-child interaction during play, children commonly used words such as *bu2yao4* 'no' and *yao4* 'want' to respond to parents' questions, or used words such as *zhe4* 'this', *shi4* 'is', and *wo3* 'I' in simple declarative sentences, which may explain why predicates and closed-class words were shared by a majority of children in the present study. Closed-class words were used more commonly by the majority of the children in early language production in the toy play context, and the proportion of closed-class words increased significantly between 1;8 and 3;0. These findings demonstrate that the role of grammatical function words increases with the development of children's syntactic abilities.

Furthermore, some words that did not belong to the three lexical categories, such as adverbs, classifiers, and sentence-finial particles, were found to be shared by the majority of the children at 2;2 and 3;0 in the toy play context.

Classifiers are a special part of Mandarin. Mandarin classifiers usually occur between a numeral and a noun or between a demonstrative word and a noun (Chao, 1968; Li & Thompson, 2009; Tai & Wang, 1990; Tien, Tzeng, & Hung, 2002). The development of classifiers not only reflects the linguistic abilities of children, but also their cognitive foundation, such as conceptual structure and categorization (Tien et al., 2002). Therefore, classifiers have received much attention in Chinese language development. The present study showed that classifier *ge4* is shared by the majority of the children at 2;2 and 3;0, which is consistent with Hsu's (1996) finding that *ge4* is the most popular and the first acquired classifier in Mandarin-speaking children. In addition, sentence-final particles are widely used in spoken Mandarin to convey speakers' various feelings, emotions, and attitudes, which play an important semantic and pragmatic role in spoken Mandarin discourse (Li & Thompson, 2009). The present study also found that sentence-final particles such as *le* and *a1* were learned by Mandarin-speaking children quite early and were used by the majority of the children at 2;2 and 3;0.

Limitations and suggestions for future research

Although we investigated longitudinal changes in Mandarin-speaking children's early lexical production and composition based on longitudinal language samples that provide important information about the developmental patterns of early language acquisition in Mandarin-speaking children, the present study still has certain limitations. First, the sample size is relatively small, and the data are drawn exclusively from language performance during 15 minutes of parent-child interactions in the toy play context, which allows limited generalizability. Second, all participating children were from middle SES families; therefore, our findings may not be applicable to children from other SES groups. Third, the present study only included children between ages 1;8 and 3;0. However, the lexical and grammatical abilities of children continue to develop after the age of three years, which may cause more complex changes in lexical composition. We recommend that future studies collect spontaneous speech from larger sample sizes, in more varied contexts, and with a longer collection duration, which would increase the representativeness of language samples. In addition, further longitudinal research may include older children, or analyze the relationships between vocabulary and grammar using more complete and detailed grammatical indicators, which would aid in acquiring broader range of information about early language development.

Conclusion

In summary, the present study suggests that common nouns and predicates are dominant categories in Mandarin-speaking children in the early stages of lexical development. With vocabulary growth and the emergence of grammar, different lexical categories show different developmental patterns. A decrease in the proportion of common nouns and an increase in the proportion of predicates and closed-class words found in Mandarin-speaking children in the present study are similar to the developmental trajectories in English, Italian, and French. Furthermore, closed-class words are shared by the majority of children at 2;2 and 3;0, which demonstrates that the role played by grammatical function words gradually increases with age. These longitudinal data provide important information about the developmental trends of early lexical composition in Mandarin-speaking children, providing a valuable contrast for the examination of children's language in different contexts and language systems.

References

Bassano, D. (2000). Early development of nouns and verbs in French: Exploring the interface between lexicon and grammar. *Journal of Child Language*, **27**(3), 521–559. doi: 10.1017/S0305000900004396

Bassano, D., Eme, P., & Champaud, C. (2005). A naturalistic study of lexical development: General processes and interindividual variations in French children. *First Language*, **25**(1), 67–101. doi: 10.1177/0142723705049119

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**(1), 85–123. doi: 10.1017/s0305000900008680

Bleses, D., Makransky, G., Dale, P. S., Højen, A., & Ari, B. A. (2016). Early productive vocabulary predicts academic achievement 10 years later. *Applied Psycholinguistics*, **37**(6), 1461–1476. doi: 10.1017/S0142716416000060

Bornstein, M. H., Cote, L. R., Maital, S., Painter, K., Park, S. Y., Pascual, L., Pêcheux, M., Ruel, J., Venuti, P., & Vyt, A. (2004). Cross-linguistic analysis of vocabulary in young children: Spanish, Dutch, French, Hebrew, Italian, Korean, and American English. *Child Development*, **75**(4), 1115–1139. doi: 10.1111/j.1467-8624.2004.00729.x

Cadime, I., Moreira, C., Santos, A. L., Silva, C., Ribeiro, I., & Viana, F. L. (2019). The development of vocabulary and grammar: A longitudinal study of European Portuguese-speaking toddlers. *Journal of Child Language*, **46**(4), 653–681. doi: 10.1017/S0305000919000060

Caselli, M. C., Bates, E., Casadio, P., Fenson, J., Fenson, L., Sanderl, L., & Weir, J. (1995). A crosslinguistic study of early lexical development. *Cognitive Development*, **10**(2), 159–200. doi: 10.1016/0885-2014(95)90008-X

Caselli, 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**(1), 69–111. doi: 10.1017/s0305000998003687

Chao, Y. R. (1968). *A grammar of spoken Chinese*. Berkeley, CA: University of California Press.

Cheng, S. (1986). The earliest grammatical rules in Chinese toddlers and their probable developing process. *Chinese Journal of Psychology*, **28**(2), 93–122.

Cheung, H., Chang, C., Ko, H.-W., & Tsay, J. (2011). Taiwan Child Language Data Exchange System (Report no. NSC96-2420-H-002-030).

Chi, P. (2002). The development of language and communication abilities of Mandarin Chinese-speaking infants and toddlers: Parental report. *Journal of Special Education*, **16**, 23–64. doi: 10.6768/JSE.200209.0023

Choi, S., & Gopnik, A. (1995). Early acquisition of verbs in Korean: A cross-linguistic study. *Journal of Child Language*, **22**(3), 497–529. doi: 10.1017/S0305000900009934

Day, T., & Elison, J. T. (2022). A broadened estimate of syntactic and lexical ability from the MB-CDI. *Journal of Child Language*, **49**(3), 615–632. doi:10.1017/S0305000 921000283

Dixon, J. A., & Marchman, V. A. (2007). Grammar and the lexicon developmental ordering in language acquisition. *Child Development*, **78**(1), 190–212. doi: 10.1111/j.1467-8624.2007.00992.x

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**(2), 351–372. doi: 10.1017/S0305000901004676

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

Goldfield, B., & Reznick, J. S. (1990). Early lexical acquisition: rate, content, and the vocabulary spurt. *Journal of Child Language*, **17**(1), 171–183. doi: 10.1017/s0305000900013167

Hao, M., Shu, H., Xing, A., & Li, P. (2008). Early vocabulary inventory for Mandarin Chinese. *Behavior Research Methods*, **40**(3), 728–733. doi: 10.3758/BRM.40.3.728

Hao, M., Liu, Y., Shu, H., Xing, A., Jiang, Y., & Li, P. (2015). Developmental changes in the early child lexicon in Mandarin Chinese. *Journal of Child Language*, **42**(3), 505–537. doi: 10.1017/S030500091400018X

Hoff-Ginsberg, E. (1991). Mother-child conversation in different social classes and communicative setting. *Child Development*, **62**(4), 782–796. doi: 10.2307/1131177

Hsu, J. H. (1996). *A study of the stages of development and acquisition of Mandarin Chinese by children in Taiwan*. Taipei, Taiwan: The Crane Publishing Company.

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**(4), 735–757. doi: 10.1017/s0305000902005330

Lee, C. (2019). Communicative acts during toy play between mother and children [Unpublished doctoral dissertation]. National Taiwan Normal University, Taipei, Taiwan.

Lee, J. (2011). Size matters: Early vocabulary as a predictor of language and literacy competence. *Applied Psycholinguistics*, **32**(1), 69–92. doi: 10.1017/S0142716410000299

Li, C. N., & **Thompson, S. A.** (2009). *Mandarin Chinese: A functional reference grammar*. Taipei, Taiwan: The Crane Publishing Company.

Liu, H., & **Chen, Y.** (2015). Developmental changes in the content and composition of early expressive vocabulary in Mandarin-speaking infants and toddlers. *Bulletin of Educational Psychology*, **47**(2), 217–242.

Liu, H., & **Tsao, F.** (2010). The standardization and application of Mandarin-Chinese Communicative Developmental Inventory for infants and toddlers. *Formosa Journal of Mental Health*, **23**(4), 503–534. doi: 10.30074/FJMH.201012_23(4).0001

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

Lucariello, J., & **Nelson, K.** (1986). Context effects on lexical specificity in maternal and child discourse. *Journal of Child Language*, **13**(3), 507–522. doi: 10.1017/S0305000900006851

Lyytinen, P., **Laakso, M. L.**, **Poikkeus, A.**, & **Rita, N.** (1999). The development and predictive relations of play and language across the second year. *Scandinavian Journal of Psychology*, **40**(3), 177–186. doi: 10.1111/1467-9450.00115

MacWhinney, B. (2000). *The CHILDES project: Tools for analyzing talk (3rd ed.)*. Mahwah, NJ: Lawrence Erlbaum.

MacWhinney, B., & **Snow, C. E.** (1990). The Child Language Data Exchange System: An update. *Journal of Child Language*, **17**(2), 457–472. doi: 10.1017/S0305000900013866

Nelson, K. (1973). Structure and strategy in learning to talk. *Monographs of the Society for Research in Child Development*, **38**(Serial Nos.1-2), 1–136.

Ogura, T., **Dale, P. S.**, **Yamashita, Y.**, **Murase, T.**, & **Mahieu, A.** (2006). The use of nouns and verbs by Japanese children and their caregivers in book-reading and toy-playing contexts. *Journal of Child Language*, **33**(1), 1–29. doi: 10.1017/s0305000905007270

Parisse, C., & **Le Normand, M. T.** (2000). How children build their morphosyntax: The case of French. *Journal of Child Language*, **27**(2), 267–292. doi: 10.1017/s0305000900004116

Pérez-Pereira, M., & **Cruz, R.** (2018). A longitudinal study of vocabulary size and composition in low risk preterm children. *First Language*, **38**(1), 72–94. doi: 10.1177/014272317730484

Pine, J. M., **Lieven, E. V. M.**, & **Rowland, C.** (1996). Observational and checklist measures of vocabulary composition: What do they mean? *Journal of Child Language*, **23**(3), 573–590. doi: 10.1017/S030500900008953

Rescorla, L., **Roberts, J.**, & **Dahlsgaard, K.** (1997). Late talkers at 2: Outcome at age 3. *Journal of Speech, Language, and Hearing Research*, **40**(3), 556–566. doi: 10.1044/jslhr.4003.556

Rowe, M. L., **Raudenbush, S.**, & **Goldin-Meadow, S.** (2012). The pace vocabulary growth helps predict later vocabulary skill. *Child Development*, **83**(2), 508–525. doi: 10.1111/j.1467-8624.2011.01710.x

Southwood, F., & **Russell, A. F.** (2004). Comparison of conversation, freeplay, and story generation as methods of language sample elicitation. *Journal of Speech, Language, and Hearing Research*, **47**(2), 366–376. doi: 10.1044/1092-4388(2004/030)

Tai, J. H., & **Wang, L.** (1990). A semantic study of classifier tiao. *Journal of the Chinese Language Teachers Association*, 25(1), **25**-56.

Tardif, T. (1996). Nouns are not always learned before verbs: Evidence from Mandarin speaker's early vocabularies. *Developmental Psychology*, **32**(3), 492–504. doi: 10.1037/0012-1649.32.3.492

Tardif, T., **Fletcher, P.**, **Liang, W.**, & **Kaciroti, N.** (2009). Early vocabulary development in Mandarin (Putonghua) and Cantonese. *Journal of Child Language*, **36**(5), 1115–1144. doi: 10.1017/S0305000908009185

Tardif, T., **Fletcher, P.**, **Liang, W.**, **Zhang, Z.**, **Kaciroti, N.**, & **Marchman, V. A.** (2008). Baby's first 10 words. *Developmental Psychology*, **44**(4), 929–938. doi: 10.1037/0012-1649.44.4.929

Tardif, T., **Fletcher, P.**, **Liang, W.**, **Zhang, Z.**, & **Kaciroti, N.** (2009). Early vocabulary development in Mandarin (Putonghua) and Cantonese. *Journal of Child Language*, **36**(5), 1115–1144. doi: 10.1017/S0305000908009185

Tardif, T., **Gelman, S. A.**, & **Xu, F.** (1999). Putting the "noun bias" in context: A comparison of English and Mandarin. *Child Development*, **70**(3), 620–635.

Tardif, T., Shatz, M., & Naigles, L. (1997). Caregiver speech and children's use of noun versus verbs: A comparison of English, Italian, and Mandarin. *Journal of Child Language*, **24**(3), 535–565. doi: 10.1017/S030500099700319X

Tien, Y., Tzeng, O. J. L., & Hung, D. L. (2002). Semantic and cognitive basis of Chinese classifiers: A functional approach. *Language and Linguistics*, **3**(1), 101–132.

Trudeau, N., & Sutton, A. (2011). Expressive vocabulary and early grammar of 16- to 30-month-old children acquiring Quebec French. *First Language*, **31**(4), 480–507. doi: 10.1177/0142723711410828

Viana, F. L., Pérez-Pereira, M., Cadime, I., Silva, C., Santos, S., & Ribeiro, I. (2017). Lexical, morphological and syntactic development in toddlers between 16 and 30 months old: A comparison across European Portuguese and Galician. *First Language*, **37**(3), 285–300. doi: 10.1177/0142723717691087

Yang, J. (2015). *Measuring early vocabulary growth in Mandarin-speaking children: A corpus-based study* [Unpublished master's thesis]. National Taiwan Normal University, Taipei, Taiwan.