國立臺灣大學文學院語言學研究所

博士論文

Graduate Institute of Linguistics

College of Liberal Arts

National Taiwan University

Doctoral Dissertation

中文學齡前幼童的非詞覆誦、詞彙量及音韻能力在發展過程中之動態互動：跨序列研究

The dynamic interactions among nonword repetition, vocabulary size and phonological capacities in Mandarin-speaking preschoolers: A cross-sequential study

李乃欣

Naihsin Li

指導教授：張顯達博士；曹峰銘博士

Advisors: Hintat Cheung, Ph.D.; Feng-Ming Tsao, Ph.D.

中華民國 104 年 6 月

June, 2015

謝辭

這條博士班的路走來艱辛且漫長，常常覺得自己像大海裡飄蕩的小舟，知道自己想往哪裡去，卻總是看不到目的地。這份的論文的完成，就像長途飄搖後終於達到了一個目標。這個過程除了自己不斷的努力外，就像是「當你真心想要做一件事，全世界都會一起幫你」的體現。這些人一路上的扶持、勉勵或陪伴，是我能不斷前進的動力。

首先，感謝我的指導教授—張顯達老師及曹峰銘老師。張顯達老師是我在語言發展研究上的引領者。他的知識淵博，總是能在我迷惘時給予我方向與新的啟發。沒有他這幾年的指導，我相信這條路我會走得更艱困辛苦。曹峰銘老師除了給我許多研究及生涯規劃的寶貴建議外，也總是毫無保留的提供我所有的資源，讓我能心無旁騖的完成論文。

再來，也要感謝我的口試委員：馮怡蓁老師，胡潔芳老師以及許馨仁老師。他們在論文逐漸成形的這段過程中，用心地給予我許多仔細且具體的意見。

這本論文能夠完成，有很大部份要感謝願意參加研究的家長及小朋友，還有熱心的老師們，包括新北市慈霖幼兒園的黃淑貞園長，台北市上禾托兒所的李瑞茵所長以及台北市書宜幼稚園的李威立老師。這長達一年半的縱貫性研究，沒有他們的耐心配合及熱情參與是沒有辦法成功的。

此外，感謝培育我多年的台大語言所。所上師長們總是不吝於給我關懷、鼓勵及扶助，尤其是黃宣範老師，蘇以文老師，宋麗梅老師，馮怡蓁老師，呂佳蓉老師，美玲助教及劉姐。這份溫情我永遠銘記在心。還要感謝所上曾經幫助我或是為我加油打氣的所胞們，尤其是育穎，書珮，佳音，國樹等等。另外也要感謝台大心理系兒童語言實驗室的旅揚，霓姿，信慧及幼炘，在寫論文這段時間給予我的協助。

我的家人是我最堅強的後盾。謝謝我的公婆對於我選擇這份工作的尊重與支持。我的弟妹常以溫暖的親情伴我渡過許多低潮，也帶給我許多歡樂，感謝有他們做我的手足。我的丈夫志詠，沒有他的支持，奉獻及包容，我沒有辦法完成這

個夢想。

從我選擇走上學術這條路以來，我的爸爸是唯一不曾質疑這個選擇的人，總是無條件的支持我的夢想，並自始至終相信我可以做到。我的媽媽教我識字讀書，是我知識的啟蒙者，也是我對小小孩的興趣的啟發者。謹將此論文獻給我的父母。

摘要

本研究的目的在於探討非詞覆誦及語言知識（即辭彙量及音韻能力）在發展過程中之動態互動。首先，本研究提出一個工作模型，並預測隨著幼童詞彙量的增加，音韻能力及詞彙量會對於非詞覆誦表現有不同比重的貢獻及影響。此外，本研究也檢視非詞覆誦對於表達性詞彙發展的預測能力。此研究採用了跨序列實驗設計，追蹤三個不同年齡組的小孩（即兩歲、三歲及四歲）各一年的時間。小孩每半年接受一次測試，共受測三次。測試內容包含理解性詞彙，表達性詞彙，音韻口語輸出能力，詞彙區辨，及非詞覆誦。非詞覆誦包含了兩項次作業，分別為暫時詞覆誦（由中文現存音節所組合之非詞）及空缺詞覆誦（由中文之空缺音節所組合之非詞)。研究結果顯示幼童在非詞覆誦的表現隨年紀增長而進步，然而同齡孩童之間存在顯著的個體間差異。分析指出其個體間差異來自孩童的詞彙知識及音韻口語輸出能力之影響,而它們影響的程度則取決於詞彙量及非詞刺激材料的特性。隨著孩童詞彙量的增加，詞彙知識會支持並增進他們在非詞覆誦之表現。音韻口語輸出能力也會調節孩童在非詞覆誦的表現，但其效果發生在詞彙效應之後。當孩童達到相當的詞彙量，並皆會以詞彙資源來處理非詞時，他們的個體間差異則轉而取決於他們的音韻口語輸出能力。暫時詞覆誦作業及空缺詞覆誦作業在非詞刺激材料上的差異，則可反映不同語言層面在新詞處理時的作用。當兩者都受到音韻口語輸出能力的影響時，暫時詞覆誦會額外受到詞彙知識的影響。基於這些發現，本研究最後提出了一個修正模型來說明非詞覆誦的機制。同時本研究也證明孩童在非詞覆誦的表現可預測他們在表達性詞彙的發展。

關鍵詞：非詞覆誦，接受性詞彙，表達性詞彙，音韻口語輸出能力

Abstract

This study explored the dynamic relationship among language knowledge (i.e. vocabulary size and phonological capacities) and nonword repetition (NWR). Specifically, we proposed a three-phase working model which predicted that phonological capacities and vocabulary size might be dominating factors to NWR at different phases. Moreover, we examined the predictability of NWR to expressive vocabulary development. The study was conducted in a cross-sequential design. We recruited three cohorts of typically-developing children, respectively from the ages of 2, 3, and 4. They were followed for one year, tested at three time points, with an interval of 6 months. The children were tested with receptive vocabulary, expressive vocabulary, productive phonology, word discrimination, and NWR, which included nonce word repetition and gap word repetition. Inspection on children's NWR revealed growth with age. However, children of the same age manifested considerable individual variation. Findings of the analyses verified that productive phonology and vocabulary knowledge played roles in children's NWR development. Nevertheless, the extent to which they contributed to the variance in NWR was determined by the increase in vocabulary size and the nature of the stimuli. The effect of lexical knowledge was consistently found in children from age 2 to age 5, as evidenced by the vocabulary breadth effect in the age 2 and the lexicality effect in the older children. The findings indicated that children made use of existing vocabulary knowledge to support their encoding of novel sound forms. The mediation of productive phonology to NWR usually occurred after the mediation of vocabulary; however, productive phonology took over the role of determining NWR variation when children reached a certain level of vocabulary size and learned to retrieve for lexical support when encoding nonwords. The repetition of nonce words and the repetition of gap words were found to involved different processes. While the repetition of both types of nonwords was mediated by productive phonology, the repetition of nonce words was additionally supported by lexical knowledge. Based on the finding above, a revised model was developed to account for the processes involved in NWR. Furthermore, our study provided the evidence that NWR could predict children's subsequent expressive vocabulary knowledge.

Key words: nonword repetition, receptive vocabulary, expressive vocabulary, productive phonology

Table of Contents

List of Figures

List of Tables

Chapter 1 Introduction

1.1 Background

People have been interested in the mechanisms that govern the development of language, and have also expressed concerns for the causes of language disorders. A number of assessments have been developed to effectively identify children with language disorders. Nonword repetition (NWR) is one such measure that has been developed for this purpose, and has been found to be a powerful index to not only children's language development (Adams & Gathercole, 1996; Gathercole & Baddeley, 1989, 1990a; Gathercole & Willis, 1991; Hoff, Core, & Bridges, 2008) but also to children with language disorders (Conti-Ramsden, 2003; D'odorico, Assanelli, Franco, & Jacob, 2007; Girbau & Schwartz, 2007; Stokes & Klee, 2009; Stokes, Wong, Fletcher, & Leonard, 2006). In light of its significance to language development, looking into the nature of NWR would provide insights into the mechanisms that support language development.

Over the past 30 years, there have been considerable studies related to the NWR measure. While it has been constantly applied as a measure of either phonological memory or phonological representation in studies, a wealth of research has been carried out in parallel to explore the nature of this task. Consensus has been reached that this task taps the capacity to decode and encode phonological information, and the ability to maintain phonological information in storage. Performances in NWR could also be affected by the perception ability and the ability to organize articulatory gestures (Gathercole, 2006). The complex processing mechanism involved in the task makes it sensitive to any problems in the mechanism that would cause language difficulties or deficits. However, this complexity also makes its interpretation challenging, because a poor performance in this task could result from any of the processes involved, or could be the consequence of several causes. Therefore, it is necessary to understand the processes involved in this task, and also factors that would affect task performances, so that we can develop more accurate interpretation of children's performance in this task.

NWR performances develop with age, and demonstrate considerable variation across individuals. Different proposals have been raised regarding the major source of individual variation and developmental changes. Gathercole and colleagues (1989, 1990a) propose that NWR is a measure of phonological short-term memory (see also Gathercole, 2006). They made this proposal based on the working memory model of Baddeley and Hitch (1974). In this model, phonological short-term memory is one of the components specialized in managing verbal information. When verbal information is encountered, it will automatically enter into the phonological store in phonological codes and will be maintained in the store with the rehearsal process. The operation of subvocal rehearsal is most relevant to phonological short-term memory capacity, because a faster rehearsal rate can capture more information in the phonological store within the two-second memory decay period (Hulme, Thomson, Muir, & Lawrence, 1984). However, the driving force of subvocal rehearsal may not be at work in preschool children or younger, because past studies have shown that children do not spontaneously exploit the rehearsal strategy or other active memory strategies until age 7 (Gathercole & Adams, 1994; Gathercole, Adams, & Hitch, 1994; Henry, 1991).

The other line of research proposes that variation and development in NWR are more associated with the development in children's language knowledge, especially in the young children. For example, Metsala and Chisholm (2010) discovered that preschool children's NWR accuracy is supported by lexical knowledge. It was found that children had better repetition performances with the syllables that have lexical status, and they tended to change a nonword syllable into a word syllable. Also, their repetition of NWR may be mediated by the density of the lexical neighbors that the constituent syllable of the nonword has. The lexical effect and the neighborhood density effect were most prominent in multisyllabic nonwords.

In addition to mediation at the lexical level, NWR performance is also influenced by children's phonological analysis ability. Metsala (1999) and Bowey (1996, 2001) found that phonological analysis played an important role in the 4- to 5-year-old English-speaking children's NWR performances, even when the effect of short-term memory was controlled. The study of Li and Cheung (2014) on 4- to 5-year old Mandarin-speaking children showed that productive phonology was the major predictor of NWR, while digit span made a minor contribution. Moreover, they demonstrated that children's individual differences in NWR performance may reside in their ability to encode nonwords into appropriate phonological units.

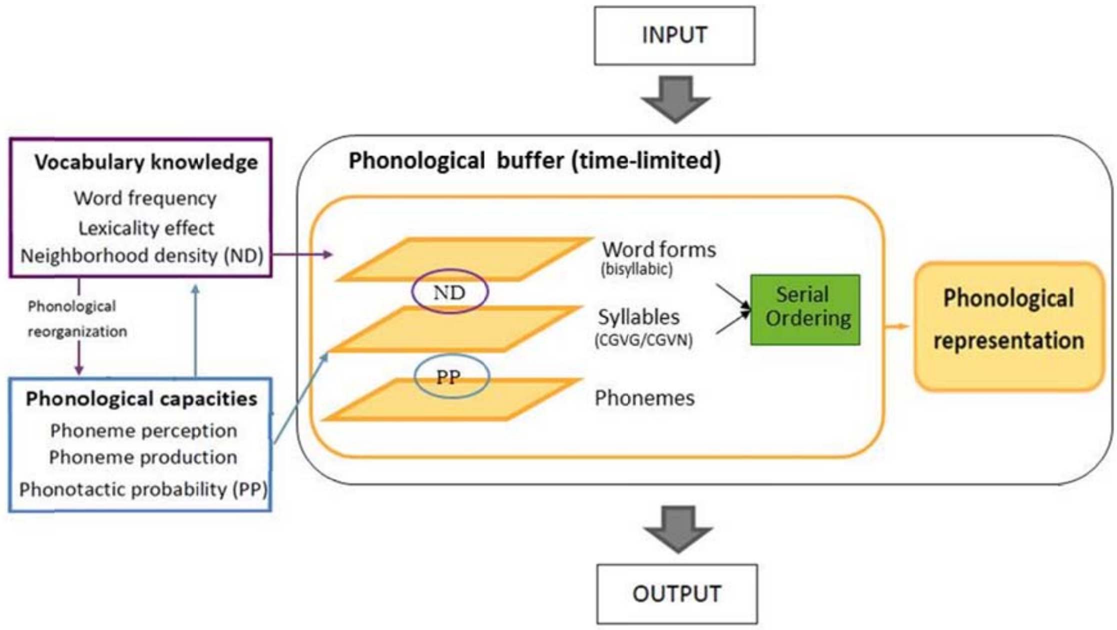
In most of the studies, the effects of lexical knowledge and phonological capacities are considered independently. However, in development they are not two unrelated constructs. In fact, increasing studies have pointed out that phonological capacities develop in conjunction with the increase in vocabulary size. It is proposed that phonological representations are shaped through the dynamics of the production-perception loop in the process of learning the forms of lexical items (Munson, Beckman, & Edwards, 2012). At the beginning, young children's phonological representations are relatively more holistic, with words or syllables as the basic units (Ferguson & Farwell, 1975; Treiman & Breaux, 1982). The emergence of phonemic representation undergoes a process of gradual reformulation and it is suggested to be propelled by vocabulary growth (Metsala, 1999; Smith, McGregor, & DeMille, 2006; Walley, 1993). This is also known as the "lexical restructuring account." Even though some scholars propose that the developing phonological system affects lexical acquisition to a greater extent than the reverse (Sosa & Stoel-Gammon, 2006; Stoel-Gammon, 2011), they also admit that this may be limited to the age before 2;6. From then on, the increase in lexicon size may be the driving force of phonological development. For example, the findings of Sosa and Stoel-Gammon (2006) suggested that phonological reorganization and the emergence of phonemic representation may take place with the attainment of 150-200 words.

Therefore, regarding the contribution of lexical knowledge to NWR, lexical knowledge does not only mediate NWR at the lexical level, but also could mediate NWR at the sublexical level by affecting phonological abilities. For example, Munson, Kurtz, and Windsor (2005) showed that vocabulary size was the best predictor of the difference in repetition accuracy between high- and low-probability sequences. Also, the study of Edwards, Beckman, and Munson (2004) demonstrated an interaction effect between children's vocabulary size and the phonotactic probability of nonwords. They discovered that children with smaller vocabulary size showed more prominent influence of the phonotactic probability of nonwords. Based on the findings, they propose that vocabulary size mediates the influence of phonotactic probability on nonword repetition by improving the specificity of phonological categories. Children with smaller vocabulary might have less established knowledge of sublexical units, because this knowledge is formed based on generalization made over lexical items.

As shown in the literature, a well-established model on NWR should not only take into consideration the effect of storage capacity, but also incorporate the influences from long-term lexical knowledge and phonological capacities. Gupta (2009) has incorporated all in a computational model of NWR. In the model, Gupta (2009) has incorporated a serial order mechanism and also linguistic representations at both the lexical and the sub-lexical levels. As suggested by Gupta (2009), the linguistic representations in this model constitute long-term knowledge, and the serial ordering device constitutes the short-term sequence memory. Therefore, when given a nonword, the participant has to decode and encode the nonwords into representations at the lexical level (i.e. word) and the sublexical levels (i.e. syllable and phoneme), and also to maintain the sequential information of the linguistic units. This process could be performed in the phonological buffer, which is subject to time decay (Barrouillet et al., 2009; Towse & Hitch, 1995). Hence, efficient decoding and encoding of the nonwords is required to maintain the verbal information. Though it is not mentioned explicitly in Gupta's model, vocabulary knowledge (i.e. word frequency, lexical effect, or neighborhood density) supports the repetition of nonwords up to the word form level, and phonological capacities (phonemic perception and production, or phonotactic knowledge) support NWR up to the syllable level. The mechanisms involved in NWR can be conceptualized as the figure demonstrated in Figure 1.

Figure 1.1

A Conceptual Model of NWR



While Gupta's (2009) computational model offers a clear and plausible illustration of the abilities and processes involved in NWR, this framework is static in nature. Nevertheless, the variables involved in this framework manifest developmental changes, particularly NWR performances, vocabulary knowledge and phonological capacities. Findings in past studies also reveal that the interactions between these variables might change over time. For example, Gathercole, Willis, Emslie, and Baddeley (1992) showed that NWR performance at age 4 could predict vocabulary knowledge at age 5.

However, among the older children, the dependence relationship between the two factors changed. The vocabulary knowledge at age 5 and age 6 predicted children's performance in NWR at age 6 and 8, respectively. Also, regarding the relationship between vocabulary and phonological development, there may be a change in the direction of influence between them (Stoel-Gammon, 2011). Therefore, what appears to be interesting is the dynamic interaction pattern of these three variables along the course of development.

Particularly, it is of interest how the growth in vocabulary or the growth in phonology influences the performances in NWR. Past studies examine the effect of vocabulary growth by comparing children with large vocabulary size with children with small vocabulary size (Edwards et al., 2004). This approach could allow us to examine how children form different processing strategies and performances when they accumulate different sizes of vocabularies. However, it does not allow us to control for children's variation in other aspects, such as their phonological analysis capacity, attention span, or other cognitive factors, which might also contribute to variation in NWR performance. Gupta's (2009) computational model simulates vocabulary growth in the model to examine NWR performance; however, it was done just for the purpose of establishing corpus for the model, and may not be assimilated to the nature vocabulary growth in children.

In addition, the relationship between vocabulary knowledge and phonological development is rarely considered in NWR studies. Even though the studies of Edwards et al. (2004) and Munson et al. (2005) have pointed out the close tie between vocabulary and phonological development, they did not measure children's phonological capacities independently, but rather manipulated the phonotactic probability of the nonword stimuli. They hold the belief that phonology is an emergent consequence of the mapping between phonetic parameters and lexicons (Munson, Beckman, & Edwards, 2012). However, it is also likely that phonological capacities have their own independent contributions, especially in cases where lexical knowledge is only allowed to exert little intervention, such as very young children who has only limited vocabulary size, or nonword stimuli that composed of non-attested syllables.

Most of the current NWR studies are cross-sectional, thus not able to demonstrate the dynamic relationship among the three variables. Longitudinal studies on the relationship between NWR and other measures have been rare (e.g. Gathercole, Willis, Emslie, & Baddeley, 1992; Melby-Lervåg, et al., 2012; Bowey, 2001), and all these studies have focused on children at 4 or above. However, if nonword repetition could potentially be used as an indicator of language abilities, it is necessary to examine its correlation with language abilities at a much younger age.

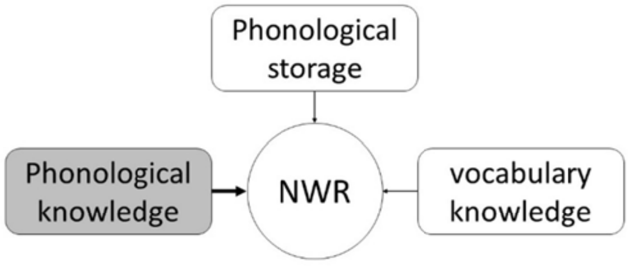
1.2 Purpose of the Study

In this dissertation, we would like to portrait the dynamic relationship among language knowledge (vocabulary size and phonological capacities) and NWR performance. Specifically, we explored the effects of vocabulary growth and phonological development on the improvement of NWR. A cross-sequential study was conducted, so that we could inspect not only cross-group differences, but also individual variation in the interaction among these capacities. The primary research questions that we address are as follows:

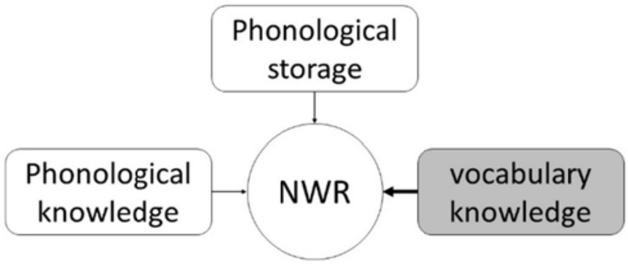
1. What are the roles of phonological capacities and vocabulary knowledge in NWR? And how do the two factors interact in NWR developmentally?
2. Does NWR predict vocabulary development?

Based on the literature, we formulated a working model to delineate the dynamic interactions among these factors in development. In this model, we propose that vocabulary knowledge and phonological capacities play roles in NWR. However, the extent to which they are involved in NWR is affected by children's vocabulary size. For children at the beginning stage of language development, their NWR performance is mainly constrained by their ability to accurately encode the sound forms. In this phase, the intervention of lexical knowledge would be small due to the child's little vocabulary knowledge (Phase I). Vocabulary knowledge begins to support the repetition of nonword when the vocabulary size grows larger. In other words, there will be greater chance for children to encounter familiar syllables in nonwords. Moreover, the increase in vocabulary size would also lead to refinement of more discrete representation of the phonemic units. Thus, children become more efficient and accurate in their encoding of the phonological representation of nonword. Their repetition of a single nonword, especially nonwords composed of unattested syllables, is still largely determined by their ability to accurately encode the sound form (Phase II). Finally, in phase III, when children have even greater vocabulary size and well-developed phonological representations which allow manipulation as abstract discrete units, their decoding and encoding of nonwords becomes less effortful. Their performance is likely to reach ceiling effect when repeating short nonwords, and the repetition of lengthy nonwords may be determined by the storage factors.

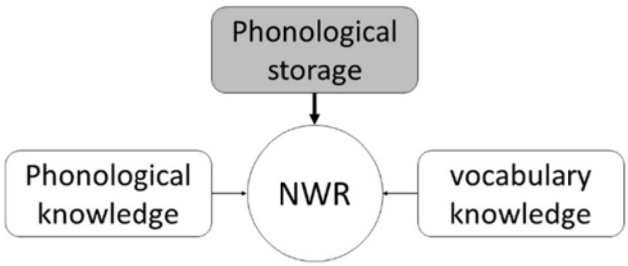
Phase I:



Phase II:



Phase III:



In addition to examining the developmental trajectory of NWR, we are also interested in its predictability to language development, particularly vocabulary development. Baddeley, Gathercole and Papagno (1998) have proposed the phonological loop, as measured by NWR, plays a crucial role in learning the novel phonological forms of new words. Also, findings in Gathercole et al. (1992) suggested that NWR could predict vocabulary knowledge of younger children, though it failed to predict that of older children. However, mixed results have been found. For example, with data derived from a three-year longitudinal study, Melby-Lervåg et al. (2012) found that NWR is not associated with vocabulary development. In this study, we would like to evaluate the hypothesis that NWR is a predictor of vocabulary development in Mandarin-speaking children.

For the purpose of our study, we recruited three cohorts of children, respectively from the ages 2, 3, 4, and followed them longitudinally, testing their NWR performance, vocabulary size and phonological abilities at three different time points, with an interval of 6 months. We included children as young as age 2 in this study, because their vocabulary and phonological system are still in development. It was interesting to examine their developments in vocabulary and phonology, and explore how these language abilities are related to NWR performance along the developmental trajectory. Some studies have examined NWR in children at this young age (Chiat & Roy, 2007; Hoff et al., 2008; Stokes & Klee, 2009). However, as far as we know, no one has followed children's phonological development and NWR performance longitudinally across different age groups.

In the analysis, first we carried out cohort-based examinations on children's development of vocabulary, phonology, and NWR, both quantitatively and qualitatively. Also, the relative contributions of phonological capacities and vocabulary knowledge to NWR performance were examined in each cohort. Then, we delineated the development of vocabulary, phonology and NWR across cohorts, from age 2 to age 5. Furthermore, the contributions of vocabulary knowledge and phonological capacities to NWR in the developmental trajectory were inspected with the hierarchical linear model approach. Finally, the predictability of NWR to vocabulary development was explored.

It is expected that this study will lead to a more comprehensive understanding of the NWR task, which is often applied to distinguish children with language problems. In addition to identifying the abilities that are involved in this task, we also explored how the factors (i.e. vocabulary knowledge and phonological analysis in this study) interact to contribute to NWR developmentally. In this way, NWR would not only function as a preliminary index of language performance, but could also potentially reflect the underlying cause of the impairment in language.

Chapter 2 Literature Review

2.1 NWR and Language Development

Young children's vocabulary development is worthy of attention due to its relationship with later language and literacy development (Adams & Gathercole, 1995; Adams & Gathercole, 1996; Gathercole, Willis, Emslie, & Baddeley, 1992). Children's vocabulary size in the early stage of language development can affect their performance in reading performance (National Institute of Child Health and Human Development 2000). A number of studies have shown that children's repetition of nonwords (NWR) can provide a quick and reliable index of children's vocabulary development in the early childhood. Children who have better performance at NWR are also more capable of learning new vocabulary items (Gathercole & Baddeley, 1990a). Moreover, low repetition group is poorer in long-term retention of new vocabulary materials. In addition, this measure can help identify children with language disorders. Children who have devastating performance in repeating nonwords, especially lengthy nonwords, are likely to be at risk of language impairment (Dollaghan & Campbell, 1998; Gathercole & Baddeley, 1990b). Since this measure is closely linked to vocabulary development, it provides a window through which we may examine what capacities are involved in word learning.

2.1.1 Relationship between NWR and language development. NWR is a good predictor to several aspects of language development during early years of life, particularly vocabulary development. The studies carried out by Gathercole and colleagues were the first to show that vocabulary knowledge is associated with NWR performance (Gathercole & Baddeley, 1989, 1990a), and this association has been replicated in numerous studies (see reviews in Gathercole, 2006). Recent studies have extended this finding to younger children. For example, Hoff, Core and Bridge (2008) found that NWR performances were significantly correlated with the expressive vocabulary percentile in children at age 2 (r = .53-.72, $p < .05$). Several studies tried to disentangle the causal relationship between vocabulary knowledge and NWR. Current findings show that NWR predicts children's vocabulary knowledge in the earlier years, while the direction of prediction changes as children grow older. For example, Stokes and Klee (2009) investigated the factors that affect the vocabulary development of children aged 24-30 months. They found that NWR was the unique predictor of their expressive vocabulary knowledge in addition to sex and age. With the cross-lagged technique, Gathercole, Willis, Emslie, and Baddeley (1992) showed that NWR performance at age 4 could predict vocabulary knowledge at age 5. Nevertheless, among the older children, the dependence relationship between the two factors changed. The vocabulary knowledge at age 5 and 6 predicted children's performance in NWR at age 6 and 8, respectively. The findings above appear to suggest that NWR is robustly associated with vocabulary development. However, this proposal is not without debate. For example, with data derived from a three-year longitudinal study, Melby-Lervåg et al. (2012) assert that phonological working memory is not associated with vocabulary development.

NWR is also found to be related to children's syntactic development and reading development. Adams and Gathercole (1996) found that 4- and 5-year-old children's ability to repeat nonwords made a significant contribution to the variance in children's ability to recall a story and the average length of the five longest utterances, independent of age, nonverbal abilities and vocabulary. Adams and Gathercole (1995) found evidence of a relationship between NWR and expressive language abilities indexed as the vocabulary diversity, the mean length of utterances in morphemes, and the syntactic complexity produced in the spontaneous speech of children at age 3. With a training study, Maridaki-Kassotaki (2002) demonstrated a strong correlation between NWR and reading skills in Greek-speaking children at ages 6 to 9. They found that children who receive one-year training on NWR showed a benefit in reading achievements. However, there might be an age effect in the correlation pattern. Gathercole, Willis and Baddeley (1991) showed that NWR was associated with reading among children at age 5, but not among children at age 4.

2.1.2 Relationship between NWR and word learning. A relationship between vocabulary learning and NWR performance has been established in several studies. For example, Stokes and Klee (2009) have found that children's performance in a fast mapping task is positively correlated with their performance in repeating nonwords ($r = .26$, $p < .001$, when age is partialled out). The study of Gathercole and Baddeley (1990a) revealed that children with good NWR performance were better in learning novel names, suggesting that NWR predicts word learning performance. Gathercole, Hitch, Service and Martin (1997) examined the association between NWR and new word learning in different conditions (i.e. word-word association, word-nonword association, new word in story context: recall of new word and recall of definition) in children at age 5. They found that NWR is associated with all the word learning conditions except for the word-word association. Therefore, Gathercole et al. (1997) suggested that the phonological memory component, as measured by NWR, is particularly involved in the acquisition of novel sound forms (see also Baddeley et al., 1998). However, when vocabulary scores were partialled out, these significant links were eliminated.

The positive relationship between NWR and word learning has been replicated in studies on younger children and studies on non-English-speaking children. For example, Weill (2011) examined the relationship between verbal working memory and new word learning in younger English-speaking children, age 24 to 30 months, and found that NWR is a significant predictor to children's performance in the receptive fast-mapping task. Lee (2005) also found a significant correlation between NWR and immediate word learning performance in Mandarin-speaking preschoolers. NWR, especially the repetition of lengthy non-attested nonwords, accounted for significant variance of the word learning performance. Storkel (2001) proposes that children's better ability to process the novel sound form would spare more capacity resources to process the semantic representation of the sound form, thus children could be better at learning novel words.

However, the association between NWR and word learning is not robustly found in all the studies. For example, in the investigation of Mandarin preschoolers' word learning, Yang (2002) showed that NWR was not a significant cause for the group differences in the word learning task, though there was a significant correlation between NWR and word learning performance (r = .30, p < .01). Ramachandra, Hewitt, & Brackenbury (2011) examined the relationship between phonological working memory, phonological sensitivity and incidental word learning in English-speaking children at age 4. It was discovered that NWR (adopted from Dollaghan & Campbell, 1998) did not make significant contribution to incidental word learning, while phonological sensitivity, as measured by rhyming and phoneme alliteration tasks, did. Similar findings have been discovered in a recent study by Abel and Schuele (2014). Yuen (2009) looked into Cantonese-speaking children between the age 3;2 to 5;1, and found no association between NWR and children's performance in the fast mapping task. These findings appear to suggest that word learning performance may be more related to language knowledge, than to NWR.

There might be some possible explanations to the discrepant findings among the studies. One is that the role of phonological memory to word learning is determined by children's concurrent language experience. Word learning could be dependent upon verbal STM in the very early stage of language development. However, when children have acquired considerable language experience/knowledge, their learning of novel sound forms would be supported by their language knowledge (Masoura & Gathercole, 2005). Another likely cause is the difference in task demand. Abel and Schuele (2014) pointed out that the significant link between NWR and word learning has usually been observed when the word learning task involves explicit teaching (Gathercole & Baddeley, 1990a). But the link is absent when an incidental word learning task was adopted, as in the studies of Ramachandra et al. (2011) and Abel and Schuele (2014). Therefore, it is likely that how the word learning is designed and instructed would incur a strategic difference in the acquisition of novel sound forms.

2.1.3 Relationship between NWR and language disorders. NWR is of clinical importance because poor performance in this task is indicative of language disorders, though this task alone may not be a sufficient index. Children with some forms of language disorders would perform poorly in repeating nonwords. Most notable is the robust impaired NWR performance observed in children with specific language impairment (SLI), and their performance deteriorates sharply with the increase in the lengths of nonwords (Dollaghan & Campbell, 1998; Gathercole & Baddeley, 1990b). Weak performance in NWR could also be observed among several clinical groups, including those with cochlear implants, stuttering, autism spectrum disorders (ASD), or language delay (LD). Given that NWR is a complex task, the performance of which involves a variety of processes (Gathercole, 2006), it is possible that different clinical groups may have similar NWR performance due to different underlying factors (Riches et al., 2011). For example, children with cochlear implants performed poorly on nonword imitation due to their constraints in auditory encoding of the nonword information. However, NWR could possibly serve as a phenotypic marker for some forms of language impairment (Bishop, North, & Donlan, 1996). For example, significant poor performances in NWR may not be observed among all the children who stutter, but only found among stuttering children with concomitant language or speech sound disorders (Smith, Goffman, Sasisekaran, & Weber-Fox, 2012). Similar findings have been observed in studies on children with ASD. For example, similarly poor NWR performance has been found between SLI children and ASD children with language impairment (ALI), while the ASD children with normal language development performed equally well with the typically developing children (Taylor, Maybery, Grayndler, & Whitehouse, 2014). However, SLI and ALI children may have different underlying causes of language deficits, as evidenced by the qualitative differences in their error patterns in NWR. There were a stronger effect of syllable length in SLI than ALI and a trend for SLI to make more errors affecting syllable structure, and drop more syllables. Also, a number of studies have suggested that NWR could be a putative marker for heritable language impairment, since family members of children with SLI (Bishop et al., 1996) or ASD (Bailey, Palferman, Heavey, & Le Couteur, 1998; Lindgren, Folstein, Tomblin, & Tager-Flusberg, 2009) show impaired performance on NWR. Therefore, poor NWR performance can be indicative of some forms of disorders at the processing of linguistic level.

2.2 Nature of Nonword Repetition Task

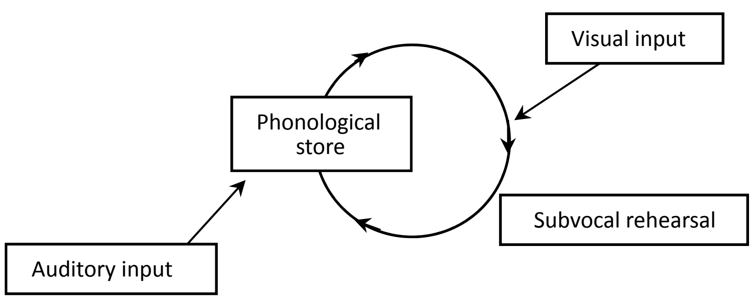
A number of processes and abilities are involved in the repetition of nonwords, including auditory processing, phonological analysis, phonological storage and verbal output abilities. Although a remarkable number of studies have been devoted to exploring the underpinning mechanism of this task (see Gathercole, 2006, for a review), consensus has not yet been reached regarding the major source of individual variation.

2.2.1 The phonological storage hypothesis. One of the most prevailing proposals regarding the nature of NWR is suggested by Gatherolce and Baddeley (1989, 1990a, 1990b), who have considered NWR as mainly a measure of phonological memory. The link was established by the association found between NWR and the conventional tests of memory storage, such as digit span (Gathercole et al., 1994, Gathercole & Baddeley, 1990b), and the association found between low repetition scores and patients with short-term memory impairments (Baddeley, 1993). This proposal was advocated by Gathercole and colleagues (see Gathercole & Baddeley, 1989, 1990a, 1990b; Baddeley et al., 1998) based on Baddeley and Hitch's influential working memory model (1974, Figure 2.1, see also Baddeley, 1986). In this model, phonological short-term memory is one of the components specialized in managing verbal information. When verbal information is encountered, it will automatically enter into the phonological store in phonological forms, and be maintained in the store with the rehearsal process. The operation of the subvocal rehearsal is most relevant to the phonological short-term memory capacity, because a faster rehearsal can capture more information in the phonological store within the two-second memory decay period (Hulme et al., 1984). Accordingly, an individual's phonological short-term memory capacity, as reflected by the NWR score, is determined by the rate at which one is able to rehearse the to-be-recalled nonwords.

However, one limitation of this account to young children's NWR performance concerns the source of variations observed in the course of development. Previous findings have shown that children do not spontaneously exploit the rehearsal strategy or other active memory strategy until a later age (Gathercole & Adams, 1994; Gathercole, Adams, & Hitch, 1994; Henry, 1991). Therefore, factors other than rehearsal rate should contribute to the developmental changes in NWR performances.

Figure 2.1

Baddeley and Hitch's (1974) Working Memory Model



Findings in the studies on working memory development may provide insights to the possible sources of developmental changes in NWR. For example, the study of Case, Kurland, and Goldberg (1982) proposed that the age-related increase in working memory capacities can be attributed to the greater efficiency in processing. They suggest that there is a total processing space which remains constant across ages, and is composed of an operating space and a storage space. The operating space is related to the execution of intellectual operations, and the processed item would be stored in the storage space. There would be a trade-off between the two spaces in processing. More efficient operation speed takes up fewer resources in working memory, thus freeing more available space for storage, hence the better recall in older children.

Towse and Hitch (1995) have proposed a task-switching hypothesis, which proposes that children alternate their attention between processing and storage during working memory span tasks. When children are engaged in the operating process, the memory traces would suffer from a time-related decay. The increase in working memory capacities in older children may result from their faster operating speed, and thus less time-based forgetting.

Barrouillet and colleagues (2004, 2009) demonstrated that both the trade-off account (Case et al., 1982) and the task-switching account (Towse & Hitch, 1995) could account for developmental or individual differences in working memory capacity. However, the task-switching model is more appropriate to account for working memory in preschool children (Barrouillet, Gavens, Vergauwe, Gaillard, & Camos, 2009).

2.2.2 The phonological analysis hypothesis. Contrary to the proposal that phonological storage is the major constraint to nonword repetition and word learning, other studies propose the role of linguistic factors. For example, Snowling, Chiat and Hulme (1991) suggested that lexical knowledge is involved in the repetition of nonwords. This is supported by the findings of the repetition advantage of real words over nonwords (Hulme, Maughan, & Brown, 1991), and wordlike nonwords over less wordlike nonwords (Gathercole, Willis, Emslie, & Baddeley, 1991). Also, children tend to change a nonsense syllable into a lexical item (Jones & Witherstone, 2010).

Other studies further propose that linguistic knowledge not only affects NWR at the level of lexical knowledge, but also at the more basic level of phonological analysis. That is, performance in NWR is constrained by the ability to efficiently process the novel verbal forms into accurate phonological representation (Bowey, 1996, 2001; Metsala, 1999). This proposal has received support from the findings that young children's phonological analysis, as measured by phonological awareness (Metsala, 1999) or output production (Li & Cheung, 2014), accounts for major proportion of variance when the effect of short-term memory is controlled. The effect of phonological analysis on NWR can also be observed in older children and adults when they encode nonwords constructed with nonnative phonological constituents (Kovács & Racsmány, 2008; Morra & Camba, 2009; Service, Maury, & Luotoniemi, 2007). For instance, when 8- to 10-year-old Italian-speaking children were asked to repeat and learn Italian nonwords which contained one Russian phoneme, their performance was more related to measures of phonological sensitivity, such as the phonological awareness of rhyme or initial consonants (Morra & Camba, 2009).

Among the studies favoring the phonological analysis account, differences should be noted regarding their assumptions of the relationships between phonological analysis and phonological storage. For example, Bowey (1996) asserts that phonological memory and phonological sensitivity may be surface manifestations of a latent phonological processing factor, possibly reflecting the clarity of underlying phonological representations of speech. However, others do not link between phonological memory and phonological analysis. For instance, Munson and colleagues (Edwards et al., 2004; Munson et al., 2005) consider NWR as a measure of children's abstract phonological encoding ability, and the relationship between NWR and word learning is due to the association of these constructs with phonological representation (Munson, 2006).

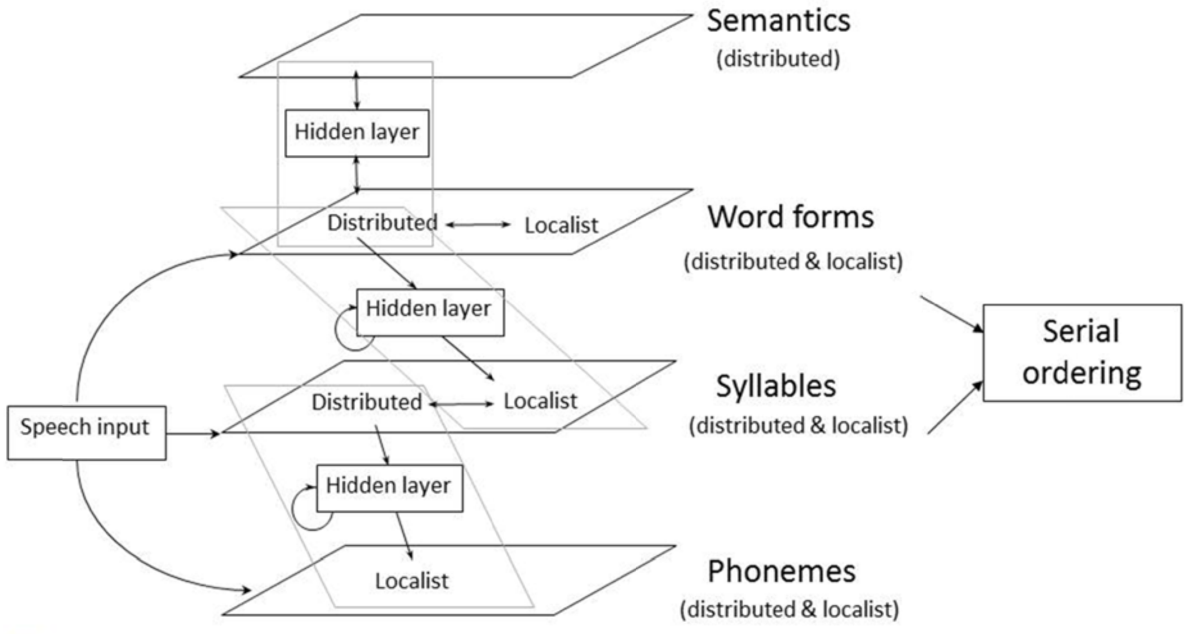
2.2.3 Gupta's (2009) computational model of nonword repetition. The previous accounts taken together implicate that NWR constitutes a domain of interaction between short-term memory and long-term memory. This concept has been demonstrated well in Gupta's (2009) computational model of nonword repetition. Gupta (2009) simulated the processes involved in nonword repetition, serial recall and nonword learning in a computation model. He proposed that when encoding a nonword or a list of nonword, one should be able to represent sublexical constituents of the nonword, and also to encode the serial orders of the sublexical units and the nonwords. Therefore, his model incorporated a serial order mechanism and also linguistic representations at the lexical and the sub-lexical levels. With an attempt to account for the mechanism of novel word learning, a representation at the semantic level was also included in his model (see Figure 2.2).

According to Gupta (2009), the presentation of a sound form would give rise to sequences of representation activated at the phonemes level and the syllable level, and of a single activation at the word form level. The serial ordering mechanism would help to encode the serial order of a sequence of activations, both at the lexical and the sublexical levels. Each of the word form level and the syllable level is composed of two sets of representations: the localist representations and the distributed representations. The localist and distributed representations at each level are bidirectionally connected (Gupta, 2009, p.113). The localist representation refers to the representation of an individual unit as the entire entity. Every unit in the localist pool (e.g. syllable level or phoneme level) has a connection to every unit in the distributed pool (e.g. word form level or syllable level). In the distributed representation, the entity would be represented as activations of a pool of units, in each of which there is a shared feature (or more than one shared features) with the entity. The distributed representation at the word form level are the phonologically structured representations of an entire word form, in a form of a string of syllables, while the distributed representation at the syllable level are the phonologically structured representation of an entire syllable, arranged in a format of syllable structure, such as CCVCC in English. The hidden layers function to translate input representation into a sequence of localist outputs that are represented in sequential order in the sub-level of representation. As suggested by Gupta (2009), this aspect of the model constitutes long-term knowledge, and the serial ordering device constitutes the short-term sequence memory.

The model constructed by Gupta (2009) is able to simulate qualitatively humanlike performance in NWR, thus providing strong evidence for the involvement of both long-term linguistic knowledge and storage factor in NWR processing and performance.

Figure 2.2

Gupta's (2009) Computational Model on NWR



2.3 Design of the Nonword Repetition Task

What NWR measures could be greatly affected by the construction of the task. Archibald and Gathercole (2006) have compared the most widely-used nonword repetition tasks in the English literature: the Children's Test of Nonword Repetition (CNRep; Gathercole, Willis, Baddeley, & Emslie, 1994) and the Nonword Repetition Test (NRT; Dollaghan & Campbell, 1998). CNRep consists of nonwords of one-, two-, three-, and four-syllable lengths. These nonwords are phonotactically and prosodically (i.e. stress pattern) legal. Due to the manipulation on stress, the nonwords may have weak syllables with a reduced vowel. In addition, some of the stimuli contain consonant clusters. NRT also consists of nonwords of one- to four-syllable of length. However, the nonword stimuli in NRT contain only single consonants. In addition, the nonwords are spoken with equal stress on each syllable. Archibad and Gathercole (2006) discovered that these two tasks showed different patterns of correlations with language impairments. Children with SLI performed significantly poorly than the age-matched group in both the two tasks. However, the SLI children performed significantly poorly than the language-matched group only in CNRep, but not NRT. The findings not only suggest that language and output factors may be involved in SLI in addition to memory problem, but also demonstrate that different psychological processes may be involved in nonword repetition that are constructed differently.

With regard to the construction of nonword stimuli, factors that would be considered include length, phonological complexity, lexicality/wordlikeness, phonotactic frequency and neighborhood density. How they may affect the NWR performances is reviewed in the following sections.

2.3.1 Construction of the stimuli.

Length. The construction of a nonword repetition task always involves the manipulation of length, because children's performance at items of different lengths may be especially informative in separating the group with language impairment from the typically-developing bilingual group (Windsor, Kohnert, Lobitz, & Pham, 2010), and also the group of SLI from other clinical groups, such as ASD (Riches et al. 2011). For example, now it has been well-established that children with language impairment usually have deteriorated performances in repeating 3- and 4-syllable nonwords (Dollaghan & Campbell, 1998; Gathercole & Baddeley, 1990b; Snowling et al., 1991).

Phonological complexity. As revealed in the study of Archibald and Gathercole (2007), CNRep is phonologically more complex than the NRT due to its inclusion of consonant clusters and also the variation in prosodic patterns. This complexity in phonological complexity may challenge SLI children to a greater extent than the controlled group with matched language abilities yet younger age.

Gallon, Harris, and van der Lely (2007) investigated the phonological deficits in children with Grammatical Specific Language Impairment (G-SLI). With this intent, they manipulated the nonword stimuli in terms of their prosodic complexity. They systematically varied the syllabic and metrical complexity of nonwords. In terms of the syllable structures, they manipulated three parameters, including onset (single consonant vs. consonant cluster), rhyme (open syllable vs. closed syllable), and word-end (vowel-final vs. consonant final). With regard to metrical structure, two parameters were considered. First was to determine whether a word contains an unfooted syllable adjoined to the beginning of a word. The other was to determine whether a word contains an unfooted syllable to the end of a word. Their study clearly demonstrated that the increase in the prosodic complexity of nonwords can result in deterioration in NWR accuracy in G-SLI.

Lexicality & wordlikeness. The study of Hulme, Maughan, and Brown (1991) demonstrated that the recall of real words is better than the recall of nonwords (i.e. *the lexicality effect*), because the latter lacks a long-term memory representation. However, the mediation of lexical knowledge does not only support the repetition of real words, but also the repetition of nonwords. This has been evidenced by the wordlikeness effect in NWR. Wordlikeness refers to the degree to which the nonwords are like real words. Superior performances are observed in the repetition of more wordlike nonwords (Dollaghan, Biber, & Campbell, 1995; Gathercole, Willis, Emslie, & Baddeley, 1991), because the repetition of wordlike items can be mediated by the mapping with an existing linguistic neighbor in lexical knowledge. On the other hand, it is less likely for low-wordlike nonwords to be mapped to well-established lexical representation, and the repetition of these nonwords could be largely dependent on phonological memory. Thus, the repetition performance is usually poorer in low-wordlike nonwords.

Phonotactic frequency and neighborhood density. The degree of wordlikeness of nonwords is rated through native speaker's subjective judgment of the nonword based on a 5- or 7-point scale. However, the judgment is in fact influenced by at least two objective factors: the similarity between a nonword and one or more particular words in the lexicon, or the phonotactic structure of the nonword itself (Frisch, Large, & Pisoni, 2000). The former is usually termed as neighborhood density, and the latter as phonotactic probability. By definition, neighborhood density refers to the number of phonologically similar words based on a difference of one sound. Phonotactic probability refers to the likelihood of occurrence of a sound sequence in a language. Phonotactic probability and neighborhood density are positively correlated (Vitevitch & Luce, 2005). Metsala and Chisholm (2010) discovered that children's repetition of NWR may be mediated by the density of the lexical neighbors that the constituent syllable of the nonword has (i.e. *the neighborhood density effect*). Also, it has been found that children had better repetition performances with nonwords containing high-frequency phoneme sequences (i.e., *the phonotactic probability effect*, Edwards, Beckman, & Munson, 2004; Gathercole, Frankish, Pickering, & Peaker, 1999; Messer, Leseman, Boom, & Mayo, 2010). Also, infants are aware of phonotactic probability as it was revealed in a discrimination task (Jusczyk, Luce, & Charles-Luce, 1994).

Vitevitch and colleagues have identified that phonotactic probability has more facilitative effects for nonwords over the neighborhood density, because “nonwords fails to strongly activate competing lexical representations” (Vitevitch & Luce, 1998; Vitevitch & Luce, 2005).

2.3.2 Procedures of NWR task. Nonword repetition is a task easy to apply to young children, because it resembles the process of acquiring the sound form of a newly-encountered word. It has often been applied to preschool children and children of older age. The nonword stimuli would be pre-recorded into the audio files, and played through speakers for the participants to repeat. In this way, the study could control for the articulatory variations of the inputs across participants and across sessions. To enhance children’s participation in this task, the NWR would proceed in a form of an imitation game, in which a puppy would be used to pretend as the speaker of the “weird language.” Some of the studies may choose to present stimuli live, but the experimenter had to cover up their mouth when delivering the nonword stimuli for the child to repeat. This is applied especially in studies which use NWR as a measure of auditory short-term memory, because they intend to eliminate any visual cues that could cause unintended facilitative effect to auditory short-term memory performance (Adams & Gathercole, 1995).

Recently, more and more studies explore NWR performances of children at a much younger age, almost as early as they are at the beginning of producing words (Chiat & Roy, 2007; Hoff, et al., 2008; Roy & Chiat, 2004; Stokes & Klee, 2009; Weill, 2011). These studies point out that the way of testing older children’s NWR may be too demanding for younger children, and some adjustments on the procedure are recommended to boost cooperation and maximize responses, such as presenting stimuli live, or making the imitative activity more interesting. For example, Stokes and Klee (2009) devised a ball-rolling activity to facilitate young children’s participation in the

NWR task. For children who repeated after the experimenter, they could get one chance to roll a ball down a chute.

2.3.3 NWR scoring. Children's performance in NWR could be scored at the word level (Gathercole et al., 1994), the syllable level or the phoneme level (Dollaghan & Campbell, 1998; Edwards & Lahey, 1998). At the word/syllable level scoring, the participant get one credit for each correctly recalled word/syllable. As for the phoneme-level scoring, each correctly repeated phoneme would get one point. However, given that nonwords with controlled number of syllables may differ in the number of phonemes they have, in phoneme level scoring, the researchers might took the percentage of the number of correctly repeated phonemes to yield a percentage-of-phonemes-correct score (usually termed PPC). In phoneme level scoring, studies may differ in some details regarding what are considered to be accurate phonemes, depending on the purpose of the NWR task. For example, in Dollaghan and Campbell (1998), which adopted NWR as a measure of phonological working memory, phoneme distortion and addition were not considered as incorrect because they do not represent any loss of information. However, in Coady, Evans, and Kluender (2010), which used NWR as a measure of phonotactic sensitivity, phoneme additions were counted as errors, because they reflected children's inability to maintain the syllable structure or phonotactic regularity of the target stimuli. Despite these differences, the systematic substitutions of young participants are usually disregarded and counted as correct, because they represent articulatory constraints in development, but not encoding deficits in memory.

Some clinical studies have compared the sensitivity of different scoring approaches. Riches et al. (2011) have asserted that comparing with the all-or-none, or word-level, scoring, phoneme-based scoring may enhance the sensitivity of the NWR assessment, because the all-or-none scoring is not able to contrast repetition with one (random) error with repetition difficulties affecting multiple phonemes per word. However, Estes et al. (2007) argued that in clinical studies, the all-or-none scoring may be more preferred, because it does not give children with language impairment any credit for partially correct repetition of nonwords. Guiberson and Rodriguez (2013) shared similar view when compared the syllable-level scoring and the phoneme-level scoring approaches in a classification study, and argued for a preference over the syllable-level scoring. As each scoring approach has its advantages and weaknesses, it is important to choose the most appropriate one based on the purpose of the study. The scoring at the large-unit level could be effective and sensitive to the identification of language deficits, while the scoring at the small-unit level could potentially reflect the underlying causes of the disruptions in NWR performance.

2.3.4. Error analysis. Examining the error patterns of the misrecalled nonwords could potentially reveal the processing mechanism of the NWR task. For example, the findings that participants tend to transform nonsense syllables into lexical items reveal the mediation of lexical knowledge in NWR (Jones & Witherstone, 2010). Moreover, inspecting the errors patterns of NWR in clinical groups may help to identify the possible causes of their underlying deficits. In the study of Edwards and Lahey (1998), they found that children with SLI exhibited different error patterns from those of typical development. The children with SLI made significantly more syllable structure and phoneme deletion errors and significantly fewer phoneme substitution errors. Riches et al. (2011) also discovered that children with autism plus language impairments manifested different error patterns from children with SLI, though both had poor NWR performances compared with the normally-developing children.

Error analyses could be conducted at different levels (e.g. syllable, phoneme, and feature) and with the identification of different processes (omission, addition, and substitution). For example, Riches et al. (2011) distinguished two types of error patterns in terms of retaining the syllable structure of nonword stimuli. They suggested that the structure-changing error, such as omission of an initial consonant or weak syllable deletion, would be an indication of difficulties with hierarchical prosodic and phonological processes. On the other hand, the structure-preserving errors, such as consonant-for-consonant substitution, reflects difficulties with simultaneous processing of metrical information (number of syllables and stress placement), and phonemic information; the encoding of phonemic information disrupts as the result. While the former is proposed to be more associated with attentional/memory processes, the latter is more associated with phonological representation (Riches et al., 2011).

Error analyses could also be conducted at the phonemic level. For example, Shriberg et al. (2009) distinguished within-class substitutions (e.g. /t/-/d/) from between-category substitutions (e.g. /t/-/tʃ/), with the class defined as manner of articulation. They proposed that a substitution preserving manner could be interpreted as a partial encoding of the target. On the other hand, poor encoding of the manner feature might indicate a poor auditory-perceptual encoding.

As revealed above, misrecalled errors in NWR could be analyzed from several different aspects. However, solid theoretical background should also be provided to justify the interpretations. It is also important to bear in mind that when examining NWR misrecalled errors of young children in clinical groups, we should distinguish developmental articulation errors related to normal phonological processes from repetition errors which result from representation or memory deficits (Dollaghan & Campbell, 1998; Guiberson & Rodriguez, 2013).

2.3.5. NWR in Mandarin—Characteristics of Mandarin NWR. NWR studies in Mandarin have been rare, and the NWR tasks applied in these studies could be divided into two types, based on their designs. The first type of NWR is devised with the purpose of assessing the overall performance in phonological working memory. Each item in this task consists of three bisyllabic nonwords, the syllables in which are all existing syllables in Mandarin (Table 2.1). The composition of nonwords is limited to a restricted set of onsets and rimes. The nonword syllables are from the combination of one of the onsets *b, d, k, g, zh, sh*, with one of the rhymes *u, a, ai, ao, an, ang*. The NWR task of this type was administered in Hu and Catts (1998), Yang (2002), and Hsu (2005).

Table 2.1

The Nonword Stimuli Used in Hu and Catts (1998)

Practice trials
1) bai3-sha4 kang1-gu4 zhan4-dao1
2) kang4-zhai1 ga1-shu3 bao3-dan1
3) zha1-ban1 dao3-shang4 gu4-kai3
Experimental trials
1) ba3-gan1 zhao1-dai3 ku4-shang4
2) shang1-kai3 dao4-ga4 zhan3-bu1
3) bang1-zhai4 kan4-du3 sha4-gao1
4) zhan4-da1 shu4-gao3 bai1-kang4
5) ga1-shao3 bang4-ku3 dan3-zhai1
6) shao1-ga1 ban4-zhu4 kang1-dai3

However, the increasing findings that a number of factors other than memory storage may influence performance in NWR have called for a more sophisticated design of the task. Therefore, in the other type of the NWR task, the length and lexicality of the nonword stimuli were manipulated to control for the influence of memory storage, lexical knowledge and phonological analysis. In this NWR task, two types of nonwords are constructed, including the nonce-words and the gap words. The nonce words are composed of existing syllables in Mandarin, while the gap words are made up of phonotactically legal but non-existing syllables in Mandarin. In the processing of a nonce word, though it cannot be mapped onto any existing semantic representation, each of its constituent syllables can be mapped onto a lexical representation. With regards to the processing of a gap word, which is absent in real-life language use, it has no stored representations at either the semantic level or the lexical level. Therefore, the processing of gap words depends on phonological analysis ability (Lee, 2005). The length of nonword items is manipulated by conjoining bisyllabic nonwords, thus the one-word lists, the two-word lists, and the three-word lists in the nonce word repetition and the gap word repetition. The NWR in Li and Cheung (2014) was a modification of Lee (2005). They added a number of nonwords to equalize the number of each syllable shape.

All these studies reveal that NWR in Mandarin-speaking children improve with the increase of age. Also, children have significantly better performances in nonce-word repetition over gap-word repetition, demonstrating the lexicality effect. However, due to the absence of a large-scale oral corpus in Mandarin, the phonotactic probabilities of nonwords were not controlled in these studies.

2.4 Phonological Development

2.4.1 Its relationship with lexical development. Children's phonological development refers to the process of reaching adult-like phonological representation, which is phoneme-based. A number of studies have demonstrated that children's phonological representation is word-based. That is, they perceive "words" as an unanalyzable phonological unit (Treiman & Breaux, 1982). Also, their production of words is the approximation of the whole-word unit (Ferguson & Farwell, 1975). The specification of the phonological representation of words is proposed to be driven by the increase of vocabulary size, which is known as the lexical restructuring account (Fowler, 1991; Metsala, 1999). As the neighborhood density of the vocabulary network increases, this would lead to restructuring of the network. Children are forced to be able to distinguish phonological similar words, such as *bat* and *pat*, causing their phonological representation to become more and more fine-grained. Also, in the process of coordinating the articulatory gestures to assimilate the sounds they hear, the co-occurring gestures would gradually be crystallized into segments (Studdert-Kennedy, 1987, 2000; Studdert-Kennedy & Goodell, 1995). Therefore, a growth in vocabulary size means not only that children know more words, but also that children are provided with richer resources from which they could generalize abstract sound patterns and processing units of their ambient language (Edwards et al., 2004).

However, another line of research proposes that vocabulary may not be the driving force of phonological development right from birth. In a review paper on the relationship between vocabulary and phonology in development, Stoel-Gammon (2011) suggests that early phonological development, at least during the period from birth to age 2;6, exerts more influence to lexical development than the other way around. Children's early lexicon is filtered by their phonological productive ability.

Given the discrepant view regarding the relationship between vocabulary and phonology in the early stage of language development, Edwards et al. (2011) suggest longitudinal studies should be conducted to examine the interaction between these two constructs in development.

2.4.2 Assessing phonological development. Traditionally, phonological development is assessed by examining children's acquisition of phonemes, either in production or in perception. Tasks used for the evaluation of phoneme production include picture naming, mispronunciation correction task, or speech imitation. Tasks used for the assessment of phoneme perception include mispronunciation detection task, accurate pronunciation identification task, speech discrimination task and gating paradigm. Also, the production errors would be examined in terms of the phonological processes that they undergo, which would be taken as a reflection of the mechanisms that children formulate speech in relation to the adult model (Ingram, 1989).

Table 2.2

Representation-related Phonological Processing Abilities and the Related Measures

Representation-related phonological processing abilities	Measures
1 Ease of forming new phonological representations	Word-learning paradigms
2 Accessibility of extant phonological representations	1. Picture naming tasks 2. Gating paradigm
3 Precision of extant phonological representations as reflected in speech perception	1. Mispronunciation detection task 2. Accurate pronunciation identification task 3. Speech discrimination task
4 Precision of extant phonological representations as reflected in articulation accuracy	1. Picture naming 2. Imitation of real words and nonwords 3. Mispronunciation correction task

More recent studies point out that these measures are in a way assessing phonological representation. Phonological representation is an abstract unit that cannot be assessed directly. Several studies have attempted to find out the best measure to tap this construct (Anthony et al., 2010; Foy & Mann, 2001). By far the study of Anthony et al. (2010) demonstrated the most comprehensive comparisons across all the possible tasks. They propose that the quality of phonological representation can be examined by at least three representation-related phonological processing abilities, including the accessibility of phonological representations, the precision of phonological representation evidenced via articulation accuracy and evidenced via speech perception (Table 2.2). These abilities have been shown to be equally good assessments of phonological representation. However, in order to obtain more accurate evaluation of the quality of phonological representation, it is recommended to apply more than one task to assess two or more representation-related abilities.

2.5 The Phonological Development of Mandarin-speaking Children

2.5.1 The phonological system of Mandarin. Compared with English, Mandarin has simpler phonological structure. It has only 12 possible syllable structures, including V, VC, VG, GV, GVC, GVG, CV, CVC, CVG, CGV, CGVC, and CGVG. There are 21 consonants that can occur in the syllable-initial position and only two consonants (i.e. *n* and *ng*) can occur in the coda position. No consonant clusters are allowed in Mandarin. The initial consonants and final nasals are optional, while the vowel is compulsory. In terms of vowel, there are 9 simple vowels, 9 diphthongs, and 4 triphthongs (Zhu & Dodd, 2000). The diphthongs can be further divided into offglides (labeled as VG in syllable structure) and onglides (labeled as GV in syllable structure). The offglides refer to the diphthongs with the first vowel sounds that are longer and more intense, such as *ai*, *ei*, *ao*, and *ou*. Onglides refer to the diphthongs in which the second vowels are more sonorous, including *ia*, *ie*, *ua*, *uo*, and *üe*. The triphthongs include *uai*, *iao*, *ui*, and *iu*.

Each syllable in Mandarin is composed of its segmental combination and tone. There are four tones in Mandarin, high-level tone (Tone 1), high rising tone (Tone 2), falling-rising tone (Tone 3), and high-falling tone (Tone 4). Tones in Mandarin have lexical status for they can distinguish meanings. For instance, the CV combination /ma/ would yield different meanings when combined with the four tones, respectively *ma1* (mother), *ma2* (hemp), *ma3* (horse), *ma4* (scold).

2.5.2 The phonological development of Mandarin-speaking children. We reviewed a number of studies on Mandarin-speaking children's phonological development. In addition to studies targeted at Mandarin-speaking children in Taiwan, we also included a large-scale study on phonological development of children speaking Putonghua, a Mandarin dialect speaking in Beijing, China. Its larger subject pool and longitudinal data might provide robust findings regarding Mandarin-speaking children's development in phonological system. However, we should be aware of the possible regional or dialectal differences in developmental patterns in interpretation.

With regards to tones, they are acquired well before the acquisition of segments. By the age of two, articulation control over tone production is mastered (Hsu, 1987; Li & Thompson, 1977; Zhu, 2006). Also, it is reported that children have little difficulty learning the tone sandhi rule, which is acquired by age 3 (Hsu, 1987; Li & Thompson, 1977). However, in tone perception, children tend to confuse Tone 2 and Tone 3 easily.

Findings regarding the acquisition of vowels have demonstrated considerable differences across studies (Table 2.3). For example, Shiu (1990) revealed that the vowel *ü* was stabilized in children's production at the age of 3; however, Hsu (1987) demonstrated that it was not stabilized until after age 6. Even within a study, we also observe individual variation. In Jeng's (1985) study, child K acquired the vowel *u* and *o* at 1;1. Nevertheless, child J did not acquire these vowel until age 1;6. Despite these differences across studies and participants, a general pattern could be observed. There is a tendency for the simple vowels to be acquired by the age of 2 (except for the vowel *ü*, which is often acquired late). The acquisition of diphthongs and triphthongs is not complete until 5.0.

When we look into Mandarin-speaking children's mastery of final coda, it is found that children in Taiwan tend to substitute /ŋ/ with /n/, different from children speaking Beijing Mandarin (or Putonghua) in Zhu and Dodd (2000). Also, it is found that it is more common for Taiwan children to omit final coda than Putonghua-speaking children speaking (Liu, 2007).

Consonant development is more often the issue of investigation in phonological development, because consonant acquisition extends a longer period of time, usually not complete until the age of 5. When examining children's acquisition of phonemes, we are not only interested in the time a particular phoneme emerges in children's production. In fact, we are more interested in when the phoneme becomes stabilized in production. Phoneme stabilization is a measure indicating the growing consistency of children's pronunciation of a certain phoneme (Zhu & Dodd, 2000). A sound is considered stable when the child produced the sound correctly on at a considerable numbers of occasions over its total productions. For example, in Zhu and Dodd's (2000) study, a phoneme should have at least an accuracy of 66.7% (two out of three productions are correct) to be considered as stable. When 90% of the children in an age group achieved this accuracy rating for a phoneme, the phoneme would be considered to have been stabilized by that age group. However, different studies may adopt different standards. The criterion of 75% is often adopted in the literature as well. Though differences are observed across studies, similar acquisition patterns were observed (Table 2.4). When the criterion of 75% is adopted, phonemes acquired before age 3 generally include *b, m, d, t, g*, and those acquired before age 4 include *k, q, x, j, z, c, s, f, h*. Retroflexes such as *zh, ch, sh* are acquired relatively late, because they are articulatorily more challenging for young children.

In general, in Mandarin-speaking children's development of phonological units, tones are acquired first, followed by syllable-final consonants and vowels. Syllable-initial consonants are acquired last (Zhu & Dodd, 2000).

Table 2.3

Mandarin-speaking Children's Development of Vowels

Study	Jeng (1985)[1]	Shiu (1990)	Hsu (1987)
Participants	Two boys: Child J: 1;3-2;5 Child K: 1;0-2;6	One boy and one girl, by 3;0	28 children, aged 1;0-6;0
1;0-1;6	Child K: 1;0 e, a → 1;1 u, o → 1;3 i → 1;4 ü Child J: 1;4 i, e, a → 1;6 u, o → 1;8 ü	i, u, a, ua	a,
1;7-2;0			u, i, /ə/, ia, ua, an
2;1-2;6		o, /ə/, e, ai, ao, ang, in, ia, ie, io, uo, ue, iao, uai, ui, ian, uan, un	ai, ei*, ie, iau, iu*, uai, ui*, ang, en, in, ing*, iang
2;7-3;0 (3.0)		ü, ei, ou, an, en, eng	ung, ian, iong*,
3;1-4;0 (3.0)			au, ou*, uan, uang, un, uo
4;1-5;0 (4.0)			
5;1-6;0			
After 6.0			ü, ue, eng, ün, üen

Notes: Chen (2005) looked into Mandarin-speaking children's development of vowels. Nevertheless, she looked at only one child. Therefore, the acquisition patterns could be influenced by individual differences. For that reason, it was not presented here.

1. Considerable variation was found between Child K and Child J's acquisition of vowels. However, *i* and *ü* tended to be the last vowels acquired
2. Zhu and Dodd (2000) also examined the development of vowels. However, they did not specify when each vowel became stabilized. In general, they found that the acquisition of vowels, particularly the simple vowels, took place mainly between the age of 1;0 and 2;0. Among simple vowels, low vowel *a* and back high vowel *u* occurred earliest, while retroflex vowel and the back vowel *o* occurred last. As for diphthongs, *ei* is acquired first, and *üe* last. Among triphthongs, *iu* is acquired first, and *ui* last.

Table 2.4

Large-scale Studies on Mandarin-speaking Children’s Development of Consonant Productions

	Cheng et al. (2003)	Hsu (1987)[1]	Wang et al. (1986)[1]	Chang & Chung (1986)	Zhu & Dodd (2000)	Cho (2008)	Lin & Lin (1994)	Zhu & Dodd (2000)	Cho (2008)
Sample	Taiwan preschoolers (Tainan), aged 2.5 to 6	28 children (1;0-6;0)	150 Taiwan preschoolers (Taipei), aged 3-6	363 Taiwan preschoolers (Taipei), aged 2-2;6	129 Beijing children, aged 1;6-4;6	852 Taiwan preschoolers (Taipei), aged 3-6	839 Taiwan preschoolers (Taipei), aged 3-5;11	129 Beijing children, aged 1;6-4;6	852 Taiwan preschoolers, aged 3-6
Measure	Picture-naming	Diary records	Picture-naming	Not specified	Picture-naming & picture description	Pic-naming & sentence imitation	Measure of Chang & Chung (1986)	Picture-naming & picture description	Pic naming & sentence imitation
criterion	70%	stabilized	75%	75%	75%	75%	90%	90%	90%
Before 2;0	b, n	-1;8 b, m, d, g,	b, p, m, d, t, n, l, g, h		d, t, m, n, h	b, d, t, g, k, l, h, m, ng, coda n, r		d, m	
2;1-2;6	b, n	1;8-2;6	b, p, m, d, t, n,		b, p, g, k, j, q,			n	

	m, d, g, j	n, j	l, g, h 2.11-3.5		x		b, p, m n, l, g, k, h, j, q		
2;7-3;0			k, q, z	b, p, m, d, t, n, l, g, j	f			b, t, f, h, x	
3;1-3;6				k, h, q, x, zh, z, c, s		p, n, x, z, c, j, q, ch		g, k	b, d, t, g, k, l, h, m, ng, n, r
3;7-4;0	l, k, h	3;4-4;0 p, t, k, h,	3.5-3.11 x, c, s, f	ch		f, s, sh	d,	p	n, j
4;1-4;6	f, t, q		4.0-4.11 j	f, sh	l, sh, r, s	zh	x, z	l, s, j, q, r	p, ch
4;7-5;0		4;4-6;0 f, l, q, x, z, c, s	4.0-4.11 j		after 4.5 zh, ch, z, c		t, c	zh, ch, sh, z, c	f, z, c, q
5;1-5;6	p, x, z, c		5.0-5.5				sh		s, x
5;7-6;0			5.6-5.11				f, ch, r, s		
After 6.0	zh, ch, sh, r, s		zh, ch, sh, r	r			zh		zh, sh

Note: Jeng (1985) and Shiu (1990) also examined Chinese-speaking children's development of phonemes. However, they looked at only one or a few more children. Therefore, the acquisition patterns could be influenced by individual differences. For that reason, the two studies were not presented here.

1. The study of Cheng et al. (2003) is cited in Chi (2009).
2. When a study adopted different age periods when demonstrating the acquired consonants, the age periods adopted in the specific study would be specified in rectangular.

Chapter 3 Method

This study aimed to explore the developmental trajectories of nonword repetition, phonological capacities and vocabulary knowledge, and the relationship among them along the course of development in young children. In order to model the developmental change across an extending period of ages, a cross-sequential design was adopted. We recruited children of three age cohorts, respectively age 2, age 3 and age 4, and followed their growth in vocabulary knowledge, phonological capacities, and nonword repetition for one year. Detailed descriptions of the participant information, experimental tasks that we administered, and the analyses of the data were provided in the following sections.

3.1 Participants

We recruited typically-developing children at age 2, age 3, and age 4. Children were recruited from kindergartens, posts on parenting websites or by word-of-mouth. Children who had been diagnosed as language delayed were not included in this study. Also, at the beginning of the test, children were screened with the "Preschool Child Developmental Checklist" for his or her age (see the Appendix A for the sample for age 2 children). All of the children passed this screening test, and showed no sign of developmental delay. While previous studies suggest that NWR performance is not independent of language (Coady & Evans, 2008), balanced bilinguals proficient in another language in addition to Mandarin were not included in this study. All of the children speak Taiwan Mandarin as their dominant language.

Table 3.1

Participant Information

		Age 2	Age 3	Age 4
Time 1	*N*	24 (12 ♂)	24 (8 ♂)	24 (11 ♂)
	Mean age	24.92 months	35.80 months	48.83 months
	Age range	23.83-26.63	35.47-38.93	47.36-50.8
Time 2	*N*	22 (12 ♂)	23 (7 ♂)	24 (11 ♂)
	Mean age	31.10 months	42.81 months	54.97 months
	Age range	29.93-32.93	41.47-44.63	53.67-56.9
Time 3	*N*	21 (11 ♂)	20 (6 ♂)	23 (11 ♂)
	Mean age	37.00 months	48.70 months	60.89 months
	Age range	35.57-39.53	47.37-50.73	59.57-62.57

At the first testing (Time 1), there were seventy-two children participating in this study, 24 children in each age cohort. Each child was followed for one year, and tested at three time points with an interval of 6 months. At Time 2, three children did not return, because the parents were unable to cooperate with the testing schedule (Age 2: 2; Age 3: 1). At Time 3, additional 4 children were lost due to parents' inability to cooperate with the schedule (Age 3: 1) or children leaving the kindergarten (Age 3: 2; Age 4: 1). Table 3.1 presented our child participants' background information.

Since children in each of the age group were followed for one year, each age cohort would overlap with another age cohort. For instance, the age 2 cohort overlapped with the age 3 cohort at age3, and the age 3 cohort overlapped with the age 4 cohort at age 4. Table 3.2 demonstrated the number of data at each chronological age. In total we had 206 data from our three cohorts.

Table 3.2

Number of Data at Each Age

Timeline	2	2.5	3	3.5	4	4.5	5
Age 2 cohort	24	22	22				
Age 3 cohort			24	23	20		
Age 4 cohort					24	24	23
Total	24	22	46	23	44	24	23

3.2 Experimental Tasks

In this study, we tested children with their expressive vocabulary, receptive vocabulary, phonological perception, phonological production and NWR. In addition, we tested children's nonverbal intelligence with the Leiter-R as a control for their cognitive abilities. However, concerning the broad age ranges we covered in this study, some constructs were tested with different tests that were more appropriate for the age. For example, vocabulary knowledge among children at age 3 or above was measured with the Receptive and Expressive Vocabulary Test (REVT, for expressive vocabulary) and Peabody Picture Vocabulary Test-Revised (PPVT-R, for receptive vocabulary). However, for children below age 3, their expressive vocabulary and receptive vocabulary were tapped with the Mandarin-Chinese Communicative Development Inventory (MCDI-T). Also, on account of younger children's smaller attention span, we downsized the number of trials children below age 3 had to complete in some of the tasks. These adjustments would be specified in the following introduction of the tasks.

3.2.1 Mandarin-Chinese Communicative Development Inventory (Taiwan) (MCDI-T). The MCDI-Taiwan (Liu & Tsao, 2010) is a parent-report evaluation of infants'/toddlers' word production, communicative functions, sentence complexity, and the mean length of the three longest utterances. This assessment was applied in this study to collect the data on very young children's expressive vocabulary size. Since there was no parallel measure of receptive vocabulary for children below 3 in Mandarin, this form was also used to collect our children's receptive vocabulary.

3.2.2 Peabody Picture Vocabulary Test-Revised (PPVT-R). The PPVT-R (Lu & Liu, 1994) is often used to measure age 3 to age 12 children's receptive vocabulary knowledge. Different from the MCDI-T, this test is administered by the experimenter. In this study, it was also applied to test children's receptive vocabulary at age 2.5.

The test has form A and form B. In this study, form A was adopted. In the test, the child responded by pointing to one of four line drawings that corresponds to the word spoken by the experimenter. The test has 125 items. The child started from the item appropriate to his or her age. Testing was stopped when the child made 6 errors in 8 consecutive items. The total number of items the child accurately answered was counted as his or her PPVT-R score.

3.2.3 Receptive and Expressive Vocabulary Test (REVT). The REVT (Huang, Jian, Zhu, & Lu, 2011) is designed for children at the age from 3 to 6, and consists of the scale of receptive vocabulary and the scale of expressive vocabulary. Each scale is comprised of four subtests, including label, category, definition, and reasoning. This test takes into consideration the specific linguistic features of Mandarin, in comparison with the linguistic structure of English. We adopted the scale of expressive vocabulary to assess children's ability to encode or express language based on concept.

3.2.4 Productive phonology task. This task tapped the children's productive phonology, which reflected the quality of phonological representation (Anthony et al., 2010). Colored pictures of familiar objects were used to elicit the children's production of 21 Mandarin syllable-initial consonants. Since words in Mandarin are mostly disyllabic, the position constraints were considered by testing all the target phonemes in the first syllable and the second syllable (see Appendix B).

In order to promote young children's participation, we also designed a "find-the-figure" game to elicit children's production. Five pictures were created incorporating the figures of the target words. Children were given pieces of the figures, and were asked to paste each to its corresponding figure on the pictures. While pasting, children were required to name the figure.

Each child would be given a score on onset production and a score on rhyme production. Onset and rhyme were scored separately because previous studies suggest that they play asymmetric roles in speech processing and language acquisition (Nazzi, 2005; Nespor, Peña, & Mehler, 2003). With regard to the scoring of onset, we first calculated the accuracy rate of the target phoneme in each lexical item; then we took the average of the accuracy rates across all the lexical items that contained the target phoneme. For example, for the phoneme /b/ at the word initial position, one child produced *bei1-zi* (cup), *bai2-tu4* (white rabbit), and *bang1-mang2* (help). For the lexical item *bei1-zi*, 5 out of his 6 productions of /b/ were accurate; therefore, the accuracy rate of /b/ in this lexical item was 0.83. The accuracy rates for /b/ in *bai2-tu4* and *bang1-mang2* were 0.7 and 1, respectively. Therefore, the average accuracy rate of /b/ at word-initial position was 0.8433. One point was given when the target phoneme was pronounced correctly in all the lexical items that contained it. The children's systematic mispronunciations for a single phoneme were considered inaccurate but were noted for scoring NWR. Children's scores of each of the phoneme at the word-initial position and the non-word-initial position were summed up to yield a productive phonology score, with the maximum score of 42.

The scoring of rhymes followed a similar procedure. We first calculated an average accuracy rate for each of the rhymes. Then we took the average of the rhymes with the same rhyme structure. For example, we averaged the accuracy rates of the rhymes *ai, ei, ao,* and *ou* to get a score of the VG rhyme structure. There are 6 rhyme structures in Mandarin, including V, VN, VG, GV, GVN, and GVG. The maximum score for rhyme production was 6.

3.2.5 Word discrimination. This task was designed to tap children's discrimination ability at the phonemic level. There were four practice trials and 24 test trials. The test trials were composed of 6 pairs of sound contrasts. Two pairs of the sound contrasts differed in three features (3-*f difference,* e.g. *t-m* and *b-q*); another two pairs differ in two features (2-*f difference,* e.g. *h-p* and *x-j*); and the other two pairs differ in only one feature (1-*f difference,* e.g. *d-g* and *n-l*) (see Appendix C for the stimuli). A female Taiwan Mandarin speaker produced spoken version of these stimuli in a sound-attenuated booth, and recording were made by using a DAT-recorder.

In this task, children were instructed to answer "yes" or "no" in response to the correspondence between the picture they saw and the audio label they heard. For example, for the *n-l* pair, children would see a picture of a bird. In the "yes" trial, the children were asked "*zhe4-li3 you3-mei2-you3* ***niao3****?* (Is there a bird here?)" However, in the "no" trial, the children were asked "*zhe4-li3 you3-mei2-you3* ***liao3****?* (Is there a knotweed?)"

The test was divided into four blocks. The presentation order of the six pairs of stimuli would be counterbalanced across the four blocks. Therefore, for each sound pair, there would be four times of discrimination, in which each of the sound stimuli would appear twice in the audio, and the presentation order of the sound stimuli would be counterbalanced. However, considering that younger children might have shorter attention span, children below age 3 only had to complete the first two blocks.

Each accurate response was credited with one point. The maximum score for the task was 24. However, for children at age 2 and age 2.5, the maximum score was 12.

3.2.6 Nonword repetition. Two types of nonwords were constructed based on the characteristics of Mandarin phonology, following the Mandarin NWR studies (Lee, 2005; Li & Cheung, 2014). The first type is nonce words, which are nonsense words consisting of two existing syllables. Though a nonce word cannot be mapped onto any existing semantic representation, each of its constituent syllables can be mapped onto a lexical representation. The second type is gap words, which are formed by conjoining two phonotactically legal but non-existing syllables. Gap words are absent in real-life language use and therefore have no stored representations at either the semantic level or the lexical level. Based on their compositions, the two types of nonwords are distinctive in their lexicality. While nonce word repetition is supposed to be more related to vocabulary knowledge, the repetition of gap words may tap phonological analysis.

Some gap syllables may be real syllables in other dialects in Taiwan, and children's familiarity with these dialects may affect their performance. In fact, we have checked the lexical status of the Mandarin gap syllables in Southern Min and Hakka, two of the major dialects in Taiwan. It is found that thirty-three syllables out of the 108 gap syllables are real syllables in Southern Min (despite the fact that there might be subtle difference in the actual pitch of the tone). As in the case of Hakka, sixty of the Mandarin gap-syllables were real syllables in Hakka (also disregarding the subtle difference in the actual pitch of the tone).

The nonwords were composed of early-acquired consonants only (Cheung, 2000; Zhu & Dodd, 2000) to avoid articulatory difficulties. The NWR task contained 36 disyllabic nonce words and 36 disyllabic gap words. Each type of nonword was equally divided into the one-word list, the two-word list, and the three-word list (see Appendix D).

In the task, the child was told to repeat some strange words. The task was administered in the form of a live presentation for an optimal condition for the perception of the stimuli (Chiat & Roy, 2007). The repetition of nonce words was administered first.

Concerning that the younger children's shorter attention span and their need for more motivation to engage in this task, we not only downsized the number of nonword items, but also adopted and modified the testing procedure of Stokes and Klee (2009) for children at age 2 and age 2.5. In the testing, the child was asked to imitate each nonword said by the experimenter, and then rolled a ball down a chute as a reward, whether the word was correctly imitated or not. At age 2, the child was required to complete only half of the stimuli in the one-word list and the two-word list of each NWR task (i.e. 12 disyllabic nonce words and 12 disyllabic gap words). At age 2.5, the child had to complete half of the stimuli in the one-word list, the two-word list and the three-word list of each NWR task (18 disyllabic nonce words and 18 disyllabic gap words). For children at and above age 3, they were instructed to repeat after puppies for the all the nonwords.

The recall accuracy of the nonwords was scored at the syllable level. A syllable received one point if it was correctly repeated. Any omission, deletion, addition, or substitution would be considered errors because they signaled an inability to encode or maintain the original phonological representation. However, children's systematic substitutions, as observed in their performance in the productive phonology task, were counted as correct to minimize the effect of articulatory constraint on their NWR performance. The maximum score was 72 for the nonce word repetition or the gap word repetition.

3.2.7 Leiter International Performance Scale-Revised (Leiter-R). The Leiter-R (Roid & Miller, 1997) is an intelligence test designed to assess nonverbal cognitive abilities in children and adolescent aged 2 to 20. It includes two groupings of subtests: the Visualization and Reasoning Battery and the Attention and Memory Battery. For the purpose of this study, we used 4 subtests in the Visualization and Reasoning Battery as a rapid estimate of global intellectual ability, including Figure Ground (FG), Form Completion (FC), Repeated Pattern (RP) and Sequential Order (SO). The raw scores on each of the measure would be converted to normalized scale scores. The composite of the subtest scaled scores was regarded as a Brief IQ estimate.

3.3. Procedures

The child was tested in a quiet room in the kindergarten or a sound-attenuated testing booth. Since there were a number of tests the child had to complete in each testing, each testing would be divided into two sessions. The two sessions would be controlled to be separated for no more than two weeks. Each of the session lasted for about 30 to 40 minutes.

In the first session, the child was tested with vocabulary (PPVT-R for children above age 3), output phonology, discrimination, and nonce-word repetition. In the second session, the child was tested with vocabulary (REVT for children above age 3), gap-word repetition, and Leiter-R (only at Time 2). For children below age 3, the parents were asked to fill in the MCDI-T form. They had to check the child's productive vocabulary at the 1st session, and the child's receptive vocabulary at the 2nd session. The procedure demonstrated here was the regular arrangement. However, the order of the tasks would change in accordance with children's cooperative situation in the testing in order to achieve children's optimal performance.

3.4 Analyses

A set of analyses were conducted to answer the research questions. Our research questions concerned the interactions among vocabulary size, phonological capacities and NWR. To answer this question, we first examined the growth in vocabulary, phonology, and NWR in each of the three age cohorts. The developments were described quantitatively and qualitatively. Individual variation in developmental patterns was also noted. Then, regression analyses were conducted in each age cohort to examine (1) the contribution of phonological capacities and receptive vocabulary knowledge to NWR performances, and (2) the predictability of NWR performance to expressive vocabulary knowledge. Finally, we delineated the developmental trajectories of vocabulary size, phonology capacities, and NWR based on the data across age groups (from age 2 to age 5). Given that children might demonstrate considerable variation in NWR performance, and that phonological capacities and vocabulary size might have an effect on the developmental changes of NWR (including the initial ability and the growth rate), the hierarchical linear model approach was conducted to present a more global profile of the interactions among these three factors in the development from age 2 to age 5.

Chapter 4 Vocabulary, Phonology, and NWR of the Age 2 Cohort

In this chapter, we examined the vocabulary development and phonological development of children in the age 2 cohort, and also the relationship between these two constructs in this age range. Children's performances in the nonce word repetition and the gap word repetition were also inspected. The performance data and the correlation relationship of the measures in discussion would be presented in the main text. However, a more detailed descriptive statistics of all the measures and an overall correlation matrix were provided in Appendix E(1) and F(1), respectively. In the end of this chapter, we examined the relative contribution of phonological capacities and vocabulary development to NWR performances. Preliminary analyses showed that there was no significant difference between genders; therefore, gender was not included in the following analyses, unless it was particularly specified.

4.1 Participants

Children in this cohort were recruited at age 2 (age range = 23.83-26.63 months), and tested every six months. Therefore, they were tested at age 2 (Time 1), age 2.5 (Time 2) and age 3 (Time 3), respectively. These children were recruited by posts on parenting websites or by word-of-mouth. Twenty-four children were tested at Time 1. At Time 2, we lost two participants because their parents were unable to cooperate with the testing schedule. At Time 3, all the remaining participants returned. However, it should be noted that one of the boy children (YHW) refused to perform almost all the tasks that required speech production, though he was quite cooperative in the past two testing points. As a result, his data were missing in some of the tasks at Time 3. Participant information was summarized in Table 4.1.

Table 4.1

Information of the Child Participants in the Age 2 Cohort

Test time	Time 1	Time 2	Time 3
N	24 (12 ♂)	22 (12 ♂)	22 (12 ♂)
Average age (month)	24.92	31.10	36.95
Age range	23.83-26.63	29.93-32.93	35.57-39.53

4.2 Vocabulary Development

Children in the age two cohort were still too young at their first testing time to receive any standardized tests. Therefore we measured their vocabulary knowledge, both the receptive vocabulary and the expressive vocabulary, with the MCDI-T checklist (Liu & Tsao, 2010). The parents were required to fill in the MCDI-T checklist in all the three testing sessions. At age 2.5, we included the standardized test PPVT-R (Lu & Liu, 1994) to measure children's receptive vocabulary, because we believed that children at this age might be old enough to appropriately react to a standardized language receptive task, though a language expressive task might still be too demanding. It was until when children reached age 3 that we included both standard tests on the receptive vocabulary (PPVT-R) and the expressive vocabulary (REVT, Huang et al., 2011) to assess children's vocabulary knowledge, in addition to the parent-report measure of early vocabulary (MCDI-T). For the convenience of reference, we would specify the receptive vocabulary score measured with the MCDI-T as MCDI-$_{\text{RECEPTIVE}}$, and the expressive vocabulary score measured with the MCDI-T as MCDI-$_{\text{EXPRESSIVE.}}$ Children's growth in each task would be examined respectively in the following sections.

4.2.1 The expressive vocabulary and receptive vocabulary as measured by MCDI-T. Since both children's receptive vocbaulary and expressive vocabulary were inspected with the MCDI-T checklist in this cohort, we first examined the children's performances in these two vocabulary constructs, and the relationship between these two. Children's average performances in MCDI-$_{\text{RECEPTIVE}}$ (the line with black dots) and MCDI-$_{\text{EXPRESSIVE}}$ (the line with hollow dots) across time were graphed in Figure 4.1.

In terms of the performances in MCDI-$_{\text{RECEPTIVE}}$, children at age 2 obtained an average score of 471.96 (SD = 99.77), with scores ranging between 232 and 628. At age 2.5, children attained an average score of 638.32 (SD = 60.23), with the range of 485 and 694. At age 3, children had an average score of 686.05 (SD = 10.99). Children's scores ranged between 651 and 696, suggesting a ceiling effect in using MCDI-T to assess age 3 children's receptive vocabulary size. In fact, this task has been designed for children below age 3, so it might not effectively reflect the actual variation in children's receptive vocabulary at age 3.

Correlation analyses were conducted to examine the relationship among children's receptive vocabulary size at each time points (Table 4.2). MCDI-$_{\text{RECEPTIVE}}$ at age 2 was associated with MCDI-$_{\text{RECEPTIVE}}$ at age 2.5 ($r = .59$, $p < .01$) and MCDI-$_{\text{RECEPTIVE}}$ at age 3 ($r = .49$, $p < .05$). However, the correlation between MCDI-$_{\text{RECEPTIVE}}$ at age 2.5 and MCDI-$_{\text{RECEPTIVE}}$ at age 3 only revealed a borderline significance ($r = .38$, $p = .083$), probably because children were about to reached the maximum score of the checklist at age 2.5 and age 3. The smaller variation in the children's performances at these two time points led to the weak correlations.

With regards to the performance in MCDI-$_{\text{EXPRESSIVE}}$, children at age 2 received an average score of 304.87 (SD = 109.89), ranging between 152 and 565. At age 2.5, children attained an average of 567.59 (SD = 85.47), a remarkable increase in size. The score range (409-686) was smaller compared to the score range in age 2. Children seemed to manifest remarkable growth in their expressive vocabulary in this half a year. At age 3, children's scores ranged between 574 and 696, with an average score of 665.91 (SD = 28.83), which almost reached the maximum score of this checklist. Therefore, MCDI-T might not effectively reflect actual variation in children's expressive vocabulary at age 3.

Then, we examined the correlations among children's MCDI-$_{\text{EXPRESSIVE}}$ score at each time point (Table 4.2). Correlations were found only between MCDI-$_{\text{EXPRESSIVE}}$ at age 2 and age 2.5 (r = .59, p < .01). MCDI-$_{\text{EXPRESSIVE}}$ at age 3 was not associated with either MCDI-$_{\text{EXPRESSIVE}}$ at age 2 (r = .26, p > .05) or MCDI-$_{\text{EXPRESSIVE}}$ at age 2.5 (r = .36, p > .05).

Figure 4.1

Children's Average Performances in MCDI-$_{\text{RECEPTIVE}}$ and MCDI-$_{\text{EXPRESSIVE}}$ Across Time

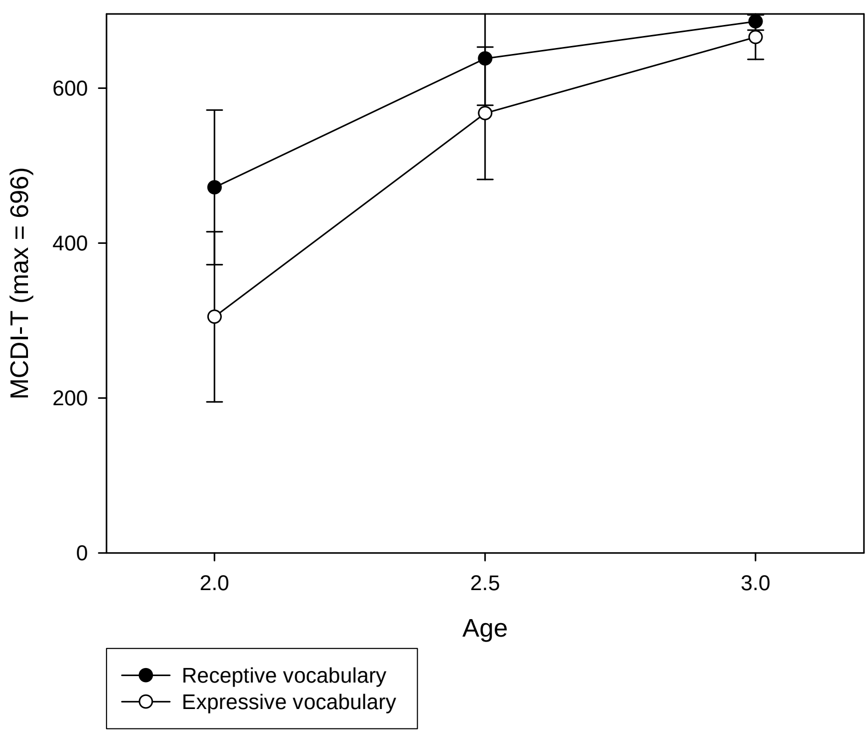


Table 4.2

Correlation Matrix of MCDI-RECEPTIVE and MCDI-EXPRESSIVE Across Time

	MCDI-R @age 2.5	MCDI-R @age 3	MCDI-E @age 2	MCDI-E @age 2.5	MCDI-E @age 3
MCDI-R @age 2	.59**	.49*	.77**	.71**	.37
MCDI-R @age 2.5		.38	.35	.81**	.33
MCDI-R @age 3			.29	.41	.86**
MCDI-E @age 2				.59**	.26
MCDI-E @age 2.5					.36

The correlations between MCDI-RECEPTIVE and MCDI-EXPRESSIVE were examined (Table 4.2). MCDI-EXPRESSIVE at each time point was highly correlated with the concurrent MCDI-RECEPTIVE (r = .77 - .86, $p < .01$). Therefore, there was close relationship between children's concurrent receptive vocabulary and expressive vocabulary. Cross-time correlations were only found between MCDI-RECEPTIVE at age 2 with MCDI-EXPRESSIVE at age 2.5 (r = .71, $p < .01$), which appeared to suggest a strong association between receptive vocabulary size and the expressive vocabulary size half a year later.

A two-way repeated measures ANOVA was conducted to compare children's performances of MCDI-EXPRESSIVE and MCDI-RECEPTIVE across time. In addition to the significant effect of time ($F(2, 42) = 153.11$, $p < .001$, $\eta^2 = .88$), the results demonstrated a significant main effect of vocabulary ($F(1, 21) = 140.26$, $p < .001$, $\eta^2 = .87$), and a significant interaction effect of vocabulary and time ($F(2, 42) = 46.63$, $p < .001$, $\eta^2 = .689$). The interaction effect was caused by the faster growth rate of expressive vocabulary in comparison with the growth rate of receptive vocabulary, as shown in Figure 4.1.

Planned analyses were conducted to examine children's growth in MCDI-$_{\text{RECEPTIVE}}$ size. The results showed that children's receptive vocabulary size increased from age 2 to age 2.5 ($t(21) = 9.47$, $p < .001$), and from age 2.5 to age 3 ($t(21) = 3.93$, $p < .001$).

Parallel analyses were also conducted to examine whether children demonstrated significant growth in MCDI-$_{\text{EXPRESSIVE}}$. The results revealed that children's expressive vocabulary size increased from age 2 to age 2.5 ($t(21) = 13.38$, $p < .001$), and from age 2.5 to age 3 ($t(21) = 5.77$, $p < .001$).

4.2.2 Receptive vocabulary size as measured by PPVT-R. Children in this age cohort were tested with the standardized test PPVT-R at age 2.5 (Time 2) and age 3 (Time 3). At age 2.5, the average score they obtained in this task was 18.82 (SD = 6.55, range 10-37). Half a year later, they attained an average score of 29.86 (SD = 8.94, range 13-47). Correlation analysis revealed that children's PPVT-R at age 2.5 and PPVT-R at age 3 were significantly correlated, $r = .62$, $p = .002$. The results of a paired *t* test revealed a significant growth effect, $t(21) = 7.34$, $p < .001$. Children's performance in PPVT-R at age 3 was significantly better than their performance at age 2.5.

Since children in this cohort were assessed with two receptive vocabulary tests, i.e. MCDI-$_{\text{RECEPTIVE}}$ and PPVT-R, we examined the correlations between children's performance in these two tasks. The results showed that PPVT-R at each time point was only associated with its concurrent MCDI-$_{\text{RECEPTIVE}}$ score. For example, PPVT-R at age 3 was only significantly associated with MCDI-$_{\text{RECEPTIVE}}$ at age 3 ($r = .49$, $p < .05$). The correlation between PPVT-R at age 2.5 and MCDI-T at age 2.5 was close to borderline significance ($r = .39$, $p = .073$). In addition, MCDI-$_{\text{RECEPTIVE}}$ at age 2 was associated with neither PPVT-R at age 3 nor PPVT-R at age 2.5. In other words, children's MCDI-$_{\text{RECEPTIVE}}$ scores could not predict children's PPVT-R scores.

4.2.3 Expressive vocabulary size as measured by REVT. In addition to measuring the age two cohort children's expressive vocabulary with the parent-report checklist, we assessed children's expressive vocabulary with the standardized test REVT. Considering that this task required children's expressive performance, and would be too demanding for children below age 3, we included this task when children in this cohort reached age 3. Children gained an average score of 34.57 (SD = 9.77, range = 15-46, percentile range = 52-99%) in this task.

We examined how REVT was associated with $\text{MCDI-}_{\text{EXPRESSIVE}}$. The results showed that REVT was correlated with $\text{MCDI-}_{\text{EXPRESSIVE}}$ at age 3 ($r = .54$, $p < .05$). Nevertheless, its correlations with $\text{MCDI-}_{\text{EXPRESSIVE}}$ at age 2 ($r = .40$, $p = .073$) and $\text{MCDI-}_{\text{EXPRESSIVE}}$ at age 2.5 ($r = .40$, $p = .076$) approached but fell short of significance.

$\text{MCDI-}_{\text{EXPRESSIVE}}$ and REVT were modestly associated, because both assess children's ability in verbally labeling objects. However, REVT provides a more advanced measurement of children's expressive vocabulary ability by additionally assessing vocabulary categorization, definition, and reasoning. Therefore, it measures not only the breadth, but also the depth of children's expressive vocabulary.

As stated in the previous section, MCDI-T checklist might not reliably reflect individual variation in vocabulary knowledge in age 3 children. Therefore, we inspected the relationship between expressive vocabulary and receptive vocabulary which were measured with the standardized tests. Correlation analyses showed that REVT at age 3 was significantly correlated with PPVT-R at age 2.5 ($r = .47$, $p < .05$), and PPVT-R at age 3 ($r = .57$, $p < .01$). The cross-time correlation seemed to suggest that the receptive vocabulary may serve as the foundation for the development of expressive vocabulary, just as we observed in the correlation between $\text{MCDI-}_{\text{RECEPTIVE}}$ at age 2 and $\text{MCDI-}_{\text{EXPRESSIVE}}$ at age 2.5. However, since we did not measure REVT at age 2.5, we could not verify whether there was a cross-time correlation between REVT at age 2.5

and PPVT-R at age 3.

4.2.4 Summary of vocabulary development from age 2 to age 3. Children in this age showed dramatic improvements in their vocabulary size, particular during age 2 to age 2.5. As measured by the MCDI-T checklist, during this half a year, children's receptive vocabulary had an average increase of 28 words per month, and their expressive vocabulary an average increase of 48 words per month. However, the flatter growth curves during age 2.5 to age 3 might not be a genuine slowdown in children's vocabulary acquisition rates, but the consequence of the limit of the checklist.

The use of the MCDI-T checklist for the assessments of vocabulary size in both the receptive and expressive aspects allowed us to make a direct comparison of these two aspects of vocabulary. The results showed that children tended to have larger receptive vocabulary size than expressive vocabulary size, just as found in previous studies (Clark, 1993; Ingram, 1974). Regarding the correlations between these two vocabulary aspects, we found that MCDI-$_{\text{RECEPTIVE}}$ at age 2 was associated with MCDI-$_{\text{EXPRESSIVE}}$ at age 2.5. The association between the earlier receptive vocabulary size and the subsequent expressive vocabulary size was also replicated in the significant correlations between PPVT-R at age 2.5 and REVT at age 3. It is likely that receptive vocabulary could serve as the foundation for the development of expressive vocabulary, at least in this age range.

Analyses in the previous sections demonstrated the average trend of children's vocabulary growth. Close inspection on the idiosyncrasies of children's performances revealed considerable variation in children's onset vocabulary knowledge, though all the children experienced growth in the size of vocabulary knowledge during this age range. It was of interest whether this difference in vocabulary knowledge would be related to performance in nonword repetition. We would investigate this hypothesis in the last section of this chapter.

We assessed the vocabulary knowledge of the children in this cohort with both the parent-report checklist MCDI-T (in all three testing sessions) and the standardized tests PPVT (at age 2.5 and age 3) and REVT (at age 3 only). Though correlations were found between the parent-report checklist and the standardized tests, the magnitude of the correlations suggested a modest-to-weak relationship between them. The differences between the MCDI-T checklist and the other two standardized tests should be noted.

4.3 Phonological Development

We examined children's phonological development in terms of their production of onsets and rhymes, and also their discrimination ability at the phonemic level.

4.3.1 Productive phonology

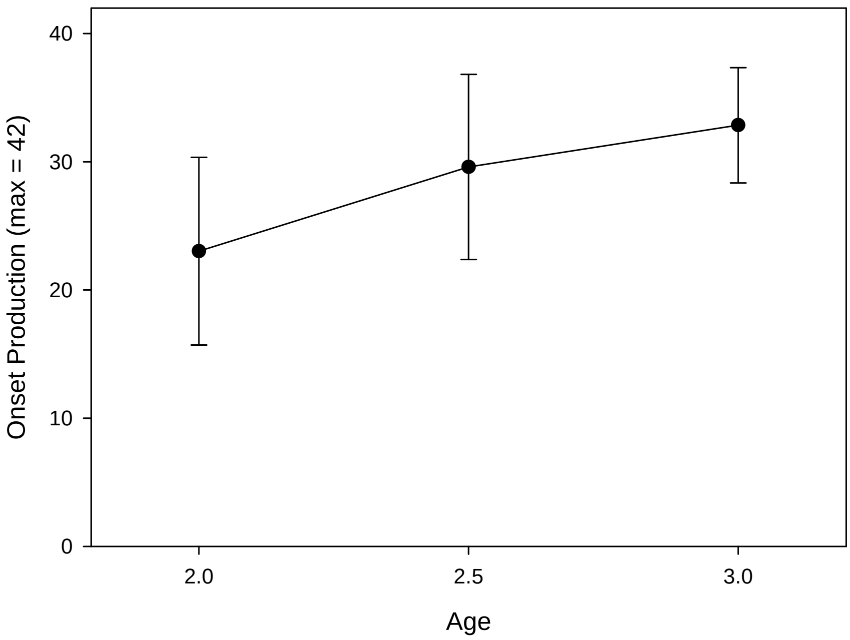
4.3.1.1 Onset production. We examined children's production of 21 onset consonants in both the word-initial position and the non-word-initial position (maximum score = 42). Children at age 2 attained an average score of 23.03 (SD = 7.32), and the scores ranged between 9.51 and 35.31. At age 2.5, children attained a score of 29.59 (SD = 7.22, range: 16.63 – 39.50). At age 3, when children reached age 3, their onset production accuracy achieved an average score of 32.85 (SD = 4.50, range 20.71 – 39.29). Children's average performances across time were plotted in Figure 4.2. Correlation analyses showed that onset production at the three time points were correlated with each other (*r* values ranged between .56-.78).

A one-way ANOVA was conducted to compare children's onset production across three time points. It demonstrated a significant main effect of time, $F(2, 40) = 42.48$, $p < .001$, $\eta^2 = .68$. Planned comparisons on onset production at each of the two time points revealed that there was significant growth both from age 2 to age 2.5 ($t(21)$ =

6.72, $p < .001$), and from age 2.5 to age 3 ($t(20) = 2.99$, $p < .01$). The growth exhibited a linear trend, $F(1, 20) = 73.23$, $p < .001$, $\eta^2 = .79$.

Figure 4.2

Children's Average Performances in Onset Production Across Time



We further examined the ages at which children reached 75%, 50-75%, or below 50% accuracy in onset phoneme production (Table 4.3). The criterion of phoneme stabilization follows the one proposed in Zhu and Dodd (2000). That is, a phoneme is considered to be stabilized when 75% of the children in the particular age group reached an accuracy rate of 66.67% in its production. With this standard we identified the phonemes that had been stabilized at each testing time point. At Time 1, when children were at age 2, they gained mastery of unaspirated oral stops *b*, *d*, *g*, and nasal stops *m*, *n*. At age 2.5, children had learned to master the articulation of other manners, including the approximant *l*, the fricative *h*, and the affricate *j*. Children's production of the aspiration feature was not yet mature until age 3. At this time, children gained mastery of aspirated oral stops, and also phonemes with other manners, such as the fricatives *x* and *s*. The sounds which remained to be difficult for children at age 3 included the labiodental fricative *f*, the affricates *q*, *z, c* and also the retroflexes *zh, ch, sh, r*.

In this age range, children's phonological production ability was still in development, which could be reflected in the variability of phonemic errors they made. However, in general the most common error types in this age range included syllable initial deletion (e.g. *zui3-ba1* "mouth" → *ui3-ba1*), stopping (e.g. *fei1-ji* "plane" → *bei1-ji1*), deaspiration (e.g. *kuai4-zi0* "chopsticks" → *guai4-zi0*), and etc.

Table 4.3

Children's Development in Onset Production

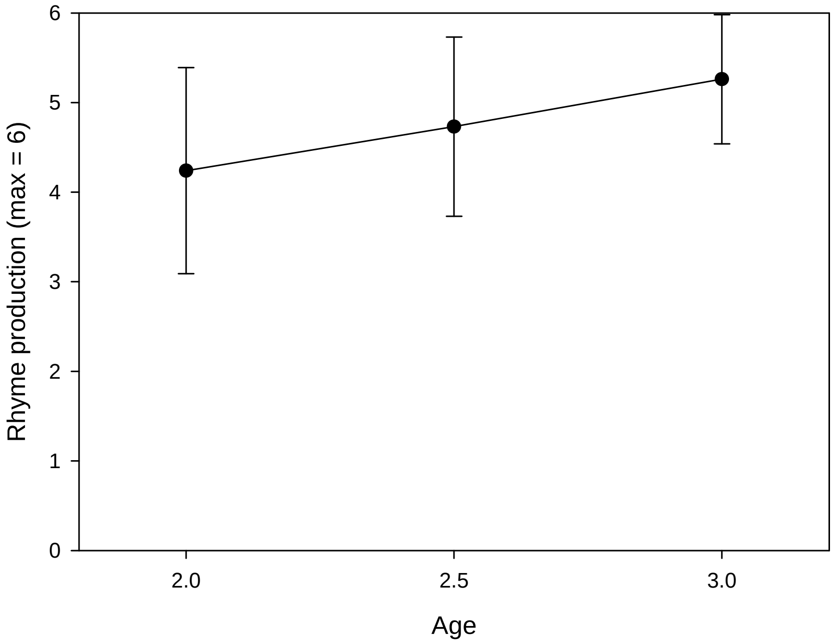
Pass rate[a]	Age 2	Age 2.5	Age 3
> 75% (stabilized)	*b, d, g, m, n*	*b, d, g, m, n, l, h, j*	*b, d, g, m, n, l, h, j, p, t, k, x, s*
50-75%	*k, h, j, q*	*p, f, t, k, q, x, z, c, s*	*f, q, z, c*
< 50%	*p, f, t, l, x, zh, ch, sh, r, z, c, s*	*zh, ch, sh, r*	*zh, ch, sh, r*

Note. [a] % of children passed the 66.7% accuracy rate

4.3.1.2 Rhyme production. In addition to the production of onset consonants, we also examined children productions of six rhyme structures, including V, VG, GV, VN, GVN, and GVG (maximum score = 6). Children attained an average score of 4.24 (SD = 1.15) in the rhyme production at age 2, and the scores ranged between 2.40 and 5.75. At age 2.5, they attained an average score of 4.73 (SD = 1.00, range: 2.88 – 5.88). At age 3, their scores ranged between 3.82 and 5.99, with an average score of 5.26 (SD = 0.72). The average performances were graphed in Figure 4.3. Correlation analyses showed that rhyme production at the three time points were correlated with each other (*r* values ranged between .62 - .85).

Figure 4.3

Children's Average Performances in Rhyme Production Across Time



A one-way ANOVA was conducted to compare children's rhyme production across three time points, and revealed a significant main effect of time, $F(2, 40) = 24.67$, $p < .001$, $\eta^2 = .55$. Planned comparisons on each of the two time points demonstrated significant growth in rhyme production from age 2 to age 2.5 ($t(21) = 3.98$, $p = .001$), and from age 2.5 to age 3 ($t(20) = 4.43$, $p < .001$ $t(20) = 4.43$, $p < .001$). The trend analysis showed that the growth exhibited a linear trend, $F(1, 20) = 31.92$, $p < .001$, $\eta^2 = .62$.

Further, we examined children's development in the production of each rhyme structure in Mandarin, particularly the rhyme structures that have stabilized at the particular age. Again, Zhu and Dodd's (2000) criterion of phoneme stabilization was applied. It was found that only the rhyme structure of V, such as the simple vowels /i/, /u/, /ə/ and others, was stabilized in the age 2 children. When children grew to age 2.5, they gradually mastered the diphthongs VG and GV, and the triphthong GVG, which are the more complex rhyme structures. Children's production of VN and GVN structures were not stabilized even when children reached the age of 3, which suggested that final nasal codas are difficulty for children to master.

4.3.1.3 The relationship between the productions of onset and rhyme. Correlation analysis was conducted to examine the association between the productions of onset and rhyme structure. The results were summarized in Table 4.4. Onset production at age 2 was found to be highly associated with rhyme structure production at the same period of time ($r(24) = .75$, $p < .01$). However, the associations of the concurrent onset production and rhyme structure production diminished in the subsequent testing. Additionally, onset production at age 2 was found to be modestly associated with rhyme structure production at age 2.5 ($r(22) = .55$, $p < .01$). Rhyme structure production at age 2 was found to be associated with onset production at age 2.5 ($r(22) = .43$, $p < .05$), and at age 3 ($r(21) = .49$, $p < .05$). In general, we found that the production of onsets and rhymes only shared strong correlation at age 2. Their association weakened as children grew up.

4.3.2 Word discrimination. This measure assessed children's word discrimination ability at the phoneme level. Since the number of items children had to complete differed across three testing points, the scores were first transformed into percentages for the subsequent analyses.

Table 4.4

Correlation Matrix of the Production of Onsets and Rhymes Across Time

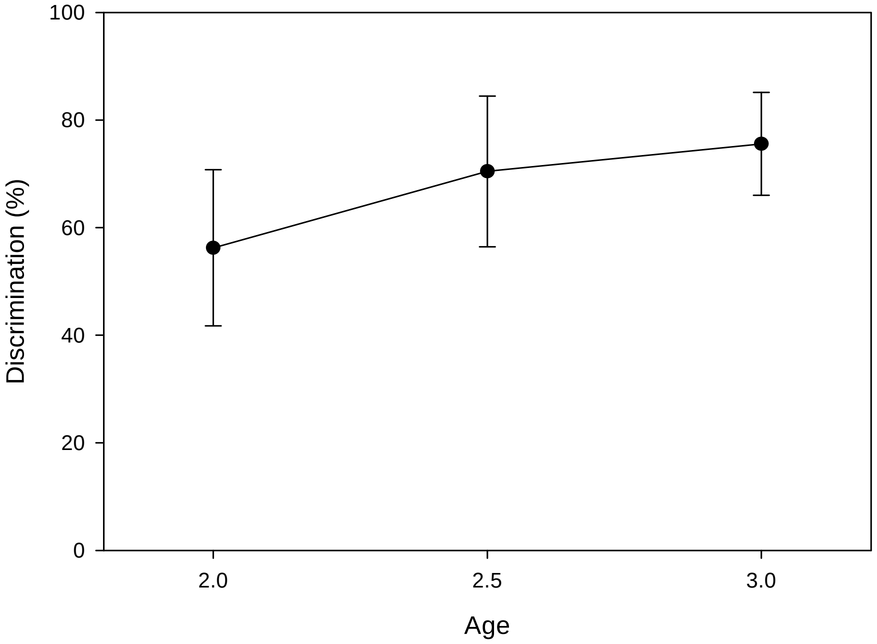
	Onset @age 2.5	Onset @age 3	Rhyme @age 2	Rhyme @age 2.5	Rhyme @age 3
Onset @age 2	.78**	.56**	.75**	.55**	.36
Onset @age 2.5		.67**	.43*	.34	.29
Onset @age 3			.49*	.36	.29
Rhyme @age 2				.85**	.62**
Rhyme @age 2.5					.80**

It has to be noted that at age 2, not all the children performed this task. Among the four children who failed to perform the task, two had difficulty following the instruction, one lost patience in watching the PowerPoint presentation, and the other one was too shy to respond. However, at age 2.5 and age 3, all the attending children completed the task.

At age 2, children attained an average accuracy rate of 56.25% (SD = 14.53), with the scores ranging between 33.33% and 91.67%. Half a year later, children gained an average accuracy rate of 70.45% (SD = 14.02, range: 50–100%). At age 3, when children reached age 3, they had an average accuracy rate of 75.57% (SD = 9.56), ranging between 58.33% and 91.67%. Children's discrimination performances across the three time points were displayed in Figure 4.4. The correlation analyses showed that discrimination performances at the three time points were not correlated with each other (p > .05).

Figure 4.4

Children's Average Performances in Word Discrimination Across Time

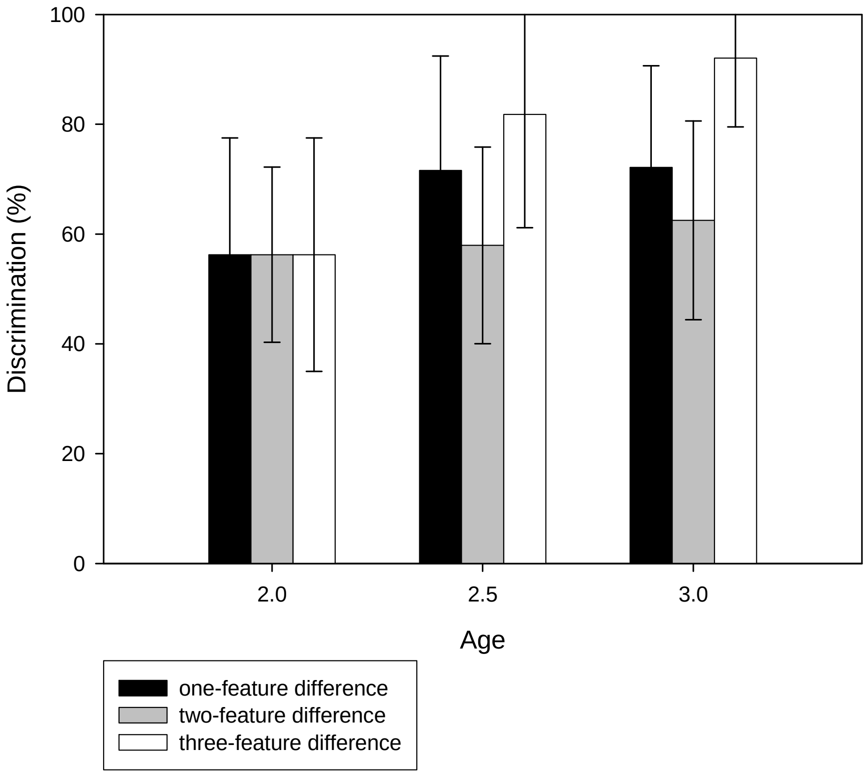


4.3.2.1 Sensitivity for different degrees of phonetic contrast. In this task we included minimal pairs that were of different degrees of phonetic contrast. We manipulated the contrasting onset consonants to differ in one phonetic feature (1-*f* difference), two phonetic features (2-*f* difference), or three phonetic features (3-*f* difference), aiming to examine the specificity of children's discrimination ability. Their average performances were presented in Figure 4.5. The one-sample *t* tests were first conducted to examine whether children's performance in discriminating the contrast pairs at each time point was above chance level. The results demonstrated that at age 2 children's discrimination performances in all contrast pairs approached borderline significance ($p < .083$, one-tailed). Their performances were above chance level at age 2.5 and at age 3 (all $p < .05$, one-tailed).

A repeated measures ANOVA was conducted with the variables of time and degree of phonetic contrast on children's discrimination performance. The results revealed a significant interaction effect ($F(4, 68) = 6.25$, $p < .001$, $\eta^2 = .27$), and both the main effects were significant (time: ($F(2, 34) = 9.93$, $p < .001$, $\eta^2 = .37$); contrast degree: ($F(2, 34) = 19.26$, $p < .001$, $\eta^2 = .53$)).

Figure 4.5

Children's Average Performances in Discriminating Different Degrees of Phonetic Contrasts



Three one-way ANOVAs were conducted to examine children's sensitivity for different degrees of phonetic contrast at each time point. Analysis on the data at Time 1 showed no significant effect of phonetic contrast, $F(2, 38) = .02$, $p > .05$, $\eta^2 = .001$. In other words, children at age 2 showed equal sensitivity for the phonetic contrasts of 3-*f* difference, 2-*f* difference, and 1-*f* difference. However, a significant phonetic contrast effect was found in Time 2 ($F(2, 42) = 10.36$, $p < .001$, $\eta^2 = .33$) and in Time 3 ($F(2, 42)$

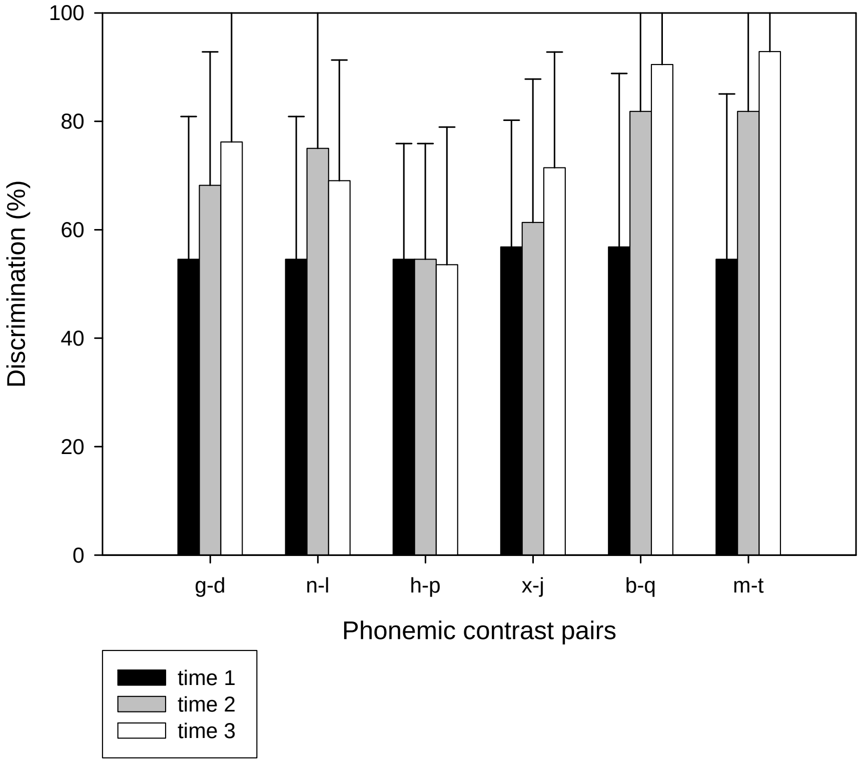
= 20.93, $p < .001$, $\eta^2 = .50$). At Time 2, children were more sensitive to 3-*f* difference contrasts than to 1-*f* difference contrasts, while they showed the poorest sensitivity to 2-*f* difference contrasts. At Time 3, children were still more sensitive to 3-*f* difference contrast, whereas they demonstrated equal sensitivity to the contrasts of 1-*f* difference and the contrasts of 2-*f* difference.

We also conducted additional one-way ANOVAs to inspect children's development in their discrimination of each type of phonetic contrast. The results showed that children had greater improvement in discriminating phonemic pairs with prominent phonetic contrasts (i.e. those with 3-*f* difference), $F(2, 34) = 24.86$, $p < .001$, $\eta^2 = .59$, especially in the first half of the year, $t(17) = 5.66$, $p < .001$. On the other hand, children sensitivity for less prominent phonetic contrasts did not manifest significant improvements over time [1-*f* difference: $F(2, 34) = 2.50$, $p = .097$, $\eta^2 = .13$; 2-*f* difference: $F(2, 34) = .446$, $p = .641$, $\eta^2 = .026$]. Though there was a noticeable increase in children's discrimination of phonemic pairs with 1-*f* difference from age 2 to age 2.5, as shown in Figure 4.5, the increase only approached significance, $t(17) = 2.09$, $p = .052$.

We had hypothesized that the more dissimilar the phonetic contrasts were, the easier they could be discriminated. However, this hypothesis was only partially born out. It was unexpected to find that children had better discrimination performance in the contrasts of 1-*f* difference than in the contrasts of 2-*f* difference. We looked into children's performance in each of the phonetic contrast pairs (Figure 4.6), and found that the contrast pairs of 2-*f* difference, i.e. *h-p* and *x-j*, had the poorest performances throughout almost all the testing points. There were some possible explanations for this finding.

Figure 4.6

Children's Average Performances in Each of the Contrast Pairs



First, it was possible that children's discrimination performance was affected by the frequencies of the word stimuli, i.e. word frequency. The more frequent a word occurs in speech, the more likely that the child could construct a well-specified and robust representation of that specific word. Therefore, it is more probable for the children to detect a mispronunciation of that specific word. To verify this hypothesis, we checked up the frequencies of our word stimuli in the Chinese Spoken Wordlist, which was derived from the transcripts of spoken data in the spoken Taiwan Mandarin corpora, developed by the Academia Sinica, Taiwan (Tseng, 2013). However, concerning that this spoken data consist of adult-to-adult conversations, which might not resemble the status of these words in adult-to-child conversations, we also checked up the frequencies of the word stimuli in the spontaneous spoken data in Taiwan Corpus of

Child Mandarin (TCCM, Cheung, 1998). The frequencies of these words in these two corpora were summarized in Table 4.5. If the word frequency hypothesis is supported, then the target words in the 2-*f* difference contrasts (i.e. *hu3* and *xi3*) should have fewer frequencies than target stimuli in other contrast types, specifically target words with 1-*f* difference contrasts (i.e. *gou3* and *niao3*). This explanation seemed to account for the performance in discriminating *h* vs. *p*, but not the performance in discriminating *x* vs. *j*, which is also frequent in spoken data.

In addition to the effect of the absolute frequencies of the target words, it is possible that children's discrimination of the target words could be affected by the frequencies of the distractors. That is, if one distractor is particularly less frequent than its target, it would be more difficult for the child participant to tell them apart. However, the distractors used in this task were almost equally less frequent in speech (Table 4.5). Therefore, this account failed to explain why children demonstrated the poorest performance in discriminating the 2-*f* difference contrasts.

Table 4.5

Frequencies of the Target Word Stimuli and the Distractors in Two Different Spoken Corpora

Target word	*gou3*	*niao3*	*hu3*	*xi3*	*biao3*	*mao1*
AS spoken corpus	29	4	3	31	1	7
TCCM	70	49	0	111	4	23
Distractor	*dou3*	*liao3*	*pu3*	*ji3*	*qiao3*	*tao1*
AS spoken corpus	1	0	0	8	8	3
TCCM	0	0	0	9	0	0

Children's discrimination of the sounds could be affected by the age of acquiring the phonemes (AoA). According to our analysis on this group of children's onset production, the target onset phonemes in the 2-*f* difference contrast, i.e. *h* and *x*, were stabilized at age 2.5 and age 3, respectively. On the other hand, the target onset phonemes in the 1-*f* difference contrast and 3-*f* difference, i.e. *g, n, b,* and *m*, were already stabilized at age 2. Since these sounds were stabilized at the earlier age, children were more familiar with these sounds and could be able to discriminate them from other similar phonemes.

Children's superior performance in discrimination 1-*f* difference contrasts than 2-*f* difference contrasts could also be accounted for by the relative acoustic saliency of the phoneme contrasts. Phonemic contrasts that are acoustically more salient could be more detectable than those that are less salient. The acoustic saliency here is referring to the persistence of the distinctive cue of the phonetic contrast pairs. For example, the *g-d* distinction was easier to detect because the distinctive acoustic cue remained to be detectable till their transition to vowels. Also, the *n-l* pair contrasts in the weaker F3 of *n*, and the acoustic cue lasts at least 50ms. However, the phonetic contrast pairs with 2-*f* difference, i.e. *h-p* and *x-j*, had swifter presence of the distinctive cue. These two pairs of sounds were most distinctive at the burst. Moreover, the syllable lengths of these two pairs of stimuli were shorter than other pairs of stimuli. If young children did not notice the burst, or their attention drifted a little while, then they could easily miss the cue and fail to discriminate the sounds.

Our findings above demonstrated that the age 2 cohort children's discrimination performances could be affected by not only the phonetic contrast degree of the stimuli, but also the word frequency, AoA, and acoustic saliency of the stimuli. However, in general, the results showed that while children at age 2 still had a chance performance at discriminating phonetic contrasts at the phonemic level, their sensitivity for these contrasts, particularly those with prominent contrast, grew with age.

4.3.3 Summary of the phonological development from age 2 to age 3. The results revealed that children's phonological capacity at age 2 were still far from adult-like. Most of the age 2 children developed stabilized production of single vowels, unaspirated stops and nasals, while the production of other consonants and rhymes remained under-developed. However, children also demonstrated appreciable variance in their phonological production ability. While some children had almost all the phonemes in place, there were some still at the beginning stage of word production. In terms of children's discrimination at phoneme level, their performance was about at chance at this age.

Half a year later, most children gained mastery of other rhyme types that are structurally more complex, including VG, GV, and GVG. Almost all the rhymes became stabilized except for the rhymes with a nasal coda (i.e. VN and GVN). With regard to onset consonants, children had mastered the articulation manners other than stops, such as the fricative *h*, the glide *l*, and the affricate *j*. However, children were still not able to control aspiration and manner of retroflex. These findings together exhibited the development in children's motor planning ability in carrying out more complex articulatory gestures. In addition, we also saw a prominent growth in children's discrimination ability. They showed growth both in discriminating minimal pairs with great contrasts and those with less prominent contrasts.

At age 3, children finally had stabilized performance in controlling the feature of aspiration in onset production. The aspirated stops *p, t,* and *k* became stabilized, as well as the aspirated fricatives *s* and *x*. Children's performance in rhyme production and discrimination retained at the similar levels as in age 2.5.

The findings in general revealed that children from age 2 to age 3 showed steady growth in their phonological capacities. Qualitative analyses suggested a greater leap during age 2 to age 2.5. In this period of time, most children had gained mastery of the production of more complex rhyme structures, such as VG, GV, and GVG. Moreover, they had acquired manners of articulation other than stops in onset production (i.e. glides, fricatives, and affricates), which reflected their enhanced ability in carrying out complex articulatory gestures.

Findings in the discrimination task also demonstrated children's refining ability in discriminating phoneme-level contrasts. For example, even though children at age 2 had random performances in discriminating contrasts in onset phonemes, their sensitivity for the phonetic contrasts increased with age. There was a noticeable increase in children's discrimination of phonemic pairs with 3-*f* difference, especially during the age 2 to age 2.5 span. These improvements, either in production or in perception, altogether implicated that children had developed more refined phonological representation in this age range, even though there were noticeable individual variation in children's phonological capacities.

4.3.4 The relationship between vocabulary knowledge and phonological capacities. Looking into the correlations between phonological capacities and vocabulary knowledge, particularly the receptive vocabulary, might provide some insights to the relationship between these two constructs. The receptive vocabulary, rather than the expressive vocabulary, is focused here, because expressive vocabulary is greatly confounded by children's phonological production ability, and their correlation relationship is expectable. Indeed, our inspection of the correlation between phonological production of onsets and expressive vocabulary (either MCDI-$_{\text{EXPRESSIVE}}$ or REVT) revealed significant correlations (r = .43-.45, p < .05).

We examined the correlation between phonological production of onsets and receptive vocabulary (either MCDI-$_{RECEPTIVE}$ or PPVT-R). No significant correlation was found between onset production and MCDI-$_{RECEPTIVE}$. However, onset production at age 2 showed a significant correlation with PPVT-R at age 3 (r = .67, p = .001). Similarly, we found a significant correlation between onset production at age 2.5 and PPVT-R at age 3 (r = .46, p < .05). As for the relationship between discrimination and vocabulary knowledge, no correlation was found among them. The findings demonstrated a relationship between the phonological capacities, particularly phonological production, with the subsequent vocabulary development.

4.4 NWR Development

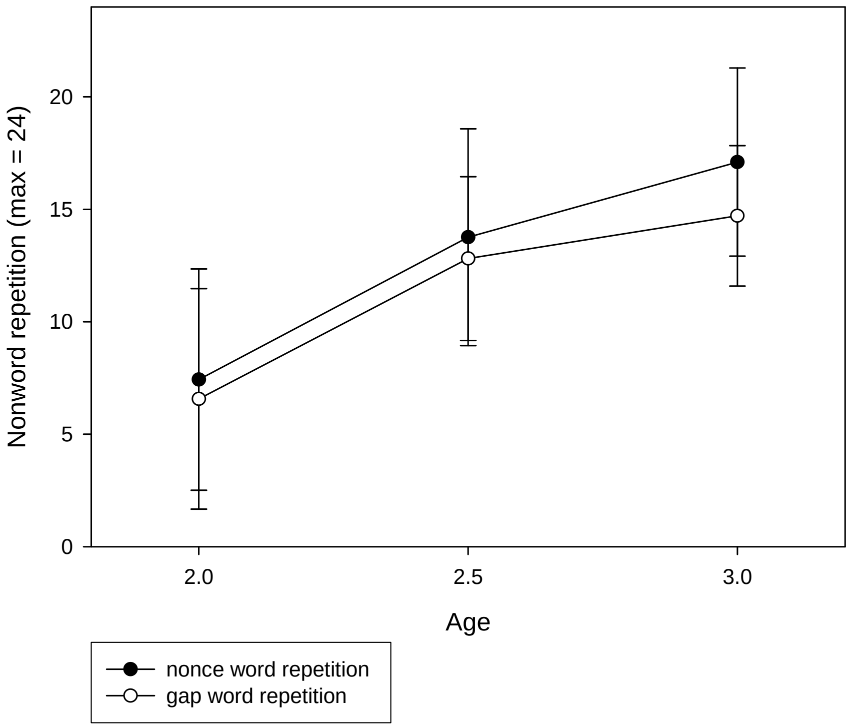
Our study is among the few studies which applied the nonword repetition task to children as young as two (Chiat & Roy, 2007; Hoff et al., 2008; Stokes & Klee, 2009). One concern of applying this task to young children is their ability to follow the instructions and the compliance to verbal output. In our study, all the children were cooperative to complete the task, except for two children. The child LLJ refused to respond to this task until age 3, so her data were missing in age 2 and age 2.5. Also, one missing data in Time 3 belonged to the child YHW, who refused to respond to any of the production tasks at Time 3.

Concerning that young children tend to have shorter attention span, we adjusted the number of items that children had to complete at each time point. At age 2, when children were age 2, they only had to complete half of the items in the one-word lists and the two-word lists in the nonword repetition task. At age 2.5, children had to perform half of the items of the entire nonword repetition task (up to the three-word lists). At age 3, children had to complete the entire nonword repetition task, as children in the other cohorts did. Children's performances were presented in Table 4.6. Since children in this cohort completed different numbers of items at different time points, the data in the following analyses uniformly consisted of children's performances of just half of the items in the one-word lists and the two-word lists in the task.

Children's NWR performances across the three time points were plotted in Figure 4.7. The line with black dots represents children's performance in the nonce word repetition. Children obtained an average score of 7.43 (SD = 4.92) at age 2, and an average score of 13.76 (SD = 4.82). When they reached age 3, they gained an average score of 17.10 (SD = 4.18). Children's gap word repetition performance is displayed with the line in hollow dots. Children attained an average score of 6.57 (SD = 4.90) at age 2, and an average score of 12.81 (SD = 3.64) at age 2.5. At age 3, children gained an average score of 14.71 (SD = 3.12).

Figure 4.7

Children's Average Performances in Nonce Word Repetition and Gap Word Repetition (Up to the Two-word List)



Correlation analyses revealed that children's performances in nonce word repetition at each time point was correlated with each other. Modest correlations were found between nonce word repetition at age 2 and age 2.5 ($r = .49$, $p < .05$), and nonce word repetition at age 2 and age 3 ($r = .57$, $p < .01$). The correlation between nonce word repetition at age 2.5 and age 3 exhibited high association ($r = .82$, $p < .001$). Regarding children's performances in gap word repetition, significant correlations were found between age 2 and age 2.5 ($r = .55$, $p = .01$), and between age 2.5 and age 3 ($r = .59$, $p < .01$). The correlation between gap word repetition at age 2 and gap word repetition at age 3 did not reach significance ($r = .38$, $p > .05$).

Children's performances in the nonce word repetition and gap word repetition were associated. Nonce word repetition performance at each time point was found to be significantly associated with the concurrent performance in gap word repetition ($r = .55\text{-}.78$, $p < .05$). Nonce word repetition at age 2 was found to be associated with gap word repetition at age 2.5 ($r = .43$, $p = .05$), but not gap word repetition at age 3 ($r = .19$, $p > .05$). Nonce word repetition at age 2.5 was not associated with gap word repetition at age 3 ($r = .33$, $p > .05$). The findings suggested that the relationship between nonce word repetition and the long term performance in gap word repetition was modest to weak. On the other hand, gap word repetition at age 2 was significantly associated with nonce word repetition at age 2.5 ($r = .58$, $p < .01$), and nonce word repetition at age 3 ($r = .67$, $p = .001$). Gap word repetition at age 2.5 was also correlated with nonce word repetition at age 3 ($r = .59$, $p < .01$). The findings implied that there was a close relationship between gap word repetition and the subsequent performances in nonce word repetition.

Table 4.6

Correlation Matrix of the Nonce Word Repetition and Gap Word Repetition Across Time

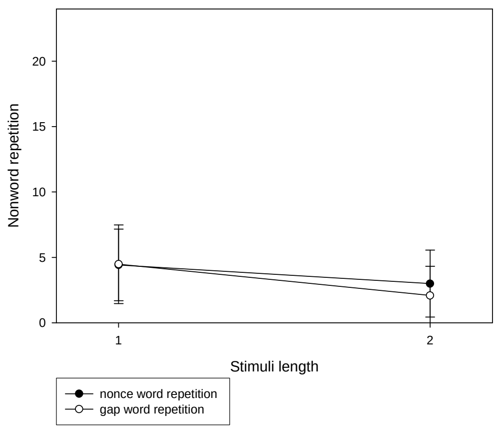
	Nonce @age 2.5	Nonce @age 3	Gap @age 2	Gap @age 2.5	Gap @age 3
Nonce $_{@age\ 2}$	.49*	.57**	.78**	.43*	.19
Nonce $_{@age\ 2.5}$		.82**	.58**	.62**	.33
Nonce $_{@age\ 3}$			.67**	.59**	.55**
Gap $_{@age\ 2}$				.55**	.38
Gap $_{@age\ 2.5}$					.59**

To examine whether children improved in NWR across time, the repeated measures ANOVA was conducted with the variables of time, lexicality and length. Significant differences were found among NWR performances across the three time points, $F(2, 38) = 77.11$, $p < .001$, $\eta^2 = .80$, stimuli with different lexicality, $F(1, 19) = 7.19$, $p = .015$, $\eta^2 = .28$, and stimuli of different lengths, $F(1, 19) = 106.14$, $p < .001$, $\eta^2 = .84$. We also observed a significant interaction effect of lexicality and length, $F(1, 19) = 21.00$, $p = .006$, $\eta^2 = .33$. The interaction effect resulted from the differences in the gaps between nonce-word repetition and gap-word repetition at different lengths. Simple main effects showed that there was an advantage for the nonce words in the repetition of longer stimuli item, i.e. when repeating the two-word lists ($t(22) = 2.75$, $p = .012$). Simple contrasts analysis further presented the presence of the lexicality effect in the two word lists at each time point (age 2.5: $t(20) = 2.74$, $p < .05$; age 3: $t(20) = 3.10$, $p < .01$), except for the repetition of the two-word lists at age 2 ($p > .05$) (Figure 4.8). However, when repeating a single nonword, children had similar performance in the nonce word repetition and gap word repetition at age 2 ($t(22) = -.12$, $p > .05$), at age 2.5 ($t(20) = -.50$, $p > .05$), and at age 3 ($t(20) = 1.75$, $p > .05$).

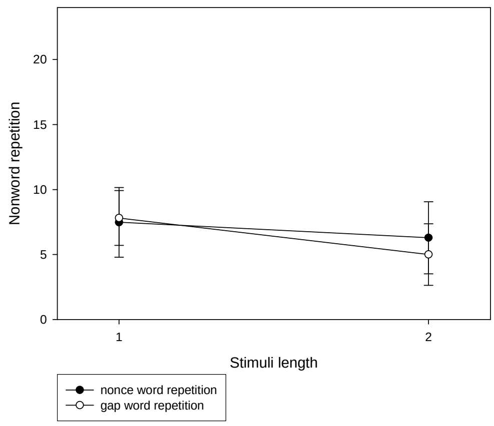
Figure 4.8

NWR Performances as the Function of Length and Lexicality in (a) Time 1, (b) Time 2 and (c) Time 3

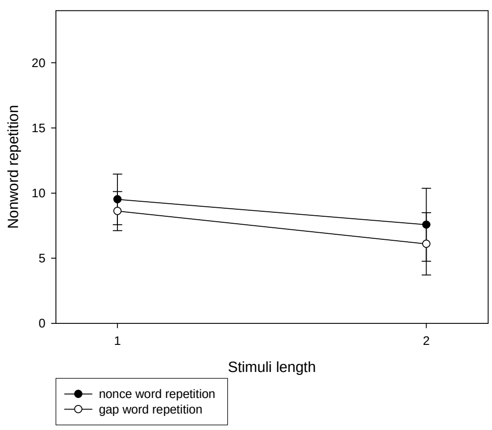
a)



b)



c)

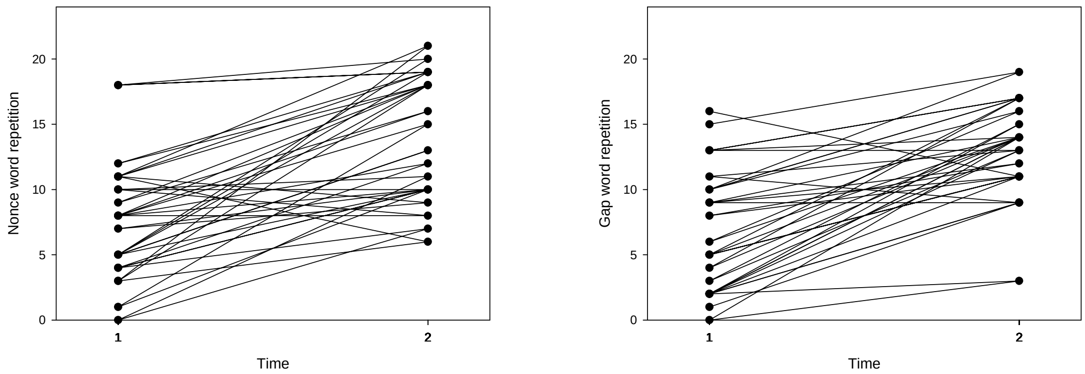


Each child's growth curves in nonce word repetition and gap word repetition were displayed in Figure 4.9(a) and Figure 4.9(b), respectively. Similar patterns have been observed in children's individual growth curves in both types of nonwords. Children in this cohort demonstrated considerable individual variation in their NWR performances. However, they were more or less parallel in the growth rates.

Figure 4.9

Children's Individual Growth Curves of (a) Nonce Word Repetition and (b) Gap Word Repetition

a) b)



4.5 NWR, Phonological Capacities and Vocabulary Knowledge

In this section, we examined the predictability of children's phonological capacities and vocabulary knowledge to their nonword repetition performance. Children's phonological capacities were represented by children's phonological production ability (as measured by onset production), and word discrimination. We did not consider children's production of rhymes in the following analyses, because children's acquisition of most rhymes were stabilized at the age of 3, and the rhyme production might not properly reflect the individual variation in their phonological capacities. With regards to vocabulary knowledge, we considered receptive vocabulary, rather than expressive vocabulary, as the main contributor based on the notion that receptive vocabulary and expressive vocabulary are two manifestations depending on one unitary lexicon system (Melka, 1997), and receptive vocabulary is less affected by children's motor planning ability.

Preliminary stepwise regression analyses revealed that children's vocabulary knowledge and productive phonology were associated with NWR, while word discrimination did not. Therefore, word discrimination was not included in the following analyses.

In the following regression analyses, children's nonce word repetition and gap word repetition at age 2.5 and age 3 were entered as the dependent variables respectively. As mentioned previously, the number of nonword items children completed differed across time. Analyses in the previous section showed that in addition to lexicality, length of the stimuli item would affect children's NWR performance. To control for the effect of length on the factors involved in NWR, the regression analyses throughout the three testing points were conducted based on children's minimum completion of NWR, that is, half of the stimuli up to the two word lists. To specify the relative contribution of productive phonology and receptive vocabulary to NWR, we carried out the hierarchical regression analyses. Children's Leiter-R score was always entered as the first step to control for the variation in nonverbal intelligence. In the first set of the analyses, the productive phonology at age 2 was entered prior to the receptive vocabulary at age 2. In the second set of the analyses, their entry sequence was reversed.

The analyses on NWR at age 2.5 demonstrated that both the nonce word repetition and the gap word repetition at age 2.5 were best accounted for by receptive vocabulary at age 2 ($\beta = .50$, $p < .05$, and $\beta = .78$, $p < .001$, respectively) (Table 4.7). Productive phonology at age 2 also played a role in predicting nonce repetition at age 2.5, but it could not predict gap word repetition at age 2.5.

As for the NWR performance at age 3, children's nonce word repetition at age 3 could be predicted by both their productive phonology at age 2 ($\beta = .46$, $p < .01$) and receptive vocabulary at age 2 ($\beta = .58$, $p = .001$) (Table 4.8). Compared with the productive phonology, the receptive vocabulary was of higher predictive power. With regards to children's gap word repetition, children's receptive vocabulary predicted their performance of repeating gap words at age 3 ($\beta = .49$, $p < .05$) (Table 4.9). It was also noticed that children's nonverbal intelligence seemed to play a role here, with the β value at .50 ($p < .05$). We found no role for productive phonology in the gap word repetition.

Table 4.7

The Hierarchical Regression Analyses on Nonce Word Repetition and Gap Word Repetition at Age 2.5 (Time 2)

Dependent variable		Independent variable	△R^2	△F	p	β	p
Nonce repetition-$_{age\ 2.5}$	1	Leiter-R	.00	.05	> .05	-.12	> .05
	2	Productive phonology-$_{age\ 2}$	.28	7.16	< .05	.41	< .05
	3	MCDI-$_{RECEPTIVE}$-$_{age\ 2}$	.22	7.57	< .05	.50	< .05
	2	MCDI-$_{RECEPTIVE}$-$_{age\ 2}$	.37	10.70	< .01	.50	< .05
	3	Productive phonology-$_{age\ 2}$	.13	4.54	< .05	.41	< .05
Gap repetition-$_{age\ 2.5}$	1	Leiter-R	.00	.02	> .05	.16	> .05
	2	Productive phonology-$_{age\ 2}$	.04	.78	> .05	-.04	> .05
	3	MCDI-$_{RECEPTIVE}$-$_{age\ 2}$	.54	21.84	< .001	.78	< .001
	2	MCDI-$_{RECEPTIVE}$-$_{age\ 2}$	.58	24.81	< .001	.78	< .001
	3	Productive phonology-$_{age\ 2}$	.00	.04	> .05	-.04	> .05

Table 4.8

The Hierarchical Regression Analyses on Nonce Word Repetition at Age 3 (Time 3)

Step	Independent variable	$\triangle R^2$	$\triangle F$	p	β	p
1	Leiter-R	.01	.20	> .05	.03	> .05
2	Productive phonology-$_{\text{age 2}}$	.41	12.66	< .005	.46	< .01
3	MCDI-$_{\text{RECEPTIVE}}$-$_{\text{age 2}}$	.29	16.60	= .001	.58	= .001
2	MCDI-$_{\text{RECEPTIVE}}$-$_{\text{age 2}}$	.53	20.66	< .001	.58	= .001
3	Productive phonology-$_{\text{age 2}}$	.17	9.65	< .01	.46	< .01
1	Leiter-R	.01	.20	> .05	.11	>.05
2	Productive phonology-$_{\text{age 2.5}}$	.14	3.04	> .05	.35	>.05
3	MCDI-$_{\text{RECEPTIVE}}$-$_{\text{age 2.5}}$	.23	6.16	< .05	.48	<.05
2	MCDI-$_{\text{RECEPTIVE}}$-$_{\text{age 2.5}}$	.24	5.89	< .05	.48	<.05
3	Productive phonology-$_{\text{age 2.5}}$	.12	3.40	>.05	.35	>.05
1	Leiter-R	.01	.20	> .05	.06	> .05
2	Productive phonology-$_{\text{age 2.5}}$	.14	3.04	> .05	.31	>.05
3	PPVT-$_{\text{age 2.5}}$	.04	0.73	> .05	.20	>.05
2	PPVT-$_{\text{age 2.5}}$	.09	1.84	> .05	.20	>.05
3	Productive phonology-$_{\text{age 2.5}}$	.09	1.80	>.05	.31	>.05

Table 4.9

The Hierarchical Regression Analyses on Gap Word Repetition at Age 3 (Time 3)

Step	Independent variable	△R^2	△F	p	β	p
1	Leiter-R	.22	5.48	< .05	.50	< .05
2	Productive phonology-$_{\text{age 2}}$	.07	1.82	> .05	.10	>.05
3	MCDI-$_{\text{RECEPTIVE}}$-$_{\text{age 2}}$	.21	7.08	< .05	.49	< .05
2	MCDI-$_{\text{RECEPTIVE}}$-$_{\text{age 2}}$	.27	9.64	< .01	.49	< .05
3	Productive phonology-$_{\text{age 2}}$	.01	0.27	> .05	.10	>.05
1	Leiter-R	.22	5.48	< .05	.48	< .05
2	Productive phonology-$_{\text{age 2.5}}$	.04	0.85	> .05	.17	>.05
3	MCDI-$_{\text{RECEPTIVE}}$-$_{\text{age 2.5}}$	.10	2.69	>.05	.32	>.05
2	MCDI-$_{\text{RECEPTIVE}}$-$_{\text{age 2.5}}$	.11	2.89	>.05	.32	>.05
3	Productive phonology-$_{\text{age 2.5}}$	.03	0.76	> .05	.17	>.05
1	Leiter-R	.22	5.48	< .05	.45	< .05
2	Productive phonology-$_{\text{age 2.5}}$	.04	0.85	> .05	.16	>.05
3	PPVT-$_{\text{age 2.5}}$	.01	0.17	>.05	.09	>.05
2	PPVT-$_{\text{age 2.5}}$	.02	0.50	>.05	.09	>.05
3	Productive phonology-$_{\text{age 2.5}}$	.02	0.50	> .05	.16	>.05

Parallel regression analyses were conducted to examine the contribution of Leiter-R, receptive vocabulary and productive phonology at age 2.5 to NWR performance at age 3. It was found that nonce word repetition could be accounted for by receptive vocabulary at age 2.5 ($\beta = .48$, $p < .05$) (Table 4.8), while gap word repetition was mainly accounted for by Leiter-R ($\beta = .48$, $p < .05$) (Table 4.9). Since children at age 2.5 received a standardize test on receptive vocabulary, i.e. PPVT-R, we replaced the receptive vocabulary score as measured by MCDI-T with PPVT-R in the regression analyses. It turned out that neither PPVT-R nor productive phonology accounted for significant variance in nonce word repetition (Table 4.8) or gap word repetition (Table 4.9). Only Leiter-R was found to make significant contribution to gap word repetition.

Overall, the results of the regression analyses revealed an early role of receptive vocabulary to NWR. At age 2.5 and age 3, the repetition of nonwords, both nonce words and gap words, could be predicted by the receptive vocabulary size at age 2. Moreover, the receptive vocabulary size at age 2.5 could predict nonce word repetition at age 3, though not gap word repetition at age 3. In other words, children who had larger receptive vocabulary tended to have better NWR performances later. However, the effect of receptive vocabulary seemed to be restricted to the use of $\text{MCDI-}_{\text{RECEPTIVE}}$ measure. The effect of receptive vocabulary was not observed when the measure PPVT-R was considered.

While we found an effect of receptive vocabulary on NWR, we would expect to see a lexical advantage in the repetition of nonce words, because these nonwords consist of lexical syllables in Mandarin. However, statistical analyses in section 4.4 showed that the lexical advantage of nonce words was not prominent in this age cohort until the children reached age 2.5 or older. Therefore, it is likely that in this age range, receptive vocabulary contributes to NWR by providing a bank of phonological vocabulary forms that are partially or sufficiently encoded, and could be easily accessed to support the decoding and encoding of novel words. In other words, the receptive vocabulary effect on NWR at this age might be more related to the breath of vocabulary knowledge.

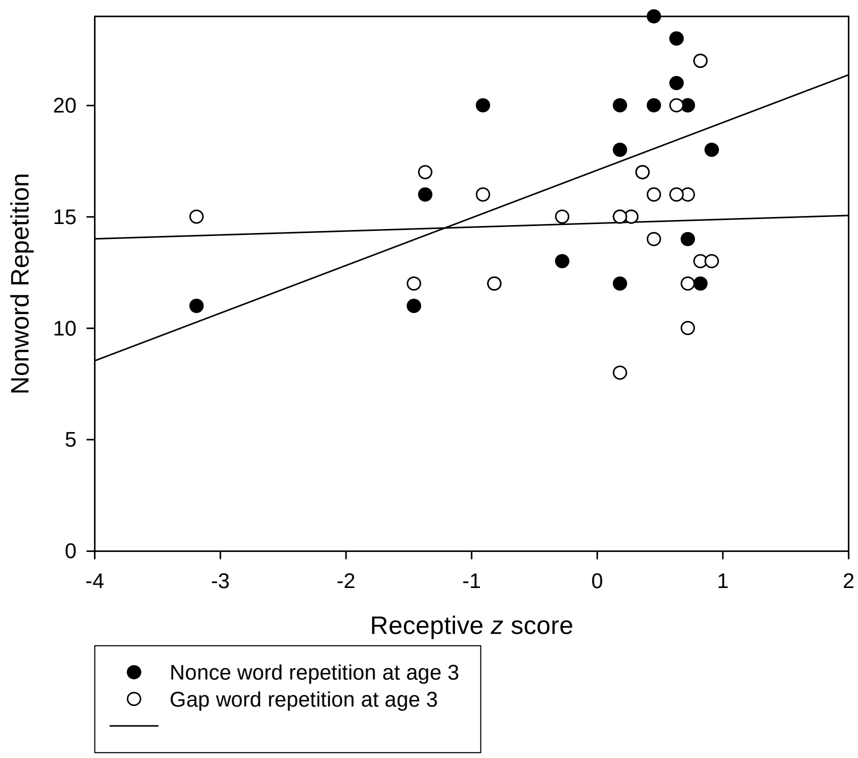
The phonological analysis account asserts that the effect of vocabulary to NWR was through the mediation of productive phonology. That is, the increase in receptive vocabulary facilitated the improvement in children's productive phonology, which further enhanced their performance in NWR. However, this hypothesis was not supported statistically. When we controlled for the effect of the productive phonology to its concurrent NWR, there was still a significant and robust effect of receptive vocabulary. Also, the discussion on the relationship between vocabulary knowledge and phonological capacities (in Section 4.3.4) suggested that in this cohort phonological capacities appeared to exert influences on vocabulary development, rather than the other way around.

The lexicality effect emerged at age 2.5 and age 3, however, only in the lengthy stimuli item. The presence of the lexicality effect implies that the emergence of sensitivity to the lexical status of the novel sound stimuli. This awareness is supposed to be nurtured by the increase in vocabulary size. Therefore, we hypothesized that the lexicality effect might be more prominent in children with larger vocabulary sizes. Therefore, we divided children into two groups (good vs. poor) based on their *z* scores in receptive vocabulary (measured by MCDI-$_{\text{RECEPTIVE}}$) at each testing time. If our hypothesis is true, then there would be an interaction effect of lexicality and vocabulary group. We carried out the repeated measures ANOVAs on children's performances at age 2, age 2.5, and age 3, and found a near-significant interaction effect of lexicality and vocabulary group at age 3, $F(1, 19) = 4.33$, $p = .051$. Figure 4.10 showed that children with larger receptive vocabulary size at age 3 tended to have better performance in repeating nonce words, even though they were not better at repeating the gap words, when compared to children with smaller receptive vocabulary size.

Children's productive phonology was found to play a role in accounting for the subsequent performance in nonce word repetition, though it had less predictive power than receptive vocabulary. Children's productive phonology at age 2 could predict nonce word repetition at age 2.5 and age 3. Nevertheless, productive phonology at age 2.5 did not show predictability to NWR at age 3. Productive phonology at age 2, but not productive phonology at age 2.5, was particularly involved in NWR, probably because phonological production at age 2 revealed greater individual differences among children, thus tended to have greater power to predict variation in the subsequent nonce word repetition.

Figure 4.10

A Scatter Plot on the Relationships of Receptive Vocabulary Z Scores at Age 3 with Nonce Word Repetition and Gap Word Repetition at Age 3



Productive phonology was generally not associated with gap word repetition at any ages. On the contrary, the repetition of gap words appeared to be related to children's nonverbal intelligence, which was measured with the Leiter-R.

Our working model hypothesized that for children with little vocabulary, their NWR performance could be mainly constrained by their productive phonology. This phase was expected to be observed in this cohort in this study since children were still at the early stage of language development. To further explore this hypothesis, we examined the relationship between productive phonology and NWR performance among those children with expressive vocabulary size below 200 words (Marchman & Bates, 1994; Sosa & Stoel-Gammon, 2006). Among the 24 children in this cohort, only 5 children had a vocabulary size smaller than 200 words. The correlation analyses showed that higher correlation was found in these children's productive phonology at age 2 and their subsequent NWR performances (r = .61 with nonce word repetition at age 2.5; r = .78-.79 with nonword repetition at age 3), compared with the correlations between their receptive vocabulary at age 2 and the subsequent NWR performance. The correlations did not reach significance due to the small sample size. However, the data revealed the tendency that NWR performance was more correlated with productive phonology among children with little vocabulary size.

To summarize the discussions above, we have found that during this age range, children's NWR was predicated by their receptive vocabulary knowledge. Receptive vocabulary knowledge supported the repetition of nonwords by providing a bank of fully- or partially-specified phonological forms that children could access when encoding a novel sound form, regardless of its degree of wordlikeness. A lexical advantage for the more wordlike nonwords (i.e. nonce words) did not emerge until children grew older. The lexical advantage may emerge with the expansion of vocabulary size. With the increase in receptive vocabulary size and also the refinement of the phonological representations of the vocabulary items, children would gradually develop a sensitivity for the phonological similarity between a newly-encountered sound form and the lexical forms that they have acquired. This could be supported by the co-occurrence of children's dramatic increase in receptive vocabulary size from age 2 to age 2.5 and the first emergence of the lexicality effect in children's NWR at age 2.5. Also, another supporting evidence came from the finding that the lexicality effect was more prominent in age 3 children with larger receptive vocabulary size.

We found a role of productive phonology, however, only to the repetition of nonce words, and with less predictive power than receptive vocabulary. On the other hand, the repetition of gap words was related to nonverbal intelligence. However, the post hoc analyses on the NWR performance of children with little vocabulary showed that these children's NWR might be more associated with their productive phonology. A larger sample of children with the same characteristic should be recruited to further verify this hypothesis.

The findings in general provided some implications to the processes that children tackle newly-encountered sound forms. When young children first run into a novel sound form, they may first rely on the phonological forms in their vocabulary bank to decode and encode it (thus, the vocabulary breadth effect). When encoding nonce words, it is easier for children to find matched or similar phonological forms that could help them construct the representation of the sound form, or at least part of its representation, and maintain the representation in the phonological storage. Since this process could be performed more swiftly or efficiently, children could further use phonological knowledge to reconstruct incomplete traces of the constructed representation (thus, the productive phonology effect). However, when encoding gap words, it is more difficult for children to find similar sound forms that could help them construct the representation of the nonword. At the same time, they also have to retain the sound in the phonological storage. The storage and retrieval process in encoding gap words would greatly tax children's ability to allocate their attentions between these two different demands (thus, the nonverbal intelligence effect). Therefore, the difference between the nonce word repetition and gap word repetition in terms of the contributing factors might result from the cognitive load of encoding in each of the task.

It is also of interest whether NWR at age 2 could predict children's subsequent vocabulary development. Correlation analyses showed that nonce word repetition at age 2 was correlated with REVT at age 3 (r = .60, p = .005). Gap word repetition at age 2 was correlated with MCDI-$_{\text{EXPRESSIVE}}$ at age 2.5 (r = .51, p < .05), and REVT at age 3 (r = .47, p < .05). It was also modestly correlated with receptive vocabulary at age 2.5, either measured by MCDI-$_{\text{RECEPTIVE}}$ (r = .37, p = .098) or measured by PPVT-R (r = .38, p = .088); however, neither achieved conventional significance level.

Regression analyses revealed that nonce word repetition at age 2 could account for 20.2% of the variance in REVT at age 3 (β = .52, p < .05), when the nonverbal intelligence and age were controlled. After controlling for the nonverbal intelligence variable, gap word repetition at age 2 could account for 20.8% of variance in MCDI-$_{\text{EXPRESSIVE}}$ at age 2.5 (β = .57, p < .05). The findings suggested that children's NWR performances were particularly associated with their subsequent expressive vocabulary development.

Chapter 5 Vocabulary, Phonology, and NWR of the Age 3 Cohort

This chapter presented the vocabulary development, phonological development, and the nonword repetition development of children in the age 3 cohort. The relative contributions of phonological capacities and vocabulary development to NWR performances were also investigated. Preliminary analyses showed that there was no significant gender difference among the tasks; therefore, gender was not included in the following analyses, unless it was particularly specified. The performance data and the correlation relationship of the measures in discussion would be presented in the main text. However, a more detailed descriptive statistics of all the measures and an overall correlation matrix were provided in Appendix E(2) and F(2), respectively.

5.1 Participants

Children in this cohort were recruited at age 3 (age range = 35.47-38.93 months), and tested every six months. Ten of the children in this cohort were recruited by posts on parenting websites, and the other fourteen children were from three kindergartens in Taipei City and New Taipei City, Taiwan. Twenty-four children were tested at age 3. At Time 2, one child did not return because his parents were unable to cooperate with the testing schedule. At Time 3, two girl children dropped out because they left the kindergartens, and another boy moved overseas. Participant information was summarized in Table 5.1.

Table 5.1

Information of the Child Participants in the Age 3 Cohort

	Time 1	Time 2	Time 3
N	24 (8 ♂)	23 (7 ♂)	20 (6 ♂)
Average age (month)	36.80	42.77	48.70
Age range	35.47-38.93	41.47-44.63	47.37-50.73

5.2 Vocabulary Development

Children in this age cohort were old enough to receive standardized tests on vocabulary. Their expressive vocabulary was tested with the expressive vocabulary scale in REVT. Though REVT also had the receptive vocabulary scale, this study assessed children's receptive vocabulary with the PPVT-R, which is more widely-adopted in Mandarin studies on children language development and thus allows comparisons across studies. Children's growth in each task would be examined respectively in the following sections.

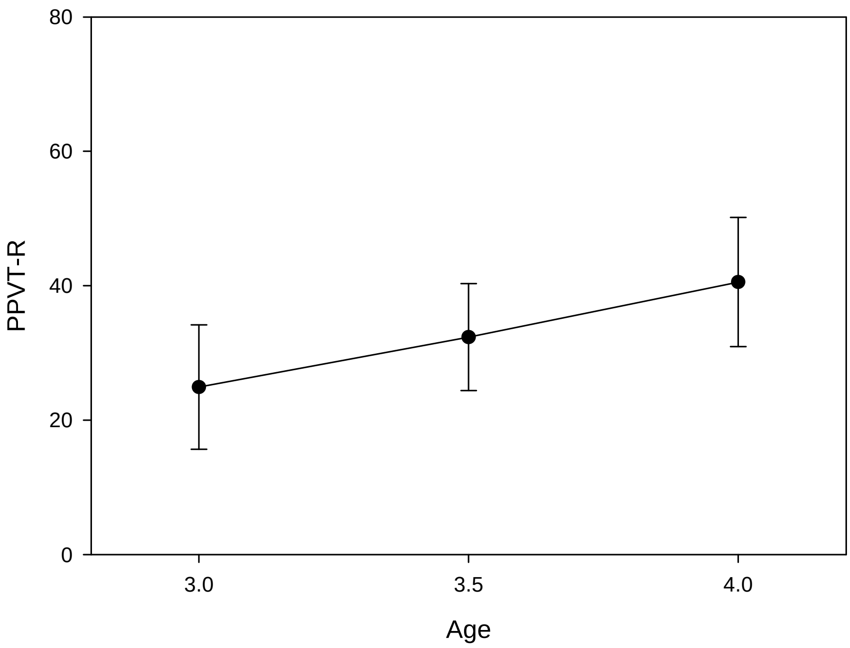
5.2.1 Receptive vocabulary size as measured by PPVT-R. Children at age 3 obtained an average score of 24.92 (SD = 9.25, range 9-49) on PPVT-R. At age 3.5, there was a prominent increase to the average score of 32.35 (SD = 7.96), ranging between 16 and 51. At age 4, children reached an average score of 40.55 (SD = 9.60, range = 19-59). Children's average PPVT-R scores of the three time points were depicted in Figure 5.1. Correlation analyses showed that PPVT-R scores of age 3, age 3.5 and age 4 were correlated with each other, with r ranging from .50 to .64 ($p < .05$).

One-way ANOVA on the effect of time revealed a significant effect, $F(2, 38) = 33.69$, $p < .001$, $\eta^2 = .64$. Paired t tests showed that there were significant increases in children's PPVT-R score from age 3 to age 3.5 ($t(22) = 4.31$, $p < .001$), and from age

3.5 to age 4 ($t(19) = 4.74$, $p < .001$). A trend analysis showed that children's growth in PPVT-R was in a linear trend. The findings suggested that children from age 3 to age 4 demonstrated steady growth in receptive vocabulary.

Figure 5.1

Children's Average Performances in PPVT-R Across Time

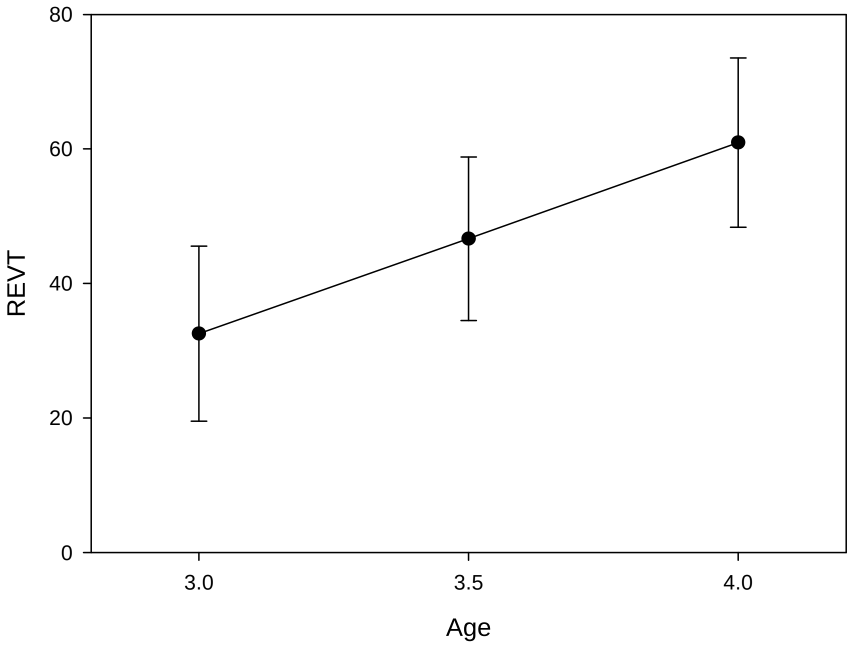


5.2.2 Expressive vocabulary size as measured by REVT. At age 3, children received an average score of 32.54 (SD = 13.02), ranging between 15 and 54. At age 3.5, an average score of 46.65 was obtained (SD = 12.16, range = 26-72). At age 4, when children reached age 4, they attained an average score of 60.95 (SD = 12.59), with the range from 40 to 86. Children's average REVT scores at the three time points were graphed in Figure 5.2. REVT scores of age 3, age 3.5 and age 4 were correlated with each other. Children's performance of REVT at each of the two adjacent time points reached a correlation as high as .8 ($p < .05$). REVT scores at age 3 and REVT scores at age 4 also revealed a correlation coefficient of .63 ($p < .05$).

One-way ANOVA on the effects of time revealed that the effect was significant, $F(2, 38) = 104.02$, $p < .001$, $\eta^2 = .85$. Paired *t* tests showed that children's REVT scores increased from age 3 to age 3.5 ($t(22) = 9.04$, $p < .001$), and from age 3.5 to age 4 ($t(19) = 8.40$, $p < .001$). The trend analysis exhibited a linear trend in children's REVT scores over time, demonstrating that children had steady growth in expressive vocabulary.

Figure 5.2

Children's Average Performances in REVT Across Time



5.2.3 The correlation between receptive vocabulary and expressive vocabulary. The correlation relationship between the receptive vocabulary (as measured by PPVT-R) and the expressive vocabulary (as measured by REVT) of the three time points were demonstrated in Table 5.2. Each of the concurrent PPVT-R and REVT scores showed correlations, though at age 3.5 the correlation did not achieve the conventional level of significance ($p = .058$).

On inspecting the long-term relationship, we found that PPVT-R score at age 3 were highly correlated with REVT at age 3.5 ($r = .76$, $p < .01$) and modestly correlated with REVT at age 4 ($r = .50$, $p < .01$). REVT score at age 3 was correlated only with PPVT-R score at age 4 ($r = .63$, $p < .01$), but not PPVT-R score at age 3.5 ($p > .05$).

Table 5.2

Correlation Matrix of PPVT-R and REVT Across Time

	PPVT @age 3.5	PPVT @age 4	REVT @age 3	REVT @age 3.5	REVT @age 4
PPVT @age 3	.54**	.50*	.75**	.76**	.50*
PPVT @age 3.5		.64**	.29	.40	.26
PPVT @age 4			.63**	.65**	.57**
REVT @age 3				.80**	.63**
REVT @age 3.5					.83**

5.2.4 Summary of vocabulary development from age 3 to age 4. Children in this age showed steady improvements in their vocabulary sizes. With respect to the development in receptive vocabulary, children in this cohort gained an average 8 points every half a year. As for the development in expressive vocabulary, there was an average of 14-point increase every half a year.

We also observed interactions between receptive vocabulary and expressive vocabulary in development. Cross-time correlations were observed not only between earlier receptive vocabulary knowledge and latter expressive vocabulary, but also between earlier expressive vocabulary and latter receptive vocabulary. This was different from what we observed in the age 2 cohort. In the age 2 cohort, associations were found between receptive vocabulary at an earlier time and latter expressive vocabulary development, but the predictability of expressive vocabulary to latter receptive vocabulary was not observed. However, according to what we found here, children in the age 3 cohort might demonstrate a different interactive patterns between receptive vocabulary and expressive vocabulary in development.

5.3 Phonological Development

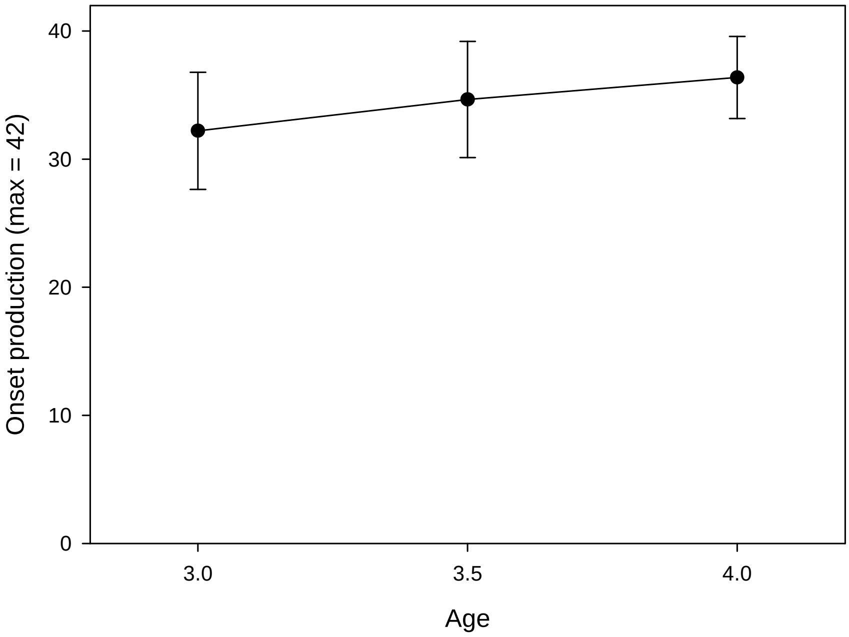
5.3.1 Productive phonology.

5.3.1.1 Onset production. Children at age 3 (Time 1) obtained an average score of 32.21 (SD = 4.57) in onset production, and the scores ranged between 19.60 and 37.06. There were small but steady increases at age 3.5 (mean = 34.66, SD = 4.53, range = 21.70-39.42) and age 4 (mean = 36.38, SD = 3.20, range = 25.16-40.11). Correlation analyses showed that onset production at the three time points were correlated with each other, with r values ranging between .69-.91, $p < .01$. Children's average performance at each time point was demonstrated in Figure 5.3.

A one-way ANOVA was conducted to examine children's growth in onset production across time. There was a significant main effect of time, $F(2, 38) = 20.78$, $p < .001$, $\eta^2 = .52$. Planned comparisons were conducted to compare onset productions at each of the two adjacent time points, and found that children's onset production performance showed significant improvement from age 3 to age 3.5 ($t(22) = 6.63$, $p < .001$), and from age 3.5 to age 4 ($t(19) = 2.16$, $p = .044$). The trend analysis showed that the growth was linear.

Figure 5.3

Children's Average Performances in Onset Production Across Time



We further examined the onsets that have become stabilized at the particular age, following the criterion of phoneme stabilization proposed in Zhu and Dodd (2000). At Time 1, when children were at age 3, most children gained mastery of oral and nasal stops, alveo-palatal affricates and fricative, alveolar liquid, and velar fricative. More than 25% of the children were still not mature with the production of *f* and alveolar affricates, and most children still had difficulty with the retroflexes. The findings were similar to what we observed in the age two cohort when the children reached age 3. By age 4, children mastered most of the onset phonemes in Mandarin except for the retroflexes. The results were summarized in Table 5.3.

Error patterns found with children in this age range were not as diverse as those of children in the age 2 cohort, thus revealing their growing maturity in phonological output ability. The common error patterns observed in this age range included stopping (e.g. *cao3-mei2* "strawberry" → *tao3-mei2*), gliding (e.g. *re4-gou3* "hotdog" →

le4-gou3), fronting (e.g. *shou3-biao3* "watch" → *sou3-biao3*).

Table 5.3

Children's Development in Onset Production

Pass rate[a]	Age 3	Age 3.5	Age 4
> 75% (stabilized)	*b, p, m, d, t, n, l, g, k, h, j, q, x, s*	*b, p, m, f, d, t, n, l, g, k, h, j, q, x, c, s*	*b, p, m, f, d, t, n, l, g, k, h, j, q, x, z, c, s*
50-75%	*f, z, c*	*z,*	
< 50%	*zh, ch, sh, r*	*zh, ch, sh, r*	*zh, ch, sh, r*

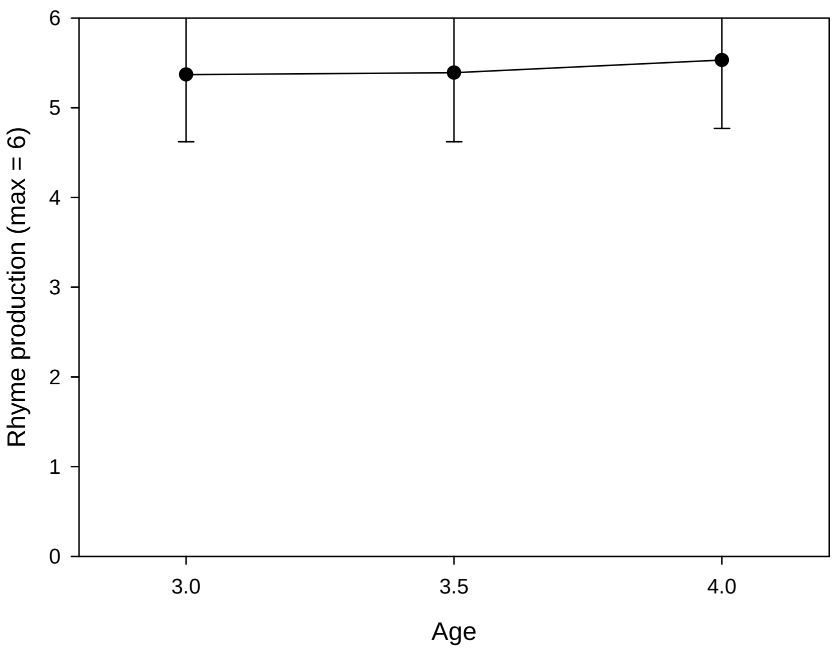
Note. [a] % of children passed the 66.7% accuracy rate

5.3.1.2 Rhyme production. In this measure, we assessed children's productions of six rhyme structures, including V, VG, GV, VN, GVN, and GVG. At age 3, children obtained an average score of 5.37 (SD = 0.75), which ranged between 3.66 and 6.00 (maximum score = 6). The mean and the range suggested that children were about to reach ceiling in their production of rhymes. There were just slight increases in the average rhyme production scores at age 3.5 (mean = 5.39, SD = 0.77, range = 3.50-6.00) and at age 4 (mean = 5.53, SD = 0.76, range = 3.8-6.00). Thus, children presented mature and stable performance in rhyme production. Correlation analyses showed that rhyme production at the three time points were highly correlated with each other, with r value ranging between .90-.95, $p < .001$.

A one-way ANOVA was conducted to compare children's rhyme production across the three time points. There was a significant main effect of time, $F(2, 38) = 3.27$, $p = .049$, $\eta^2 = .15$. Paired t tests were conducted to compare rhyme productions at each of the two time points. The only significance was found when comparing rhyme production at age 3 and rhyme production at age 4, $t(19) = 2.09$, $p = .05$. The trend analysis showed that the growth was linear ($F(1, 19) = 4.38$, $p = .05$).

Figure 5.4

Children's Average Performances in Rhyme Production Across Time



We further examined the rhyme structures that have become stabilized at the particular age. Again, Zhu and Dodd's (2000) criterion of phoneme stabilization was applied. It was found that children at age 3 gained mastery of almost all of the rhyme types in Mandarin, except for those containing a final nasal coda. The findings were similar to the rhyme production performances of the age 2 cohort children at age 3. The GVN type became stabilized by the age of 3.5, while the VN type became stabilized at age 4.

5.3.1.3 The relationship between the production of onsets and rhymes. Correlation analysis was conducted to examine the associations between the productions of onset and rhyme structure across time (Table 5.4).

Significant correlations were found between concurrent onset and rhyme production abilities. Performance at age 3 revealed the highest correlation, $r = .72$, $p < .01$, while the correlations at age 3.5 and age 4 were modest (both $r = .48$, $p < .05$). Onset production at age 3 was significantly correlated with rhyme production at age 3.5 ($r = .62$, $p < .01$) and age 4 ($r = .66$, $p < .01$), while similar patterns were also found between rhyme production at age 3 and onset production at age 3.5 ($r = .61$, $p < .01$) and age 4 ($r = .59$, $p < .01$). The findings suggested that onset production was associated with rhyme production in this age cohort, though the correlation was modest.

Table 5.4

Correlation Matrix of the Production of Onsets and Rhymes Across Time

	Onset @age 3.5	Onset @age 4	Rhyme @age 3	Rhyme @age 3.5	Rhyme @age 4
Onset @age 3	.91**	.69**	.72**	.62**	.66**
Onset @age 3.5		.72**	.61**	.48*	.56*
Onset @age 4			.59**	.58**	.48*
Rhyme @age 3				.94**	.90**
Rhyme @age 3.5					.95**

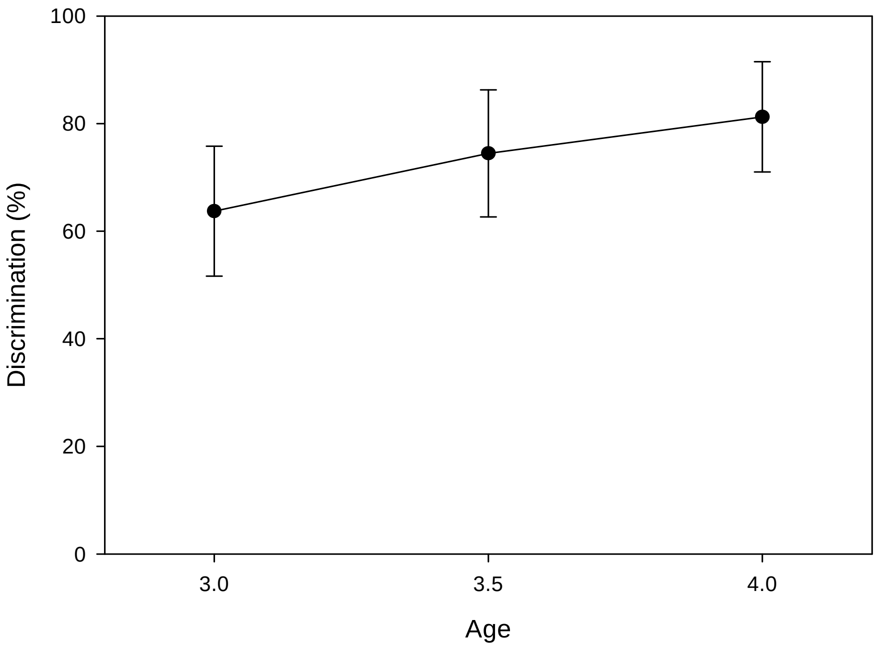
5.3.2 Word discrimination. This measure assessed children's word discrimination ability at the phoneme level. Unlike children in the age two cohort, children in this cohort completed the same number of items across three testing sessions. Therefore, the subsequent statistical analyses were conducted with the raw scores.

At age 3, children gained an average of 15.29 points (SD = 2.89, range = 12-22), attaining an accuracy rate of around 64%. At age 3.5, the average score increased to

17.87 (SD = 2.83, range = 12-24), attaining an accuracy rate of 74.46%. Half a year later, children attained an average of 19.50 points (SD = 2.46, range = 16-24), an accuracy rate of 81.25%. Correlations were found between discrimination at age 3 and age 3.5 ($r = .56, p < .01$), and between age 3 and age 4 ($r = .50, p < .05$). The correlation between the discrimination at age 3.5 and discrimination at age 4 approached but fell short of significance ($r = .39$, $p = .092$). This could be due to the fact that most of the children had good performances in the discrimination task at age 3.5 and age 4, and the little variation in children's scores rendered the correlation insignificant.

Figure 5.5

Children's Average Performances in Word Discrimination Across Time

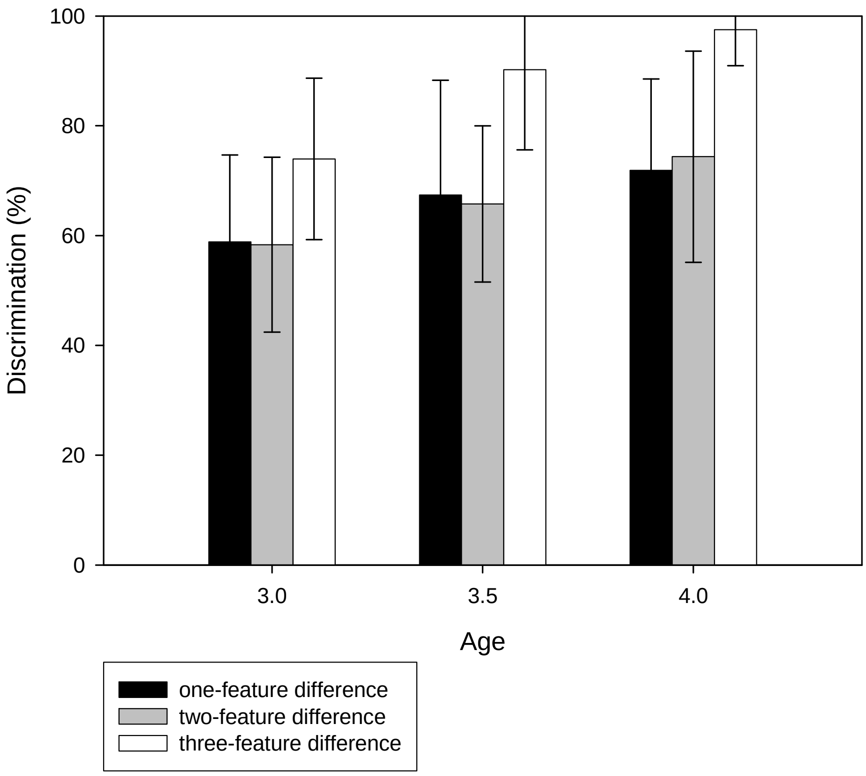


5.3.2.1 Sensitivity for different degrees of phonetic contrast. Figure 5.6 presented children's performances in discriminating minimal pairs of different degrees of phonetic contrast. One-sample *t* tests were first conducted to examine whether children's performance in discriminating the contrast pairs at each time point was above chance level. The results showed that children's performances were above the chance level across all contrast pairs and all testing sessions (all $p < .05$, one-tailed).

Paired t tests further showed that there was significant growth in their discrimination ability from age 3 to age 3.5 ($t(22) = 5.05$, $p < .001$), and from age 3.5 to age 4 ($t(19) = 2.3$, $p < .05$). The growth was linear and steady, as reflected by the trend analysis, $F(1, 19) = 55.36$, $p < .001$.

Figure 5.6

Children's Average Performances in Discriminating Different Degrees of Phonetic Contrasts



A two-way ANOVA was conducted with the variables of time and degree of phonetic contrast on children's discrimination performance. The results revealed significant main effects of time ($F(2, 38) = 28.72$, $p < .001$, $\eta^2 = .60$), and contrast degree ($F(2, 38) = 43.60$, $p < .001$, $\eta^2 = .70$). There was no significant interaction effect ($F(4, 76) = 1.19$, $p > .05$).

The significant main effect of time suggested that children's sensitivity for phonetic contrasts at the phoneme level increased over time. Post hoc analyses showed that there was significant growth in their discrimination ability from age 3 to age 3.5 ($t(22) = 5.05$, $p < .001$), and from age 3.5 to age 4 ($t(19) = 2.3$, $p < .05$).

Also, statistical analyses were conducted to examine the main effect of phonetic contrast. The results revealed that the significance of the main effect resulted from the significance found between 3-*f* difference contrasts and 2-*f* difference contrasts ($t(19) = 8.03$, $p < .001$) and 3-*f* difference contrasts and 1-*f* difference contrasts ($t(19) = 7.81$, $p < .001$). Children showed similar sensitivity for contrasts with 1-*f* difference and contrasts with 2-*f* difference ($t(19) = .08$, $p > .05$).

The findings in general suggested that children's sensitivity for contrasts at the phoneme level increased over time. Though children showed above chance sensitivity for all the contrast types, they were more capable of discriminating minimal pairs with more prominent contrasts (i.e. those with 3-*f* difference) than those with 1-*f* difference or those with 2-*f* difference.

5.3.3 Summary of phonological development from age 3 to age 4. Findings in this section showed that children in this age range mastered most of the onsets and rhyme types. With regards to onset production, children mastered all onsets except for retroflexes by age 4. Particularly during this age range, the production of *f*, *z* and *c* became stabilized. At the age of 4, children still had difficulty in producing retroflexes, the accuracy rates of which remained low throughout this age range. As for rhymes, children mastered most of them. However, some children still had difficulty with those containing final nasal codas. Compared with children in the age 2 cohort, children in this age range manifested relatively more mature mastery with the phonemes in Mandarin Chinese.

In terms of children's discrimination at phoneme level, children's performance was above chance at age 3, suggesting a sensitivity for contrasts at the phonemic level, either when the contrasts were small or prominent. However, children were still better at discriminating more prominent contrasts, while their performance in discriminating the less-prominent contrasts remained poor. At age 3, only about a quarter or less of the children could discriminate 1-*f* difference contrasts and 2-*f* difference contrasts. It was not until when children reached age 4 that more than half of the children could discriminate these less-prominent contrasts.

A modest correlation was found between children's phonological production ability (as reflected by onset production) and discrimination at age 3 (r = .43, p < .05). However, their association diminished at age 3.5 and age 4.

5.3.4 The relationship between vocabulary knowledge and phonological capacities. The correlation analyses were conducted to examine phonological capacities, including onset production and word discrimination, and the receptive vocabulary.

With regards to the correlation between onset production and receptive vocabulary, PPVT-R at age 4 showed marginal correlation with onset production at age 3 (r = .42, p = .069) and onset production at age 3.5 (r = .42, p = .066). We also found a marginal correlation between onset production at age 4 and PPVT-R at age 3.5 (r = .40, p = .084).

The analysis between word discrimination and receptive vocabulary revealed a significant correlation between PPVT-R at age 3 and word discrimination at age 3.5 (r = .51, p < .05). Marginal correlations were found between PPVT-R at age 4 and word discrimination at age 3 (r = .41, p = .072) and at age 3.5 (r = .40, p = .08).

As shown in the analyses above, phonological capacities and receptive vocabulary demonstrated modest-to-weak correlations. There were tendencies for earlier onset production ability or discrimination ability to correlate with the subsequent receptive vocabulary. However, we also observed correlations between earlier receptive vocabulary and the subsequent phonological capacities, either phonological perception or phonological production. The findings appeared to suggest that during this age range, there might be an interaction between these two constructs.

5.4 NWR Development

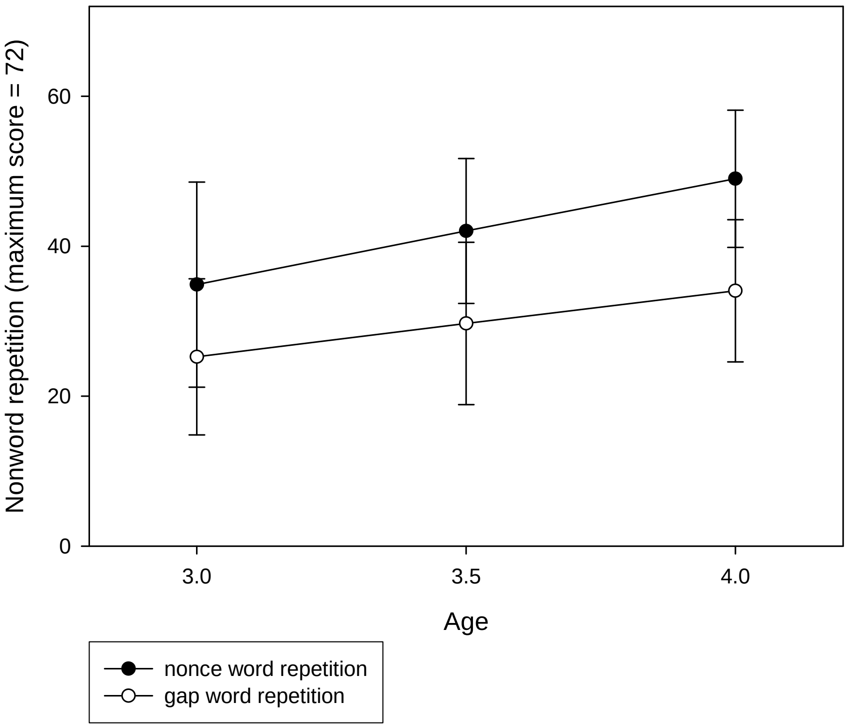
Different from children in the age 2 cohort, children in this cohort performed the complete set of the nonword repetition. In total there were 36 disyllabic words in either the nonce word repetition or the gap word repetition, respectively. Each correctly repeated syllable would be credited with one point; thus the maximum score for the nonce word repetition or the gap word repetition was 72. Children's repetition performances in each type of the nonwords were graphed in Figure 5.7.

With regards to children's performance in nonce word repetition, they obtained an average score of 34.88 (SD = 13.68) at age 3, and an average score of 42.04 (SD = 9.67) at age 3.5. When they reached age 4, they gained an average score of 49 (SD = 9.15). In the gap word repetition, children attained an average score of 25.25 (SD = 10.41) at age 3, and an average score of 29.70 (SD = 10.84) at age 3.5. At age 4, children gained an average score of 34.05 (SD = 9.48).

Correlation analyses were first conducted to examine (1) the correlations of repetition performances within each type of nonwords across time, and (2) the correlations between the nonce word repetition and the gap word repetition. The analyses on nonce word repetition showed that their performances at each time point was highly correlated with each other, with the r value ranged between .83-.86 ($p < .001$). Similar results had been observed in children's performances in gap word repetition, with the *r* values ranged between .77 and .90 ($p < .001$).

Figure 5.7

Children's Average Performances in Nonce Word Repetition and Gap Word Repetition



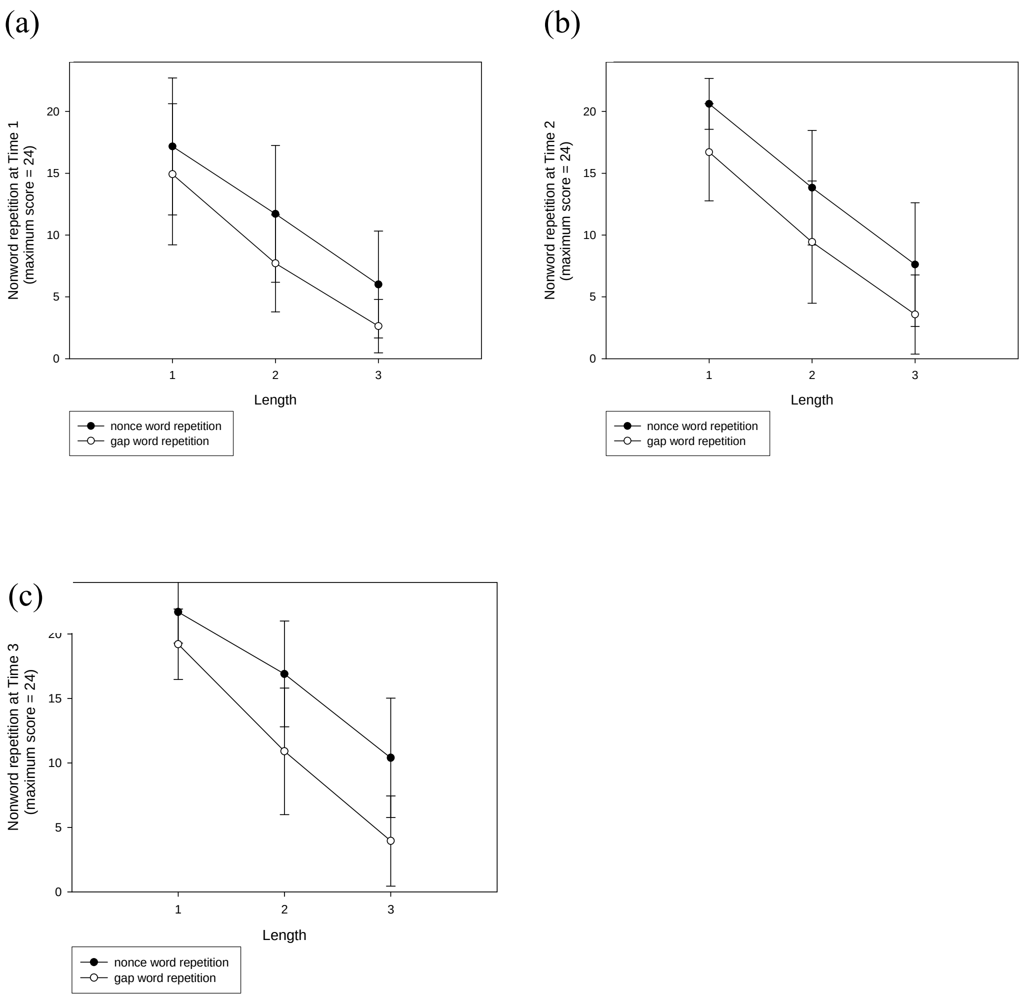
There were significant correlations between children's performances in the nonce word repetition and gap word repetition. Nonce word repetition performance at each time point was found to be highly associated with the concurrent performance in gap word repetition ($r = .85$-$.87$, $p < .001$). Nonce word repetition at age 3 was found to be associated with gap word repetition at age 3.5 ($r = .84$, $p < .001$) and gap word repetition at age 4 ($r = .84$, $p < .001$). Nonce word repetition at age 3.5 was also associated with gap word repetition at age 4 ($r = .87$, $p < .001$). The findings suggested that the relationship between nonce word repetition and the long term performance in gap word repetition was quite strong, which was different from what we had found in the age 2 cohort, who demonstrated modest-to-weak correlations. On the other hand, gap word repetition at age 3 was significantly associated with nonce word repetition at age 3.5 (r = .87, p < .001), and nonce word repetition at age 4 (r = .82, p < .001). Gap word repetition at age 3.5 was also correlated with nonce word repetition at age 4 (r = .79, p < .001). The findings implied that there was a close relationship between gap word repetition and the subsequent performances in nonce word repetition, which replicated the findings in the age 2 cohort.

To examine whether children improved in NWR across time, the repeated measures ANOVA was conducted with the variables of time, lexicality and length. The results demonstrated an interaction effect of lexicality and length ($F(2, 38) = 4.93$, $p = .012$, $\eta^2 = .21$), and also an interaction effect of lexicality and time ($F(2, 38) = 4.25$, $p = .022$, $\eta^2 = .18$). In addition, we found significant main effects of time ($F(2, 38) = 62.50$, $p < .001$, $\eta^2 = .77$), lexicality ($F(1, 19) = 283.45$, $p < .001$, $\eta^2 = .94$), and length ($F(2, 38) = 249.53$, $p < .001$, $\eta^2 = .93$). The three way interaction was at the borderline of significance, $F(4, 76) = 10.81$, $p = .062$, $\eta^2 = .11$.

Even though the three way interaction fell somewhat short of significance, it suggested that we should look at simple interaction effects. First we examined the length by lexicality interaction at each level of the time factor. Children's performance at each time point was graphed in Figure 5.8. The analyses showed that the length by lexicality interaction was statistically significant only at age 4, $F(2, 38) = 9.65$, $p < .001$, $\eta^2 = .34$, but not at age 3 or age 3.5 (both $p > .05$). The results suggested that at age 3 and age 3.5, children's performance in repeating nonce words was better than their performance in repeating gap words; however, both decreased in parallel as the length of the stimuli item increased. However, at age 4, the discrepancies between nonce word repetition and gap word repetition increased as the length of the stimuli item increased, though both exhibited downward trends (Figure 5.8 (c)). There was a difference in the decline rates of nonce word repetition and gap word repetition in relation to length at age 4. Nonce word repetition showed a gentler decline than gap word repetition did, which probably implied the raising support from lexical knowledge to nonce word repetition at age 4.

Figure 5.8

NWR Performances as the Function of Length and Lexicality in (a) Time 1, (b) Time 2 and (c) Time 3

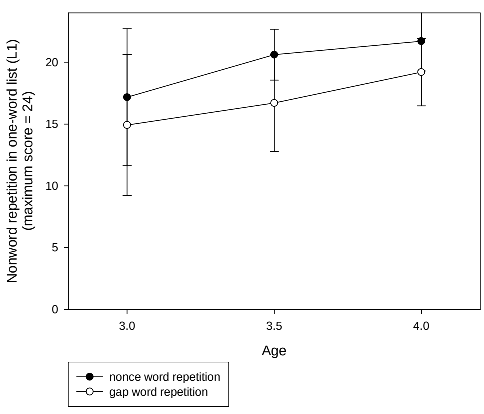


We also examined the lexicality by time interaction at each level of the length factor. Children's performances were graphed in Figure 5.9. The analyses showed that the lexicality by time interaction was statistically significant only in the three-word list (Length 3), $F(2, 38) = 4.14$, $p = .024$, $\eta^2 = .18$, but not in the one-word list (Length 1) or the two-word list (Length 2, both $p > .05$). The results suggested that in the repetition of the one-word lists and the two-word lists, children had better performance in nonce word repetition than performance in gap word repetition. However, both improved in similar rates, as reflected in the parallel growth curves. However, in the repetition of the three-word lists, children exhibited growth in their nonce word repetition performance over time, while showing little improvement in their repetition of gap words. In fact, the simple contrasts analysis on the gap word repetition three-word list across time showed that there was a significant increase from age 3 to age 3.5 ($t(22) = 1.80$, $p < .05$, one-tailed), but no significant improvement from age 3.5 to age 4 ($t(19) = 0.82$, $p > .05$, one-tailed). It appeared that children's performance in repeating three consecutive gap words levelled off at age 4.

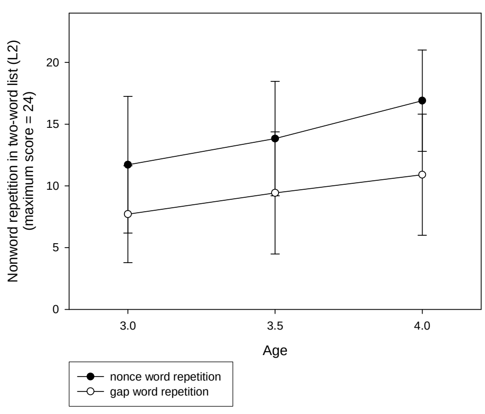
Figure 5.9

NWR Performances as the Function of Time and Lexicality in (a) the One-word Lists, (b) the Two-word Lists and (c) the Three-word Lists

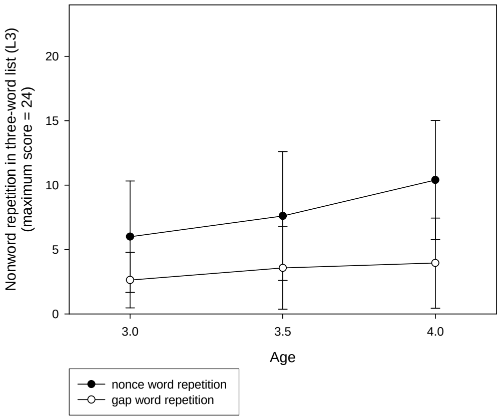
(a)



(b)



(c)



In addition to inspecting the two simple interaction effects, we also conducted the simple main effect analysis to examine the lexicality effect. We compared children's performances in nonce word repetition and gap word repetition at each time point and at each length. The results showed that the nonce word repetition performance always better than the gap word repetition performance (all $p < .05$). Therefore, children in this cohort were sensitive to the lexicality effect.

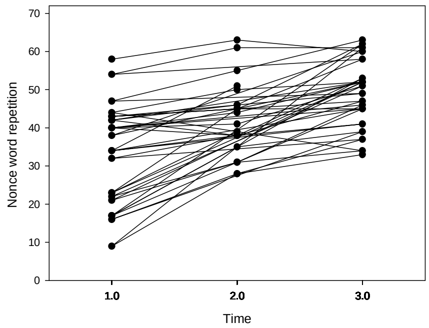
To summarize the findings above, lexicality effect was pervasive in children's performances. Children had better performance in repeating nonce words than performance in repeating gap words. Moreover, they performances decreased with the increase in the length of the stimuli items. Children's nonword repetition ability improved with age. Nevertheless, they appeared to show a ceiling in the encoding of gap word three-word list since age 3.5.

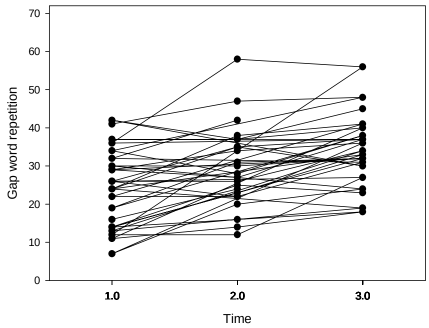
Each child's growth trajectories in the nonce word repetition and gap word repetition were plotted in Figure 5.10(a) and Figure 5.10(b), respectively. For both types of nonwords, children demonstrated noticeable variability in their initial performance. Though minor differences were observed in children's growth rates, they were more or less parallel.

Figure 5.10

Children's Individual Growth Curves of (a) Nonce Word Repetition and (b) Gap Word Repetition

a) b)





5.5 NWR, Phonological Capacities and Vocabulary Knowledge

In this section, we examined the predictability of children's phonological capacities and vocabulary knowledge to their nonword repetition performance. Preliminary stepwise regression analyses revealed that children's vocabulary knowledge and productive phonology (as measured by onset production) were associated with NWR, while word discrimination did not. Therefore, word discrimination was not included in the following analyses.

To specify the relative contribution of productive phonology and receptive vocabulary to NWR, we carried out the hierarchical regression analyses. Children's Leiter-R score was always entered as the first step to control for the variation in nonverbal intelligence. In the first set of the analyses, the productive phonology was entered prior to the receptive vocabulary (PPVT-R). In the second set of the analyses, their entry sequence was reverse.

Analyses on nonce word repetition and gap word repetition at age 3.5 (Table 5.5) revealed that productive phonology at age 3 was a more powerful predictor to nonword repetition at age 3.5, either the nonce word repetition or the gap word repetition, than PPVT-R at age 3. Productive phonology accounted for 24% of the variance in nonce word repetition even after the effects of nonverbal intelligence and receptive vocabulary knowledge were controlled (β = .50, p = .001). In fact, nonverbal intelligence and receptive vocabulary knowledge seemed to be involved in nonce word repetition, since their β values were at the marginal significance (p = .05 and p = .055, respectively). In the analysis on gap word repetition, the finding was partly in accordance with the findings in nonce word repetition. That is, productive phonology at age 3 accounted for significant variance in gap word repetition at age 3.5 (β = .45, $p < .05$). However, nonverbal intelligence and receptive vocabulary played no roles here.

Then we examined the predictability of Leiter-R, productive phonology, and PPVT-R at age 3 to NWR at age 4. The results on nonce word repetition revealed that only productive phonology made significant contribution (β = .41, $p < .05$) (Table 5.6). After Leiter-R and PPVT-R were controlled, productive phonology still accounted for 16% of the variance in nonce word repetition. Different results had been found in gap word repetition (Table 5.7). While the parallel analysis was conducted on gap word repetition at age 4, neither productive phonology nor receptive vocabulary made significant contributions. Leiter-R was found to account for 37% of the variance in gap word repetition at this age. However, its effect was not significance when children's productive phonology and receptive vocabulary at age 3 were considered in the model.

We also investigated the relationship between productive phonology and PPVT-R at age 3.5 to NWR at 4. Similar results were found. Productive phonology could account for 14% of the variance in nonce word repetition after controlling for the effect of Leiter-R and PPVT-R, though the effect was just at borderline significance (p = .053) (Table 5.6). Children's nonverbal intelligence appeared to play a role here, as Leiter-R accounted for 30% of the variance in nonce word repetition ($p < .05$). On the other hand, none of the variables made significant contributions to gap word repetition, except for an effect of nonverbal intelligence (β =.64, $p < .05$) (Table 5.7).

Table 5.5

The Hierarchical Regression Analyses on Nonce Word Repetition and Gap Word Repetition at Age 3.5 (Time 2)

Dependent variable		Independent variable	$\triangle R^2$	$\triangle F$	p	β	p
Nonce repetition-$_{age\ 3.5}$	1	Leiter-R	.40	13.73	= .001	.34	= .050
	2	Productive phonology-$_{age\ 3}$	.24	13.02	<.005	.50	= .001
	3	PPVT-R-$_{age\ 3}$	.07	4.20	= .055	.33	= .055
	2	PPVT-R-$_{age\ 3}$	.06	2.29	>.05	.33	= .055
	3	Productive phonology-$_{age\ 3}$	.24	15.36	= .001	.50	= .001
Gap repetition-$_{age\ 3.5}$	1	Leiter-R	.24	6.56	< .05	.34	>.05
	2	Productive phonology-$_{age\ 3}$	.20	6.86	< .05	.45	< .05
	3	PPVT-R-$_{age\ 3}$	.01	0.24	> .05	.11	> .05
	2	PPVT-R-$_{age\ 3}$	.01	0.16	>.05	.11	> .05
	3	Productive phonology-$_{age\ 3}$	.20	6.64	< .05	.45	< .05

Table 5.6

The Hierarchical Regression Analyses on Nonce Word Repetition at Age 4 (Time 3)

Step	Independent variable	$\triangle R^2$	$\triangle F$	*p*	β	*p*
1	Leiter-R	.30	7.58	< .05	.32	> .05
2	Productive phonology-$_{\text{age 3}}$	.17	5.51	< .05	.41	< .05
3	PPVT-R-$_{\text{age 3}}$	.02	.68	>.05	.20	>.05
2	PPVT-R-$_{\text{age 3}}$	.03	0.87	>.05	.20	>.05
3	Productive phonology-$_{\text{age 3}}$	.16	5.01	< .05	.41	<.05
1	Leiter-R	.30	7.58	< .05	.57	< .05
2	Productive phonology-$_{\text{age 3.5}}$	.16	5.01	< .05	.39	= .053
3	PPVT-R-$_{\text{age 3.5}}$	.02	0.73	> .05	-.19	> .05
2	PPVT-R-$_{\text{age 3.5}}$	.04	1.07	>.05	-.19	> .05
3	Productive phonology-$_{\text{age 3.5}}$	.14	4.38	= .053	.39	= .053

Table 5.7

The Hierarchical Regression Analyses on Gap Word Repetition at Age 4 (Time 3)

Step	Independent variable	$\triangle R^2$	$\triangle F$	p	β	p
1	Leiter-R	.37	10.39	= .005	.41	>.05
2	Productive phonology-$_{age\ 3}$	.09	2.96	>.05	.30	>.05
3	PPVT-R-$_{age\ 3}$	.02	.068	>.05	.20	>.05
2	PPVT-R-$_{age\ 3}$	.03	0.87	> .05	.20	>.05
3	Productive phonology-$_{age\ 3}$	.09	2.63	> .05	.30	>.05
1	Leiter-R	.37	10.39	= .005	.64	< .05
2	Productive phonology-$_{age\ 3.5}$	.06	1.82	> .05	.24	>.05
3	PPVT-R-$_{age\ 3.5}$	.02	0.43	> .05	-.15	>.05
2	PPVT-R-$_{age\ 3.5}$	.02	0.65	>.05	-.15	>.05
3	Productive phonology-$_{age\ 3.5}$	.05	1.52	>.05	.24	>.05

The findings in the above regression analyses revealed the predictive role of productive phonology to children's subsequent performances in repeating nonwords, particularly the nonce words. For example, children's productive phonology at age 3 was associated with their nonce word repetition at age 3.5 and age 4, and with gap word repetition at age 3.5. There was also a modest association between children's productive phonology at 3.5 and their nonce word repetition at 4.

The effect of productive phonology was further supported in the analysis in which we compared children with good or poor productive phonology in terms of their performances in nonce word repetition and gap word repetition. We divided children into two groups (good vs. poor) based on their *z* scores in onset production at each testing time. The repeated measures ANOVAs on the lexicality by production group effect at each testing time were performed. The results demonstrated significant production group effects at age 3 ($F(1,22) = 14.23$, $p = .001$) and at age 3.5 ($F(1,21) = 6.80$, $p < .05$), but not at age 4 ($F(1,18) = 0.93$, $p > .05$). At age 3 and at age 3.5, children with better productive phonology tended to have better performance in repeating nonwords. There was no difference between their performances in repeating nonce words or gap words, as evidenced by the insignificant lexicality by production group interaction (both $p > .05$). It was only at age 4 that we observed a marginal significant lexicality by production group interaction ($F(1,18) = 4.19$, $p = .056$). At age 4, children with better productive phonology had better performances in nonce word repetition in comparison with gap word repetition (though the difference was not significance, $t(18) = 1.52$, $p > .05$), while children with poor productive phonology had similar performances in nonce word repetition and gap word repetition.

While productive phonology was found to predict subsequent NWR performances, receptive vocabulary, as assessed by PPVT-R, played little role in NWR. We only found an association between children's receptive vocabulary at age 3 to nonce word repetition at age 3.5; however, the association did not reach the standard significance level. The finding was consistent with the finding in the age 2 cohort, in which we found that PPVT-R could not account for significant variance in children's subsequent NWR performances, though a receptive vocabulary effect was observed when the receptive vocabulary size assessed by MCDI-T was considered.

Though we did not see the role of vocabulary knowledge in the regression analyses, the pervasive and robust lexicality effect that we observed throughout this age range revealed the mediation of the vocabulary knowledge to NWR at the lexical level.

Analyses on this cohort showed that children in this age range were sensitive to the lexical status of the nonwords at the syllable level, regardless of age and length. Children in this cohort always had better performance in the repetition of nonce words than the repetition of gap words.

Does the lexicality effect differ in accordance with children's vocabulary sizes? We divided children into two groups (good vs. poor) based on their z scores in receptive vocabulary (as measured by PPVT-R) at each testing time, and compared the two groups in terms of their nonce word repetition and gap word repetition performances. The repeated measures ANOVAs revealed no lexicality by vocabulary group interaction at age 3 ($F(1,22) = 0.13$, $p > .05$), age 3.5 ($F(1,21) = 0.002$, $p > .05$) or age 4 ($F(1,18) = 0.09$, $p > .05$). However, there were significant main effects of vocabulary group at age 3 ($F(1,22) = 8.16$, $p < .01$) and age 4 ($F(1,18) = 4.46$, $p < .05$). Though the effect of vocabulary group at age 3.5 did not reach significance ($F(1,21) = 2.24$, $p > .05$), there was a tendency for children with larger vocabulary size to have better NWR performance. The findings suggested that children in this age range made use of their receptive vocabulary knowledge to support their repetition of nonce words, despite the difference in the vocabulary size.

The results of the regression analyses also manifested contributions from nonverbal intelligence to NWR. For example, Leiter-R significantly accounted for variance in nonce word repetition at age 3.5. Moreover, we found an increase in the contribution of Leiter-R to NWR at age 4 in the model involved productive phonology and receptive vocabulary at age 3.5, comparing to that in the model involved productive phonology and receptive vocabulary at age 3. The increase in the predictive power of nonverbal intelligence to NWR might be related to the improvements in productive phonology. As children's productive phonology matured with age, their performances gradually stabilized and the individual differences became smaller. Therefore, the variation in

NWR could not be predicted by productive phonology, but by children's ability to handle the complex processes of processing and storing the nonword stimuli at the same time.

The analyses above demonstrated how phonological capacities and vocabulary knowledge supported the repetition of nonwords in the age 3 cohort. It is also of interest whether NWR at age 3 could predict children's subsequent vocabulary knowledge. Correlation analyses showed that nonce word repetition at age 3 was correlated with REVT at age 3.5 ($r = .67$, $p < .001$) and REVT at age 4 ($r = .67$, $p = .001$). Gap word repetition at age 3 showed significant correlations with REVT at age 3.5 ($r = .68$, $p < .001$), REVT at age 4 ($r = .65$, $p < .005$), and PPVT-R at age 4 ($r = .50$, $p < .05$).

Regression analyses showed that nonce word repetition at age 3 could account for 22.8% of the variance in REVT at age 3.5 ($\beta = .51$, $p < .005$) and 24.6% of the variance in REVT at age 4 ($\beta = .54$, $p = .005$), when the nonverbal intelligence and age were controlled. On the other hand, gap word repetition at age 3 could account for 17.9% of the variance in REVT at age 3.5 ($\beta = .48$, $p < .01$) and 16.4% of the variance in REVT at age 4 ($\beta = .47$, $p < .05$). Neither nonce word repetition nor gap word repetition at age 3 accounted for significance variance in the subsequent PPVT-R when the variables of age and nonverbal intelligence were controlled.

Chapter 6 Vocabulary, Phonology, and NWR of the Age 4 Cohort

The chapter examines the age 4 cohort's development in vocabulary, phonology, and NWR. The relative contribution of phonological capacities and vocabulary development to NWR performances were also investigated at the end of the chapter. Consistent with findings in the age 2 cohort and age 3 cohort, the analyses showed that there was no significant gender effect; hence, the gender variable was no included in the following analyses. The performance data and the correlation relationship of the measures in discussion would be presented in the main text. However, a more detailed descriptive statistics of all the measures and an overall correlation matrix were provided in Appendix E(3) and F(3), respectively.

6.1 Participants

Children in this cohort were recruited at age 4 (age range = 47.36-50.80 months), and tested every six months. Therefore, they were tested at age 4 (Time 1), age 4.5 (Time 2) and age 5 (Time 3), respectively. Nine of these children were recruited by posts on parenting websites, while the other fifteen were recruited from kindergartens in Taipei City and New Taipei City. Twenty-four children were included at Time 1. All of them returned at Time 2. However, at Time 3, we lost one participants because the child left the kindergarten. The background information of the participants was summarized in Table 6.1.

Table 6.1

Information of the Child Participants in the Age 4 Cohort

	Time 1	Time 2	Time 3
N	24 (11♂)	24 (11♂)	23 (11♂)
Average age (month)	48.83	54.97	60.89
Age range	47.36-50.80	53.67-56.90	59.57-62.57

6.2 Vocabulary Development

Children in this age cohort receive standardized tests on vocabulary, just as children in the age 3 cohort did. Their expressive vocabulary was tested with the expressive vocabulary scale in REVT, and their receptive vocabulary was tested with the PPVT-R. Children's growth in each task would be examined respectively in the following sections.

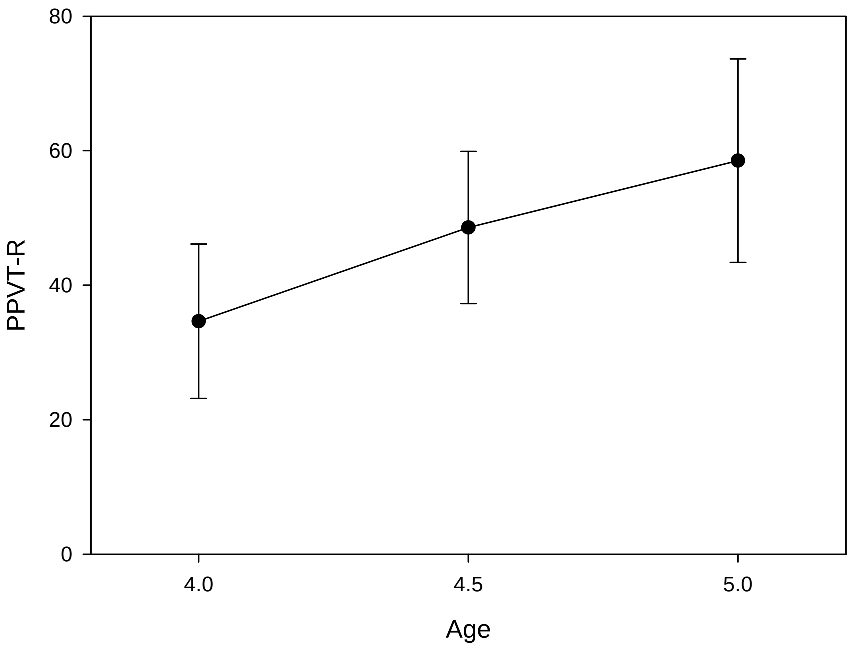
6.2.1 Receptive vocabulary size as measured by PPVT-R. Children at age 4 obtained an average score of 34.63 (SD = 11.48, range 20-65) on PPVT-R. At age 4.5, there was a prominent increase to the average score of 48.58 (SD = 11.32), with the scores ranging between 28 and 74. At age 5, children reached an average score of 58.52 (SD = 15.13), and the scores ranged between 33 and 83. Correlation analyses showed that PPVT-R scores of age 4, age 4.5 and age 5 were correlated with each other, with r ranging from .48 to .68 ($p < .05$).

One-way ANOVA on the effect of time revealed a significant effect, $F(2, 44) = 50.69$, $p < .001$, $\eta^2 = .70$. Paired t tests showed that there were significant increases in children's PPVT-R scores from age 4 to age 4.5 ($t(23) = 6.65$, $p < .001$), and from age 4.5 to age 5 ($t(22) = 4.17$, $p < .001$). A trend analysis showed that children's growth in PPVT-R was in a linear trend ($F(1,22) = 71.39$, $p < .001$), as demonstrated in Figure 6.1.

The findings suggested that children in from age 4 to age 5 demonstrated steady growth in receptive vocabulary.

Figure 6.1

Children's Average Performances in PPVT-R Across Time

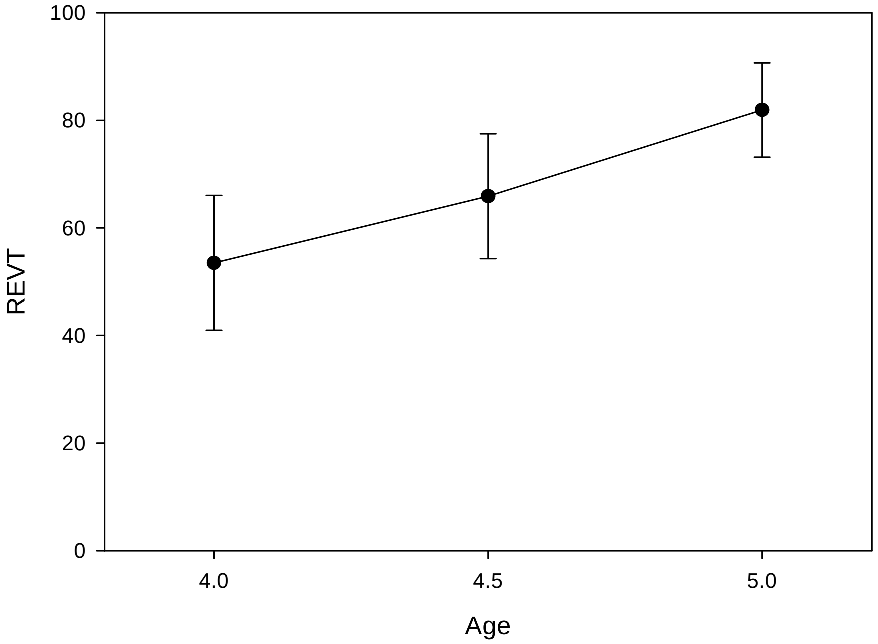


6.2.2 Expressive vocabulary size as measured by REVT. At age 4, children received an average score of 53.50 (SD = 12.52) in REVT, with the scores ranging between 27 and 81. At age 4.5, an average score of 65.88 was obtained (SD = 11.60, range = 47-87). When children reached age 5, they attained an average score of 81.91 (SD = 8.76), with the range of 63-94. REVT scores of age 4, age 4.5 and age 5 were correlated with each other ($r = .55\text{-}.67$, $p < .01$).

One-way ANOVA on the effect of time revealed that the effect was significant, $F(2, 44) = 88.43$, $p < .001$, $\eta^2 = .80$. Paired t tests showed that children's REVT scores increased from age 4 to age 4.5 ($t(23) = 5.75$, $p < .001$), and from age 4.5 to age 5 ($t(22) = 8.60$, $p < .001$). The trend analysis exhibited a linear trend in children's REVT score over time ($F(1,22) = 155.53$, $p < .001$), demonstrating that children had steady growth in expressive vocabulary (Figure 6.2).

Figure 6.2

Children's Average Performances in REVT Across Time



6.2.3 The correlation between receptive vocabulary and expressive vocabulary. The correlation relationships between the receptive vocabulary (as measured by PPVT-R) and the expressive vocabulary (as measured by REVT) of the three time points were demonstrated in Table 6.2. Each of the concurrent PPVT-R and REVT scores showed correlations ($r = .51\text{-}.68$, $p < .05$).

On inspecting the long-term relationship, we found that PPVT-R score at age 4 was not correlated with REVT at either age 4.5 or age 5. However, significant correlations were found between REVT at age 4 and PPVT-R at age 4.5 ($r = .54$, $p < .01$), and REVT at age 4 and PPVT-R at age 5 ($r = .57$, $p < .01$). Different from what we found in the age 3 cohort, findings here suggested that in this age range, expressive vocabulary knowledge could predict the subsequent development in receptive vocabulary, but not the other way around.

Table 6.2

Correlations Between PPVT-R and REVT Across Time

	PPVT @age 4.5	PPVT @age 5	REVT @age 4	REVT @age 4.5	REVT @age 5
PPVT @age 4	.59**	.48*	.51*	.38	.35
PPVT @age 4.5		.68**	.54**	.52**	.38
PPVT @age 5			.57**	.66**	.68**
REVT @age 4				.62**	.55**
REVT @age 4.5					.67**

6.2.4 Summary of vocabulary development from age 4 to age 5. Children in this age showed steady improvements in their vocabulary size. With respect to the development in receptive vocabulary, children in this cohort gained an average 12 points every half a year. As for the development in expressive vocabulary, there was an average of 14-point increase every half a year.

We also observed interactions between receptive vocabulary and expressive vocabulary sizes in development. Cross-time correlations were observed only between the earlier expressive vocabulary knowledge and the subsequent receptive vocabulary development. No correlations were found between the earlier receptive vocabulary knowledge and the subsequent expressive vocabulary. The finding was different from what we observed in the age 2 cohort and in the age 3 cohort. In the age 2 cohort, associations were found between receptive vocabulary at an earlier time and latter expressive vocabulary development. On the other hand, children in the age 3 cohort showed bi-directional cross-time interactions between receptive vocabulary and expressive vocabulary. Children in the age 4 cohort revealed an interaction pattern different from these two age cohorts. The findings in the three cohorts suggested a dynamic relationship between receptive vocabulary and expressive vocabulary in the developmental trajectory.

6.3 Phonological Development

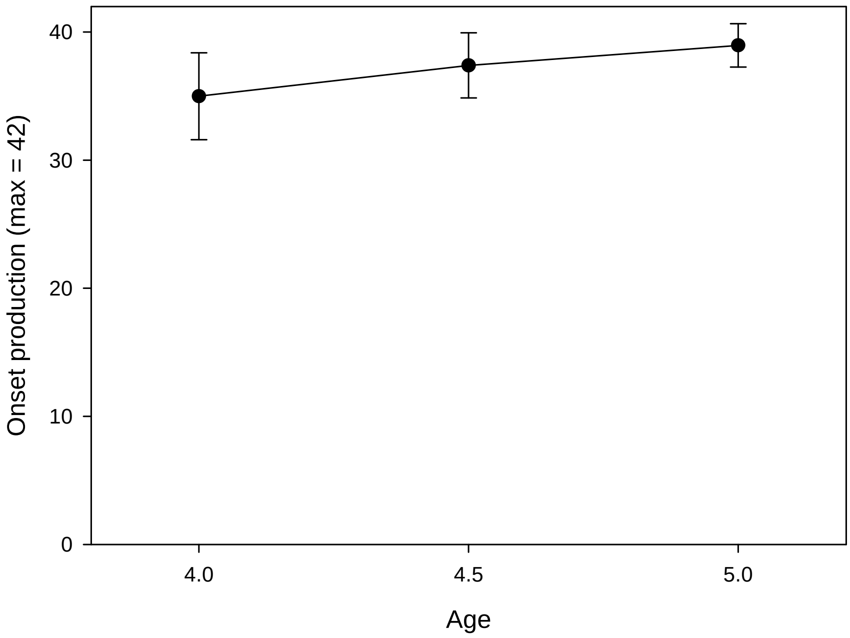
6.3.1 Productive phonology.

6.3.1.1 Onset production. Children at age 4 obtained an average score of 34.99 (SD = 3.39) in onset production (range = 21.14-38.83). There were small but steady increases at age 4.5 (mean = 37.40, SD = 2.54, range = 30.37-41.69) and age 5 (mean = 38.96, SD = 1.70, range = 34.80-41.83). Correlation analyses showed that onset production at the three time points were correlated with each other (r = .51-.78, $p < .05$). The average performances were demonstrated in Figure 6.3.

A one-way ANOVA was conducted to examine children's growth in onset production across time. There was a significant main effect of time, $F(2, 44) = 30.96$, $p < .001$, $\eta^2 = .59$. Planned comparisons were conducted to compare onset productions at each of the two adjacent time points, and found that children's onset production performance showed significant improvement from age 4 to age 4.5 ($t(23) = 4.77$, $p < .001$), and from age 4.5 to age 5 ($t(22) = 4.43$, $p < .001$). The trend analysis showed that the growth was linear ($F(1, 22) = 40.92$, $p < .001$.

Figure 6.3

Children's Average Performances in Onset Production Across Time



We further examined how the onset phonemes were stabilized in this age cohort, following the criterion of Zhu and Dodd (2000). Children in this age range had mastered all the onsets in Mandarin, except for the retroflexes, *zh, ch, sh,* and *r*. At age 4, still less than 50% of the children could accurately produce the retroflexes. It was until age 5 that about 75% of the children in this age could accurately produce these sounds. The findings suggested that retroflexes were still in development until the age of 5 (Table 6.3).

Children in this age range seldom made production errors. The common error patterns observed in this age range included fronting (e.g. *shou3-tao4* "glove" → *sou3-tao4*, *chu2-shi1* "chef" → *cu2-si1*) and gliding (e.g. *re4-gou3* "hotdog" → *le4-gou3*).

Table 6.3

Children's Development in Onset Production

Pass rate [a]	Age 4	Age 4.5	Age 5
> 75% (stabilized)	*b, p, m f, d, t, n, l, g, k, h, j, q, x, c, s*	*b, p, m f, d, t, n, l, g, k, h, j, q, x, z, c, s*	*b, p, m f, d, t, n, l, g, k, h, j, q, x, z, c, s*
50-75%	*z*	*zh*	*zh, ch, sh, r*
< 50%	*zh, ch, sh, r,*	*ch, sh, r,*	==

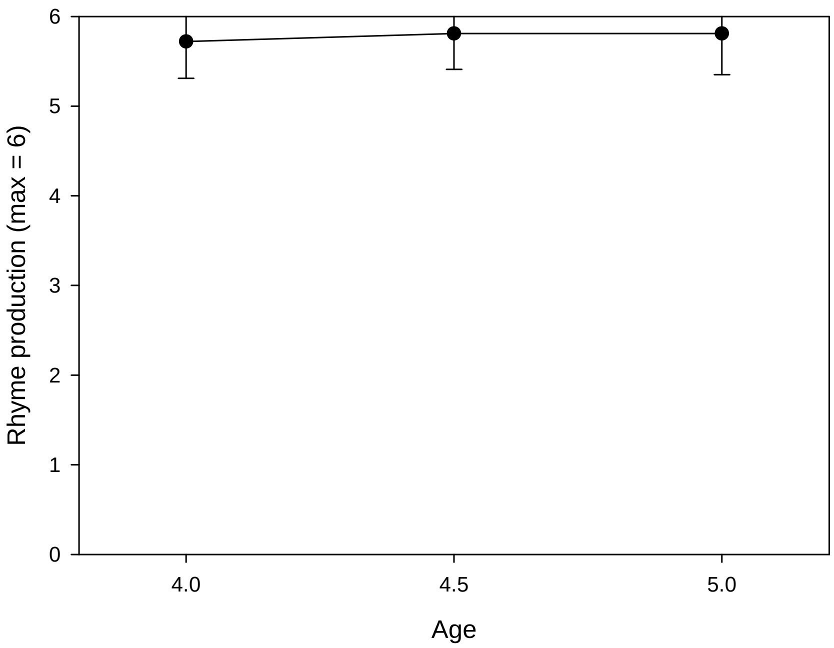
Note. [a] % of children passed the 66.7% accuracy rate

6.3.1.2 Rhyme production. Children's productions of six rhyme structures were examined. Children obtained an average score of 5.72 (SD = 0.41, range = 4.71-6.00) in rhyme production at age 4, which suggested that children had acquired most of the rhyme structures at age 4. Both at age 4.5 and at age 5, children reached an average of 5.81 (SD = 0.40 and 0.46, respectively). Children's scores at age 4.5 ranged between 4.19 and 6.00, and their scores at age 5 ranged between 4.14 and 6.00. Correlation analyses showed that rhyme production at the three time points were correlated with each other. Rhyme production at age 4 showed modest correlations with rhyme production at age 4.5 ($r = .52$, $p = .01$) and rhyme production at age 5 ($r = .53$, $p < .01$), while rhyme production at age 4.5 was highly correlated with rhyme production at age 5 ($r = .94$, $p < .001$).

A one-way ANOVA was conducted to compare children's rhyme production across the three time points. The main effect of time was not significant, $F(2, 44) = 1.09$, $p > .05$. Paired *t* tests were conducted to compare rhyme productions at each of the two time points, and neither showed significance (all $p > .05$). Children's production of rhymes had become mature at the age of 4. Thus, they showed stable performance throughout this age range.

Figure 6.4

Children's Average Performances in Rhyme Production Across Time



A further examination on children's production of each rhyme structure revealed that children had gained mastery of all of the rhyme types in Mandarin at age 4. The most difficult rhyme structures for young children, i.e. VN and GVN, had stabilized production at age 4 and onward.

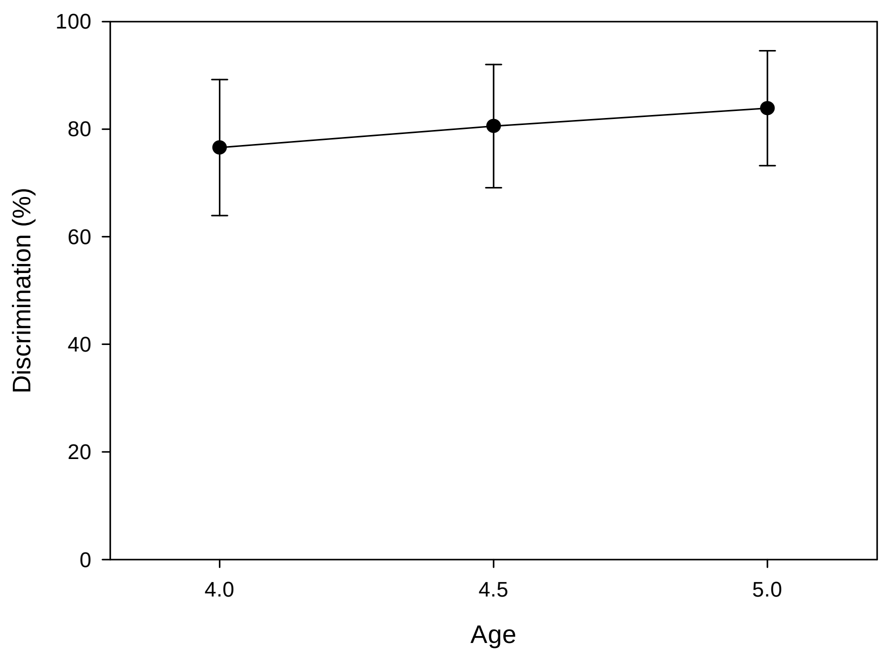
6.3.1.3 Correlation between the productions of onset and rhyme structure. Correlation analysis was conducted to examine the association between the production of onsets and rhyme structures. Children in this cohort showed no correlation between the production of onsets and rhymes. This was probably because children's production of onsets and rhymes had become stabilized at this age range, particularly the production of rhymes. Since children's production performances of onsets and rhymes became mature and demonstrated little variation across children, it was reasonable that they revealed no significant correlation.

6.3.2 Word discrimination. This measure assessed children's word discrimination ability at the phoneme level. Children in this cohort completed the same number of items across three testing sessions. Therefore, the subsequent statistical analyses were conducted with the raw scores.

At age 4, children had an average of 18.38 accurate items (SD = 3.03, range = 13-24), attaining an accuracy rate of around 76.56%. Half a year later, the average score increased to 19.33 (SD = 2.75, range = 13-24), attaining an accuracy rate of 80.56%. At age 5, children accurately responded an average of 20.13 items (SD = 2.56, range = 16-24), achieving an average accuracy rate of 83.88%. Children's average performances were graphed in Figure 6.5. Correlations were found between discrimination at age 4 and age 4.5 ($r = .64$, $p = .001$), between age 4 and age 5 ($r = .42$, $p < .05$), and between age 4.5 and age 5 ($r = .47$, $p < .05$).

Figure 6.5

Children's Average Performances in Word Discrimination Across Time

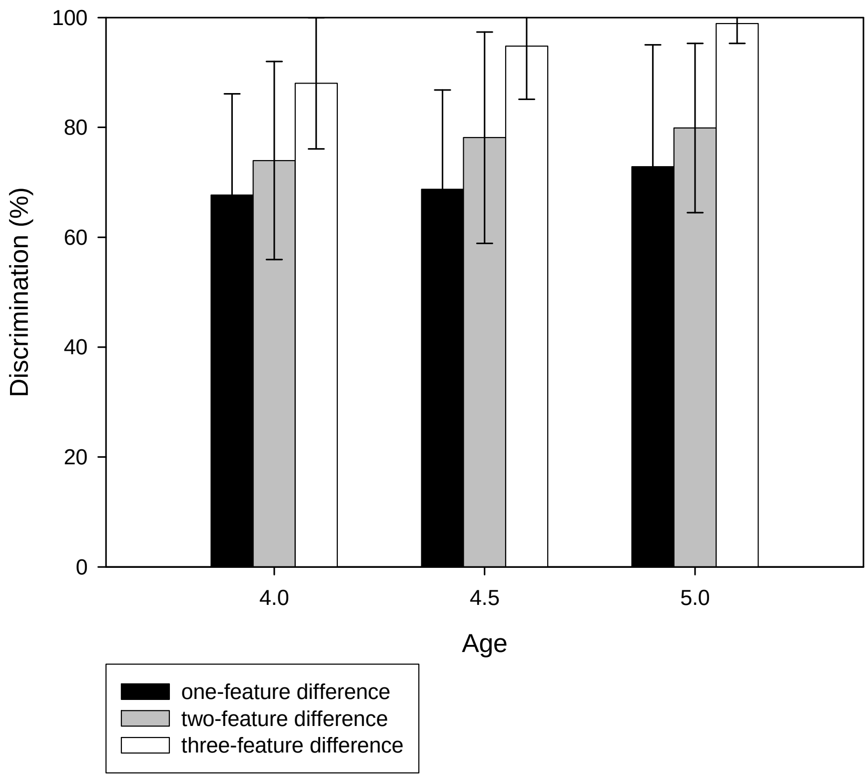


6.3.2.1 ***Sensitivity for different degrees of phonetic contrast.*** Figure 6.6 presented children's performances in discriminating minimal pairs of different degree of phonetic contrast. One-sample *t* tests showed that children's performances were above the chance level across all contrast pairs and all testing sessions (all $p < .001$, one-tailed).

A two-way ANOVA was conducted with the variables of time and degree of phonetic contrast on children's discrimination performance. The results revealed significant main effects of time ($F(2, 44) = 4.32$, $p < .05$, $\eta^2 = .16$), and contrast degree ($F(2, 44) = 44.34$, $p < .001$, $\eta^2 = .67$). The interaction effect did not reach significance ($F(4, 88) = 0.66$, $p > .05$).

Figure 6.6

Children's Average Performances in Discriminating Different Degrees of Phonetic Contrasts



The significant main effect of time suggested that children's sensitivity for phonetic contrasts at the phoneme level increased over time. Post hoc analyses showed that there was significant growth in their discrimination ability from age 4 to age 5 ($t(22) = 2.65$, $p = .015$). However, there was no significant difference either between age 4 and age 4.5 ($t(23) = 1.90$, $p > .05$), or between age 4.5 and age 5 ($t(22) = 1.52$, $p > .05$).

Also, post hoc analyses were conducted to examine the performances of the three types of contrast pairs. The results showed that children had better performance in discriminating contrasts with 3-*f* difference than contrasts with 1-*f* difference ($t(22) = 8.77$, $p < .001$) and contrasts with 2-*f* difference ($t(22) = 7.00$, $p < .001$). Children showed better sensitivity for contrasts with 2-*f* difference than contrasts with 1-*f* difference ($t(22) = 2.72$, $p = .013$).

The findings revealed that children showed sensitivity for all the contrast types, and their sensitivity for the contrasts improved in this age range. Children were more capable of discriminating minimal pairs with more prominent contrasts (i.e. those with 3-*f* difference) than pairs with less prominent contrasts (i.e. those with 2-*f* or 1-*f* differences).

6.3.3 Summary of phonological development from age 4 to age 5. Findings in this section showed that children in this age range mastered the production of most of the onsets and rhyme types. With regards to onset production, children at age 4 had acquired all the phonemes in Mandarin, except for the retroflexes. During age 4 to age 5, children gained prominent improvements in the productions of retroflexes; however, they were not stabilized until age 5. As for rhyme production, children mastered all of the rhyme types, though a few children were still having difficulty with those containing final nasal codas. The results in general suggested that children in this age range achieved near-adult performance in their phonological production.

In terms of children's discrimination ability, children showed a sensitivity for contrasts at the phonemic level, either when the contrasts were small or prominent. Children in this age range could easily discriminate phonemes of more prominent contrasts (e.g. *m-t* and *b-q*), since they had an average accuracy of 90% and above when discriminating these items. Though children performed above chance at discriminating contrasts with 1-*f* difference and contrasts with 2-*f* difference, these contrasts were still more difficult to be detected.

No significant correlations were found between children's phonological production ability and their discrimination ability.

6.3.4 The relationship between vocabulary knowledge and phonological capacities. The correlation analyses were conducted to examine phonological capacities, including onset production and word discrimination, and the receptive vocabulary.

With regards to the correlation between onset production and receptive vocabulary, no correlations were found (all $p > .05$).

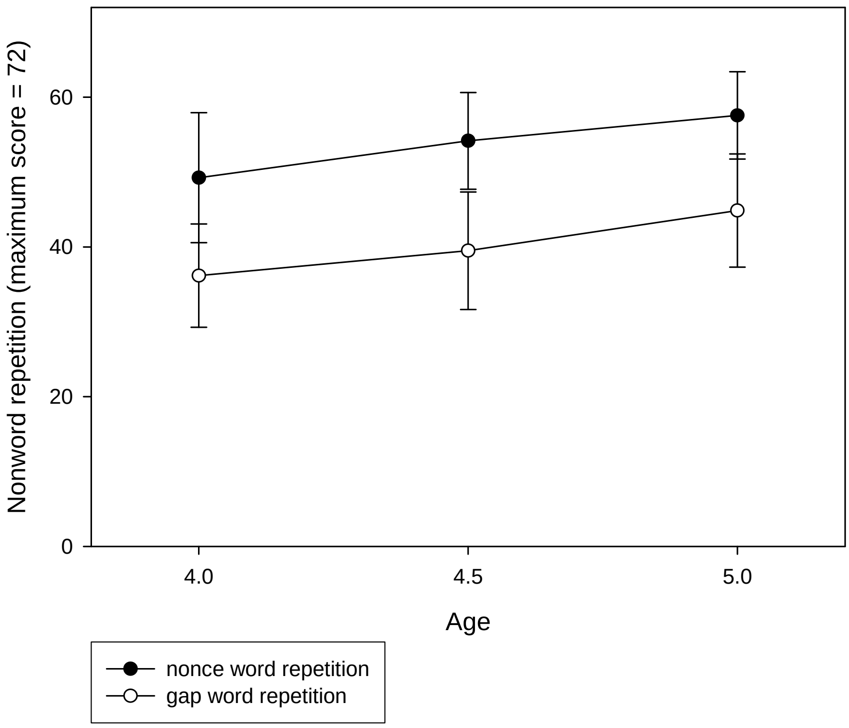
The analysis between word discrimination and receptive vocabulary revealed significant correlations between the earlier PPVT-R scores and the subsequent word discrimination performances. For example, we found a significant correlation between PPVT-R at age 4 and word discrimination at age 4.5 ($r = .44$, $p < .05$) and a marginal correlation between PPVT-R at age 4.5 and word discrimination at age 5 ($r = .39$, $p = .065$). Modest-to-weak correlations were found between PPVT-R at age 5 and word discrimination at age 4 ($r = .36$, $p = .093$) and at age 4.5 ($r = .35$, $p = .099$). However, both did not achieve the conventional significance level. In general, the results only demonstrated modest-to-weak correlations between the earlier receptive vocabulary and the subsequent word discrimination performances.

6.4. NWR Development

Children in this cohort performed the complete set of the nonword repetition. In total there were 36 disyllabic words in the nonce word repetition and the gap word repetition respectively. Each correctly repeated syllable would be credited with one point. As a result, the maximum score for the nonce word repetition or the gap word repetition was 72. At Time 1 and Time 2, 24 children finished the NWR task. At Time 3, 23 children participated in the task. Nevertheless, a child's data on gap word repetition was lost due to some technical problem. As a result, at age 5 there were 23 child data in the nonce word repetition and 22 child data in the gap word repetition. Children's average repetition performances in each type of the nonwords were graphed in Figure 6.7.

Figure 6.7

Children's Average Performances in Nonce Word Repetition and Gap Word Repetition



With regards to children's performance in nonce word repetition, they obtained an average score of 49.25 (SD = 8.68) at age 4, and an average score of 54.17 (SD = 6.46) at age 4.5. When they reached age 5, they gained an average score of 57.57 (SD = 5.82). In the gap word repetition, children attained an average score of 36.17 (SD = 6.90) at age 4, and an average score of 39.50 (SD = 7.84) at age 4.5. At age 5, children gained an average score of 44.86 (SD = 7.55).

Correlation analyses were conducted to examine (1) the correlations of repetition performances within each type of nonwords across time, and (2) the correlations between the nonce word repetition and the gap word repetition. The analyses on nonce word repetition showed that their performances at each time point was highly correlated with each other, with the r value ranging between .74-.82 ($p < .001$). Similar results had been observed in children's performances in gap word repetition, with the r values ranging between .68 and .70 ($p < .005$).

There were significant correlations between children's performances in the nonce word repetition and gap word repetition. Nonce word repetition performance at each time point was found to be highly associated with the concurrent performance in gap word repetition ($r = .70$-$.87$, $p < .001$). Nonce word repetition at age 4 was found to be associated with gap word repetition at age 4.5 ($r = .81$, $p < .001$) and gap word repetition at age 5 ($r = .66$, $p = .001$). Nonce word repetition at age 4.5 was also associated with gap word repetition at age 5 ($r = .86$, $p < .001$). The findings suggested a strong relationship between nonce word repetition and the subsequent performances in gap word repetition. On the other hand, gap word repetition at age 4 was significantly associated with nonce word repetition at age 4.5 ($r = .67$, $p < .001$), and nonce word repetition at age 5 ($r = .71$, $p < .001$). Gap word repetition at age 4.5 was also correlated with nonce word repetition at age 5 ($r = .80$, $p < .001$). Therefore, we also found a strong association between the gap word repetition and the subsequent performances in the nonce word repetition. Despite the distinction of lexicality between the nonce words and the gap words, children's performances in repeating these two different types of nonwords were closely related.

Then, the repeated measures ANOVA was conducted with the variables of time, lexicality and length. The results demonstrated an interaction effect of lexicality and length ($F(2, 42) = 44.67$, $p < .001$, $\eta^2 = .68$). We also found significant main effects of time ($F(2, 42) = 46.58$, $p < .001$, $\eta^2 = .69$), lexicality ($F(1, 21) = 450.20$, $p < .001$, $\eta^2 = .96$), and length ($F(2, 42) = 272.14$, $p < .001$, $\eta^2 = .93$). The three way interaction was at the borderline of significance, $F(4, 84) = 2.31$, $p = .065$, $\eta^2 = .10$.

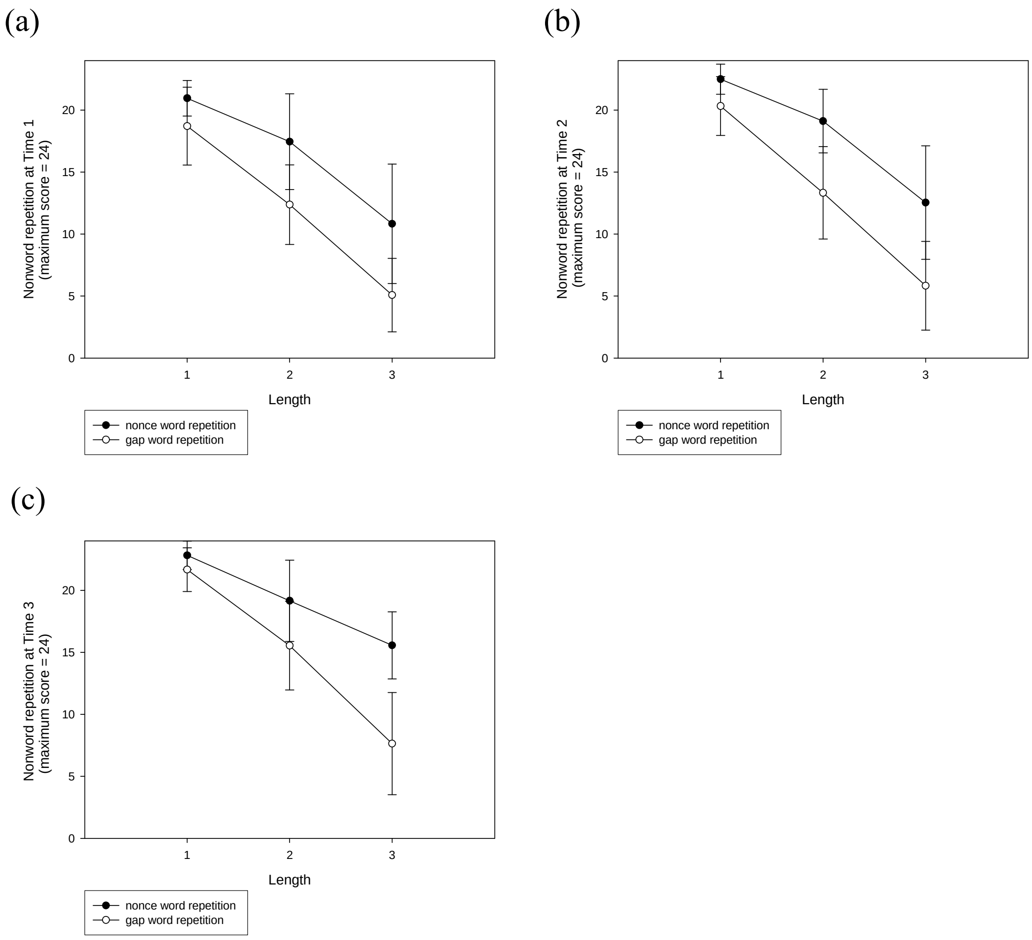
Even though the three way interaction fell somewhat short of significance, we should still look at the simple interaction effects. First, we examined the length by lexicality interaction at each level of the time factor. The results were graphed in Figure 6.8. The analyses showed that the length by lexicality interaction was statistically significant at age 4 ($F(2, 46) = 6.85$, $p < .005$, $\eta^2 = .23$), at age 4.5 ($F(2, 46) = 14.52$, $p < .001$, $\eta^2 = .39$), and at age 5 ($F(2, 42) = 36.90$, $p < .001$, $\eta^2 = .64$). Generally, children's performance in repeating nonce words was better than their performance in repeating gap words. However, the discrepancies between nonce word repetition and gap word repetition widened when children had to repeat two or more nonwords consecutively. Though the repetition performance declined as the length of the stimuli increased, there was a difference in the decline rates of nonce word repetition and gap word repetition in relation to length. Nonce word repetition showed a gentler decline than gap word repetition did.

In addition to inspecting the simple interaction effect on length and lexicality, we also conducted the simple main effect analysis to examine the lexicality effect. We compared children's performances in nonce word repetition and gap word repetition at each time point and at each length. The results showed that the nonce word repetition performance always better than the gap word repetition performance (all $p < .005$). Therefore, children in this cohort were sensitive to the lexicality effect.

To summarize the findings above, children's NWR performance improved over time. Their performance in repeating nonce words was better than the performance in repeating gap words. Moreover, it was found that their performances decreased with the increase in the length of the stimuli items. However, the decline rate of the nonce word repetition was smaller compared with that of the gap word repetition.

Figure 6.8

NWR Performances as the Function of Length and Lexicality in (a) Time 1, (b) Time 2 and (c) Time 3

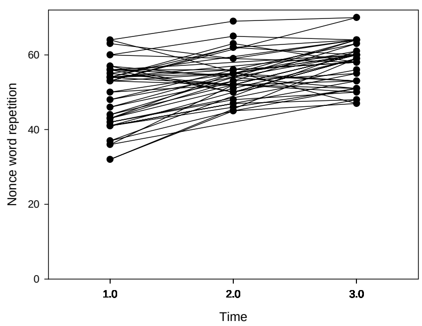


Each child's growth trajectories in the nonce word repetition and gap word repetition were plotted in Figure 6.9(a) and Figure 6.9(b), respectively. For both types of nonwords, children demonstrated noticeable variability in their initial performance. Though minor differences were observed in children's growth rates, they were more or less parallel.

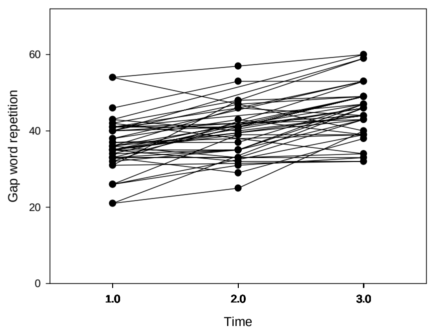
Figure 6.9

Children's Individual Growth Curves of (a) Nonce Word Repetition and (b) Gap Word Repetition

a)



b)



6.5 NWR, Phonological Capacities and Vocabulary Knowledge

In this section, we examined the predictability of children's phonological capacities and vocabulary knowledge to their nonword repetition performance. Preliminary stepwise regression analyses revealed that children's vocabulary knowledge and productive phonology (as measured by onset production) were associated with NWR, while word discrimination did not. Therefore, word discrimination was not included in the following analyses.

The hierarchical regression analyses were conducted to examine the relative contribution of productive phonology and receptive vocabulary to NWR. Children's

Leiter-R score was always entered as the first step to control for the variation in nonverbal intelligence. In the first set of the analyses, the productive phonology was entered prior to the receptive vocabulary. In the second set of the analyses, their entry sequence was reverse.

Analyses were conducted on nonce word repetition and gap word repetition at age 4.5 (Table 6.4). The analyses on the nonce word repetition revealed that productive phonology at age 4 was a more powerful predictor than PPVT-R at age 4 ($\beta = .50$, $p < .01$ vs. $\beta = .29$, $p > .05$). The finding in the analysis on gap word repetition was in accordance with the finding in nonce word repetition. That is, productive phonology was the only statistically significant predictor to gap word repetition performances at age 4.5 ($\beta = .48$, $p < .05$). Leiter-R did not make significant contribution to either nonce word repetition or gap word repetition.

Similar with previous findings, productive phonology at age 4 accounted for significant variance in children's nonce word repetition at age 5 ($\beta = .52$, $p < .05$), while receptive vocabulary and nonverbal intelligence played no role here (Table 6.5). However, the analysis on the gap word repetition at age 5 showed a different result. None of the predictive variables at age 4 served as a significant predictor to gap word repetition at age 5 (Table 6.6). Similar patterns were found when we examined the predictability of Leiter-R, productive phonology at age 4.5 and PPVT-R at age 4.5 to nonce word repetition at age 5 (Table 6.5) and gap word repetition at age 5 (Table 6.6).

The findings above showed that NWR, especially nonce word repetition, was predominantly predicted by productive phonology. The predictability of productive phonology to gap word repetition was only found when predicting gap word repetition at age 4.5 with productive phonology at age 4.

Table 6.4

The Hierarchical Regression Analyses on Nonce Word Repetition and Gap Word Repetition at Age 4.5 (Time 2)

Dependent variable		Independent variable	$\triangle R^2$	$\triangle F$	*p*	β	*p*
Nonce repetition-$_{\text{age 4.5}}$	1	Leiter-R	.06	1.35	> .05	.23	> .05
	2	Productive phonology-$_{\text{age 4}}$	.27	8.29	< .01	.50	< .01
	3	PPVT-R-$_{\text{age 4}}$	.08	2.63	> .05	.29	>.05
	2	PPVT-R-$_{\text{age 4}}$	.10	2.44	>.05	.29	>.05
	3	Productive phonology-$_{\text{age 4}}$	.25	8.27	< .01	.50	<.01
Gap repetition-$_{\text{age 4.5}}$	1	Leiter-R	.01	.22	>.05	.11	>.05
	2	Productive phonology-$_{\text{age 4}}$	.24	6.77	< .05	.48	< .05
	3	PPVT-R-$_{\text{age 4}}$	.03	0.94	> .05	.19	> .05
	2	PPVT-R-$_{\text{age 4}}$	.05	1.03	>.05	.19	> .05
	3	Productive phonology-$_{\text{age 4}}$	.23	6.40	< .05	.48	< .05

Table 6.5

The Hierarchical Regression Analyses on Nonce Word Repetition at Age 5 (Time 3)

Step	Independent variable	$\triangle R^2$	$\triangle F$	*p*	β	*p*
1	Leiter-R	.00	.02	> .05	.01	>.05
2	Productive phonology-$_{age\ 4}$	.29	8.02	= .01	.52	<.05
3	PPVT-R-$_{age\ 4}$	.05	1.45	> .05	.24	>.05
2	PPVT-R-$_{age\ 4}$	.07	1.53	>.05	.24	>.05
3	Productive phonology-$_{age\ 4}$	.27	7.60	< .05	.52	< .05
1	Leiter-R	.00	.02	> .05	-.03	>.05
2	Productive phonology-$_{age\ 4.5}$	.26	7.02	< .05	.54	< .01
3	PPVT-R-$_{age\ 4.5}$	.08	2.40	>.05	.31	> .05
2	PPVT-R-$_{age\ 4.5}$	.05	1.11	> .05	.31	>.05
3	Productive phonology-$_{age\ 4.5}$	.29	8.38	<.01	.54	<.01

Table 6.6

The Hierarchical Regression Analyses on Gap Word Repetition at Age 5 (Time 3)

Step	Independent variable	$\triangle R^2$	$\triangle F$	*p*	β	*p*
1	Leiter-R	.03	.63	>.05	.15	>.05
2	Productive phonology-$_{\text{age 4}}$	.10	2.28	>.05	.33	>.05
3	PPVT-R-$_{\text{age 4}}$	.05	1.21	> .05	.25	> .05
2	PPVT-R-$_{\text{age 4}}$	.06	1.15	>.05	.25	>.05
3	Productive phonology-$_{\text{age 4}}$	.10	2.29	>.05	.33	> .05
1	Leiter-R	.03	.63	>.05	.19	>.05
2	Productive phonology-$_{\text{age 4.5}}$	.07	1.57	>.05	.27	>.05
3	PPVT-R-$_{\text{age 4.5}}$	.00	.00	>.05	-.01	>.05
2	PPVT-R-$_{\text{age 4.5}}$	.00	.02	>.05	-.01	>.05
3	Productive phonology-$_{\text{age 4.5}}$	.07	1.48	>.05	.27	>.05

To further explore the effect of productive phonology on NWR, we compared the NWR performances between children with good production and children with poor production. We divided children into two groups (good vs. poor) based on their *z* scores in onset production at each testing time. The repeated measures ANOVAs on the lexicality by production group effect at each testing time were performed. The results showed that there was a tendency for children with good production to have better NWR performances than those with poor production, though the production group effect was only significant at age 4.5 ($F(1,22) = 5.03$, $p = .035$). There was no significant lexicality by production group interaction across time.

Though in regression analyses we did not see the role of vocabulary knowledge, the lexicality effect has been observed throughout this age range. The presence of the lexicality effect revealed the mediation of the vocabulary knowledge to NWR at the lexical level. As a result, children in this cohort always had better performance in the repetition of nonce words than the repetition of gap words.

We further examined whether the lexicality effect differed in accordance with children's vocabulary sizes. We divided children into two groups (good vs. poor) based on their *z* scores in receptive vocabulary (as measured by PPVT-R) at each testing time, and compared the two groups in terms of their nonce word repetition and gap word repetition performances. The repeated measures ANOVAs revealed no lexicality by vocabulary group interaction at age 4 ($F(1,22) = 0.43$, $p > .05$), age 4.5 ($F(1,22) = 0.10$, $p > .05$) or age 5 ($F(1,22) = 0.22$, $p > .05$). There was a tendency for children with larger vocabulary size to have better performance in repeating nonwords, even though the vocabulary group effect was significant only at age 4.5 ($F(1,22) = 7.26$, $p < .05$). Similar with the findings in the age 3 cohort, the findings suggested that children in this age range used vocabulary knowledge to support their repetition of nonce words.

The analyses above demonstrated how phonological capacities and vocabulary knowledge supported the repetition of nonwords in the age 4 cohort. It is also of interest whether NWR at age 4 could predict children's subsequent vocabulary knowledge. Correlation analyses showed that nonce word repetition at age 4 was correlated with REVT at age 4.5 ($r = .43$, $p < .05$) and REVT at age 5 ($r = .48$, $p < .05$). Gap word repetition at age 4 was not correlated with any of the vocabulary constructs at age 4.5 or age 5.

Regression analyses showed that nonce word repetition at age 4 could account for 17% of the variance in REVT at age 4.5 ($\beta = .42$, $p < .05$) and 21.9% of the variance in REVT at age 5 ($\beta = .47$, $p < .05$), when the nonverbal intelligence and age were controlled. On the other hand, gap word repetition at age 4 could not account for significant variance in REVT at age 4.5 or REVT at age 5. Neither nonce word repetition nor gap word repetition at age 4 accounted for significance variance in the subsequent PPVT-R when the variables of age and nonverbal intelligence were controlled.

Chapter 7 Developmental Trajectories of Vocabulary, Phonology, and NWR from Age 2 to Age 5

Our study was a cross-sequential design. That is, children of different age cohorts were followed for one year, tested at three occasions which spaced 6 months apart. The age 2 cohort and the age 3 cohort overlapped at age 3, and the age 3 cohort overlapped with the age 4 cohort at age 4. Therefore, the growth curves could be estimated on a combination of cross-sectional and longitudinal information.

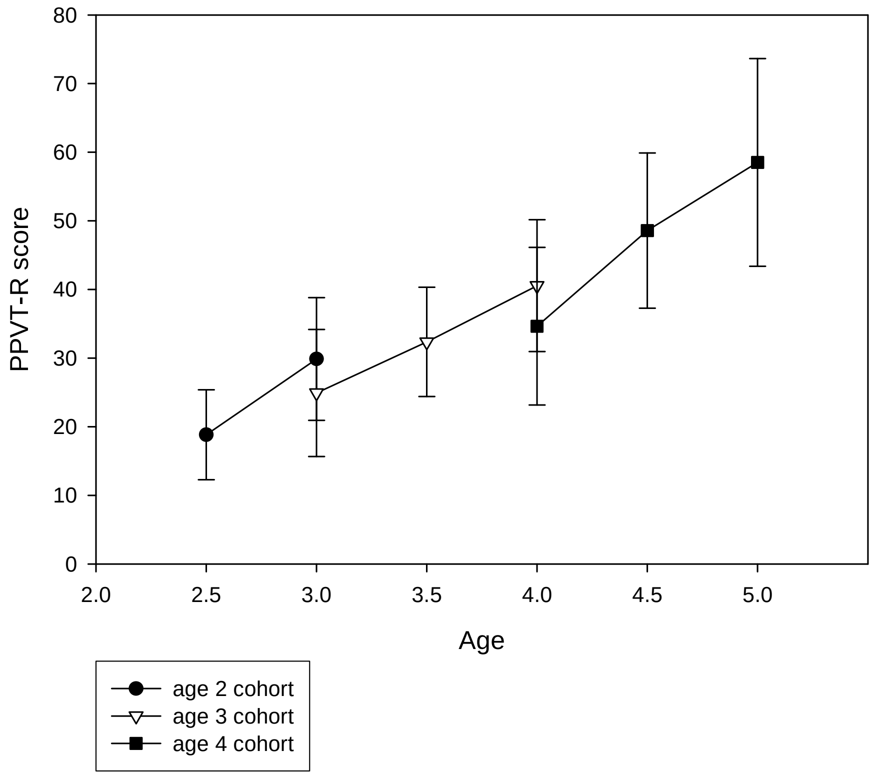
In this chapter, we tried to delineate the developmental trajectory of vocabulary knowledge, phonological capacities and nonword repetition across age 2 to age 5. Then, we would examine the developmental relationships of vocabulary knowledge and phonological capacities with NWR.

7.1 Vocabulary Development from Age 2 to Age 5

7.1.1 Receptive vocabulary size as measured by PPVT-R. We had children's PPVT-R scores from age 2.5 to age 5. Figure 7.1 demonstrated that children showed growth across ages. Independent *t* tests were conducted to compare different cohorts at the overlapping age. No significant difference was found between the PPVT-R scores of the age 2 cohort at age 3 and the PPVT-R scores of the age 3 cohort at age 3 ($t(44) = 1.84$, $p > .05$), nor a significant difference between the PPVT-R scores of the age 3 cohort and the age 4 cohort at age 4 ($t(42) = 1.83$, $p > .05$). Children from age 2.5 to age 5 showed steady growth in their receptive vocabulary.

Figure 7.1

The Growth Patterns in PPVT-R from Age 2.5 to Age 5

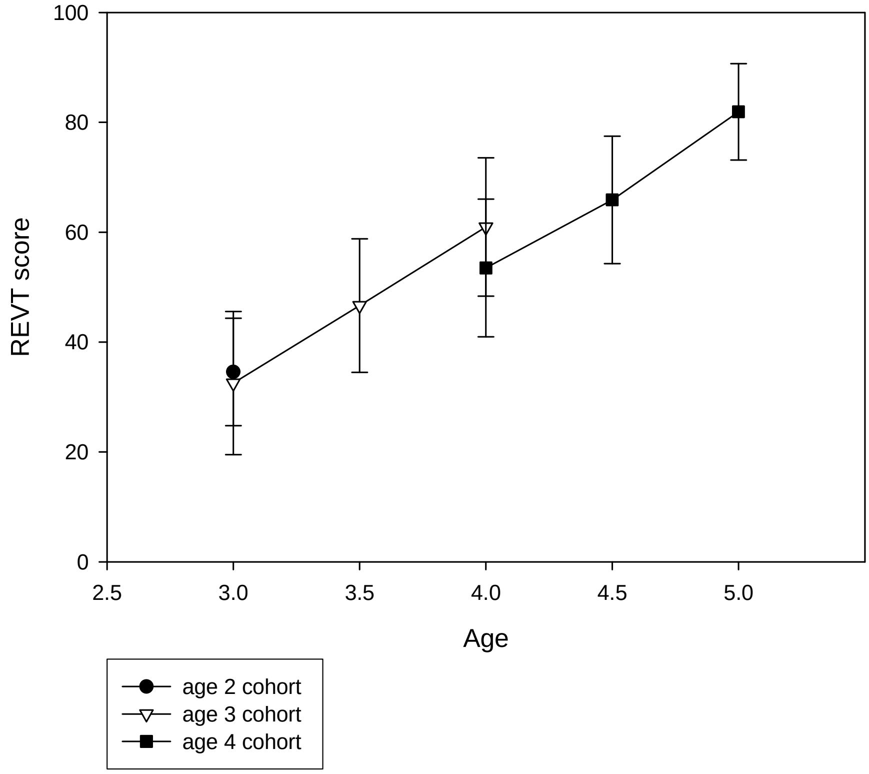


7.1.2 Expressive vocabulary size as measured by REVT. Children at age 3 and above received a standardize test on expressive vocabulary, i.e. REVT. Thus, we had children's REVT scores from age 3 to age 5. Figure 7.2 demonstrated that children showed steady growth in expressive vocabulary across ages. There was no significant difference between the REVT score of the age 2 cohort at age 3 and the REVT score of the age 3 cohort at age 3 ($t(43) = 0.58$, $p > .05$). The difference between the REVT score of the age 3 cohort and the age 4 cohort at age 4 approached but did not reach significance ($t(42) = 1.96$, $p = .057$). The findings in REVT, along with the findings in PPVT-R appeared to suggest that children grew linearly in their receptive and expressive vocabulary knowledge.

We conducted a correlation analysis on PPVT-R and REVT, while controlling the variable of age. The result showed that these two vocabulary constructs were modestly correlated with each other ($r = .50$, $p < .001$).

Figure 7.2

The Growth Patterns in REVT from Age 3 to Age 5



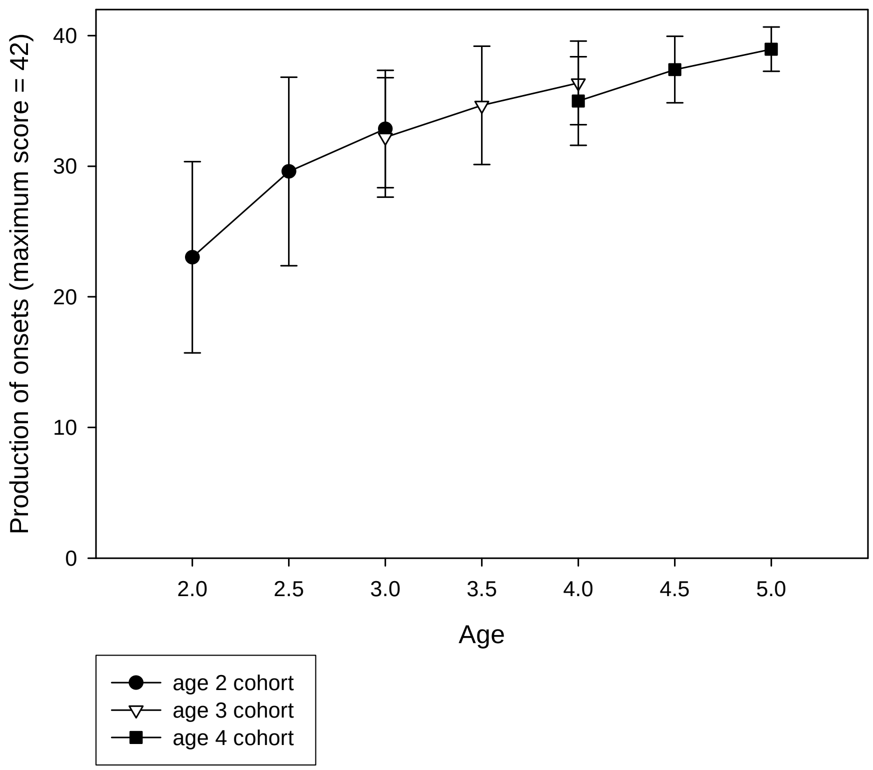
7.2 Phonological Development from Age 2 to Age 5

7.2.1 Productive phonology

7.2.1.1 Onset production. All the children in this study received a test on the phonological production throughout all the testing sessions. Figure 7.3 demonstrated that the developmental trajectory of children's production of onsets from age 2 to age 5. No significance difference was found between (1) the onset production scores of the age 2 cohort and the age 3 cohort at age 3 ($t(43) = 0.47$, $p > .05$), and (2) the onset production scores of the age 3 cohort and the age 4 cohort at age 4 ($t(42) = 1.39$, $p > .05$). The growth patterns appeared to demonstrate a quadratic trend. Children demonstrated a faster growth rate from age 2 to age 2.5, while they showed flatter and steady growth in the following ages.

Figure 7.3

The Growth Patterns in Onset Production from Age 2 to Age 5



In the previous chapters, we examined the onset phonemes that became stabilized in each age span, following the criterion of Zhu and Dodd (2000). The results of the three age cohorts were integrated and summarized in Table 7.1. The table demonstrated that at the age of 3, children had acquired 2/3 of the onsets in Mandarin. The acquired manners of articulation included stop, nasal, approximant, fricative and affricate. Also at this time, the children had learned to control the feature of aspiration. The onset phonemes that failed to stabilize at age 3 included the labio-dental fricative *f*, alveolar affricates *z, c*, and retroflexes *zh, ch, sh, r*. Retroflexes were particularly difficult for Mandarin-speaking children, because these sounds are still in development until the age of 5. Our findings were more or less similar to the findings in the past large-scale studies on Mandarin-speaking children's development in consonant production (Table 2.4).

Table 7.1

The Stabilized Onset Phonemes from Age 2 to Age 5

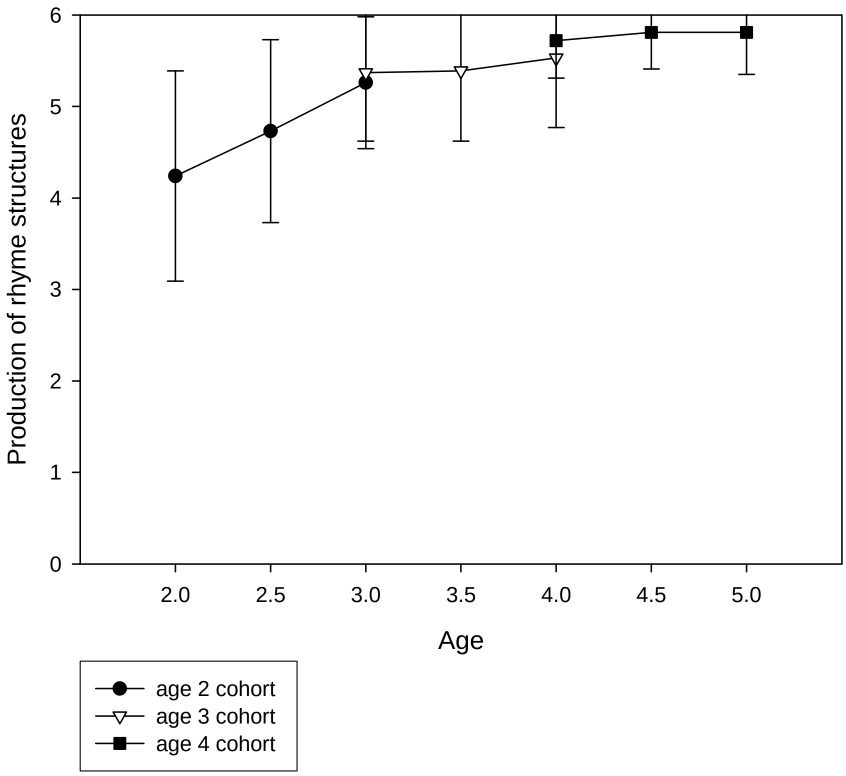
Pass rate[a]	Age 2	Age 2.5	Age 3	Age 3.5	Age 4	Age 4.5	Age 5
> 75% (stabilized)	*b, d, g, m, n*	*l, h, j*	*p, t, k, x, s, (q*[1]*)*	*f, q*	*(z*[2]*), c*	*z*	==

Note. 1 The aspirated affricate *q* was found to be stabilized in the age 3 cohort at age 3. Though *q* had not reached the stabilization rate in the age 2 cohort when the children were at 3, it actually was about to stabilize (69%). 2 The voiced affricate *z* was stabilized in the age 3 cohort at age 4. Despite not qualified for the stabilization rate, its production in the age 4 cohort at age 4 was close to stabilization (around 73%).

7.2.1.2 Rhyme production. The developmental trajectory of children's production of rhyme structures from age 2 to age 5 was presented in Figure 7.4. No significance difference was found between (1) the rhyme production score of the age 2 cohort and the age 3 cohort at age 3 ($t(43) = 0.50$, $p > .05$), and (2) the rhyme production score of the age 3 cohort and the age 4 cohort at age 4 ($t(42) = 1.09$, $p > .05$). Similar with the production of onsets, the growth patterns of rhyme productions appeared to demonstrate a quadratic trend. Children demonstrated a faster growth rate before age 3. After that, the growth curves became flatter.

Figure 7.4

The Growth Patterns in Rhyme Production from Age 2 to Age 5



We examined the rhyme structures that became stabilized in each age span following the criterion of Zhu and Dodd (2000) in the previous chapters. The results of the three age cohorts were integrated and summarized in Table 7.2. The table showed that the productions of rhyme types that do not involve a final coda had become stabilized before age 3. Rhyme types that contain a coda appeared to be difficult for young children; however, these sounds became stabilized at the age of 4. Most children at the age of 4 had mastered all the rhyme types in Mandarin, except for a few who had trouble with pronouncing the final codas. The development of rhymes is of faster pace than the development of onsets, as suggested in Zhu and Dodd (2000).

Table 7.2

The Stabilized Rhyme Structures from Age 2 to Age 4

Pass rate[a]	Age 2	Age 2.5	Age 3	Age 3.5	Age 4
> 75% (stabilized)	V	VG, GV, GVG	==	GVN	VN

A correlation analysis was conducted on the production of onsets and rhymes. The result showed that they were modestly correlated with each other (r = .51, $p < .001$) after the age factor was controlled.

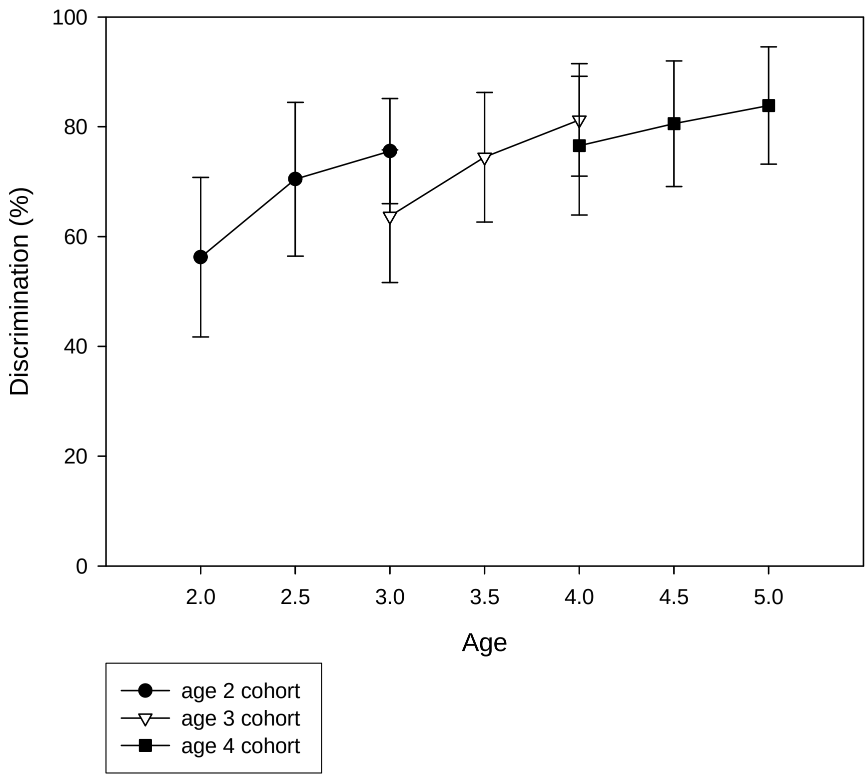
7.2.2 Word discrimination. All the children in the cohorts of this study received a test on the phonological perception in addition to the phonological production test. Figure 7.5 demonstrated the developmental trajectory of children's discrimination at the phoneme level across age 2 to age 5. It should be noted here that before age 3, children completed a shorter form of the discrimination task due to the concern for young children's attention span. Therefore, the following analyses were conducted with percentage scores.

We found no significant difference between the discrimination performances of the age 3 cohort and the age 4 cohort at age 4 ($t(42) = 1.33$, $p > .05$). However, when comparing the discrimination performances of the age 2 cohort and the age 3 cohort at age 3, the former group showed superior performance, and the difference had reached significance ($t(44) = 3.67$, $p = .001$). At age 3, children in the age 2 cohort reached an average accuracy of 75.58%, while children in the age 3 cohort had an average accuracy rate of 63.71%.

One possible explanation to this gap was that the children in the age 2 cohort did have better discrimination ability than children in the age 3 cohort. However, if this was true, then the advantage of the age 2 cohort should be accompanied with better performances in their other linguistic aspects that are proposed to be related to discrimination ability, such as their phonological production ability or vocabulary size (Best, 1984; Walley, 1993). However, as shown in previous analyses, these two groups of children were not different in those aspects. For this reason, we assumed that this would be a less probable explanation.

Figure 7.5

The Growth Patterns in Word Discrimination from Age 2 to Age 5



It was possible that the difference resulted from the practice effect. Since it was the third time the children in the age 2 cohort were asked to perform the task, they might have become more familiar with the procedure and the requirements of the task, and thus had better performance in the task. However, we think that the practice effect should be restricted to the familiarity with the procedure, not familiarity with the tested stimuli, because each testing session spaced 6 months apart. Young children might not remember the specific sounds they heard in the task. Therefore, within-cohort discrimination performances across time still to some extent reflected children's growth in discrimination ability.

In spite of the unbalanced performance between the age 2 cohort and the age 3 cohort, children from age 2 to age 5 in general exhibited growth in the discrimination at the phoneme level. Age 2 and age 3 cohorts appeared to show steady growth across time. However, children in the age 4 cohort had flatter growth curves in comparison with the other two age cohorts.

7.2.3 Summary of phonological development. From age 2 to age 5, children's phonological production abilities improved in a quadratic manner. They showed the greatest improvements before age 3, in their productions of onsets and rhymes. After age 3, they still revealed improvements, though the growth was less dramatic. In fact, by the age of 4, children had developed mature performances in rhyme production, while their production of onsets was still in development at the age of 5.

As for children's performance in discrimination phonemic level contrasts, children showed improvements mostly in their discrimination of prominent phonemic contrasts (i.e. contrasts with *3-f* difference). However, our task might not be able to distinguish between the age 2 cohort and the age 3 cohort, since they showed similar performances and growth rates across time. Children in the age 4 cohort tended to had better discrimination performance than the other two cohorts; however, the difference was not so prominent. Children's performances could be explained in terms of their insensitivity to phonemic level contrasts in metalinguistic tasks, as discovered in previous studies (Hu, 2004).

The correlation analysis on phonological production and discrimination showed that they were not significantly correlated with each other ($p > .05$), when the age factor was controlled.

7.2.4 The relationship between vocabulary knowledge and phonological capacities. The correlation analyses were conducted to examine phonological capacities, including onset production and word discrimination, and the receptive vocabulary. Only a significant correlation was found between onset production and PPVT-R ($r = .28$, $p < .001$), when the age factor was partialled out. Therefore, onset production was associated with receptive vocabulary knowledge, though the value of correlation coefficient was small. There was no significant correlation between PPVT-R and word discrimination when the age factor was controlled.

7.3 NWR Development from Age 2 to Age 5

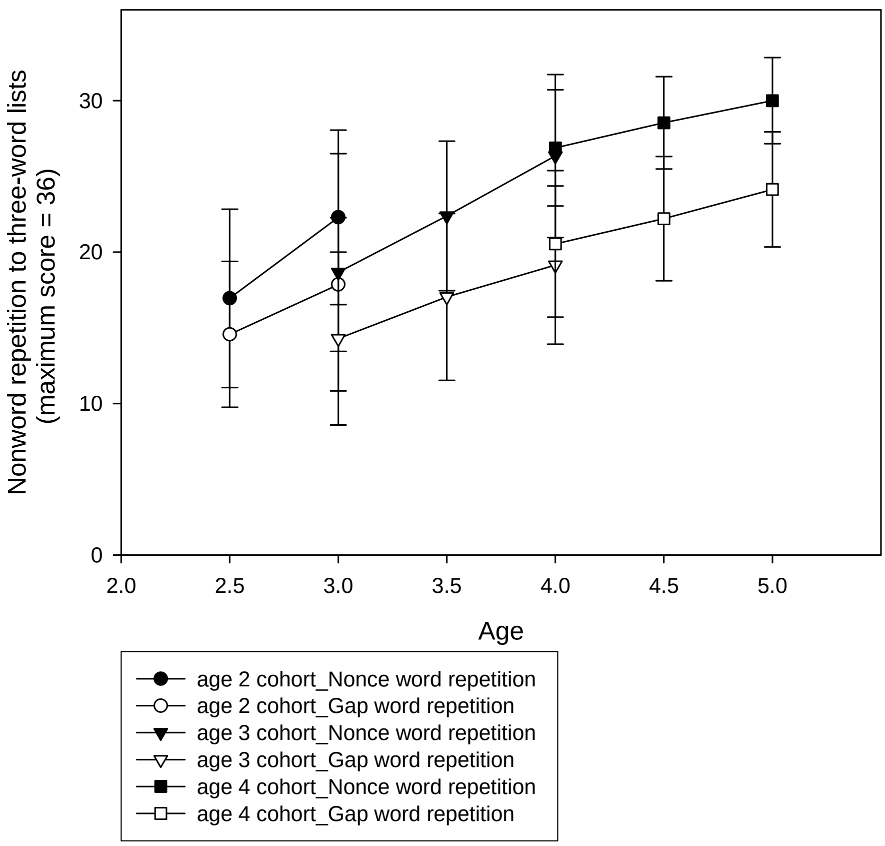
As mentioned in the chapter on methodology, children in the age 2 cohort completed a shorter form of the nonword repetition due to the concern for their shorter attention span. Since we would like to depict a developmental trajectory of NWR across age ranges, we examined children's minimum completion of NWR from age 2.5 to age 5, that is, half of the NWR task up to the three-word lists. Performances of the cohorts across the three testing occasions were presented in Figure 7.6.

Age 2 cohort and age 3 cohort at the overlapping age of 3 showed non-significant difference in their nonce word repetition ($t(43) = 1.79$, $p > .05$), but a significant difference in their gap word repetition ($t(43) = 2.32$, $p < .05$). On the other hand, age 3 cohort and age 4 cohort were not significantly different in either their repetition of nonce words ($t(38) = -0.06$, $p > .05$) or the repetition of gap words ($t(38) = 0.04$, $p > .05$).

Despite the significant difference in the age 2 cohort and the age 3 cohort regarding their NWR performances at age 3, children across the three cohorts showed a tendency of steady growth in their repetition of nonce words and gap words. The repetition of nonce words and gap words developed in parallel among each age group. However, the performance of repeating nonce words was better than gap words, except for the performances in the age 2 cohort. Previous analyses showed that the lexicality effect was not prominent until age 3.

Figure 7.6

The Growth Patterns in Nonce Word Repetition and Gap Word Repetition from Age 2.5 to Age 5



One of the main purposes of this study was to examine the developmental trajectory of nonword repetition, and the language factors associated with it. However, a prerequisite for modeling a growth curve on the NWR data is that the different cohorts were comparable. That is, children in different cohorts had comparable performances when they were at the same age (e.g. age 3), and also showed similar growth rate and growth direction (Bell, 1953). Therefore, we first carried out a test of convergence to test whether the growth curves from multiple cohorts of different ages could be regarded as a single growth trajectory.

This assumption was tested by using hierarchical linear models proposed by Miyazaki and Raudenbush (2000). In this approach, the age by cohort interaction effect is examined by comparing a full model and a reduced model. In the full model, the cohort effect was considered. The model assumed that each cohort had its own trajectory. On the other hand, the reduced model assumed that all cohorts followed a common trajectory. The details of the two models were presented in Appendix G. The two models would be compared by conducting a likelihood ratio test. If the difference between the two models was close to 0 (comparing the deviance values, which represented the fit measure of a model), then we could retain the hypothesis that the growth curves of these cohorts in fact represented a common expected developmental trajectory for the entire population.

We tested the hypothesis in children's nonce word repetition and gap word repetition, respectively (Table 7.3). For nonce word repetition, the deviance value for the full model was 976.060, with $df = 10$ and the deviance value for the reduced model was 980.200, with $df = 8$. Thus the difference between the deviance values was 4.14, with $df = 2$, $p > .05$. A similar result was found for gap word repetition. The deviance value for the full model was 993.394, with $df = 10$ and the deviance value for the reduced model was 998.355, with $df = 8$. Thus the difference between the deviance values was 4.96, with *df* = 2, $p > .05$.

Both the results showed that the addition of cohort effect to the full model reduced deviance by a non-significant amount; therefore, there was no significant difference between the models in representing the growth patterns of the three cohorts. In other words, we could regard the growth curves of the three cohorts as representing a common expected developmental trajectory for the entire population. The departures of cohort-specific growth curves from the developmental trajectory could be dismissed as chance differences.

Table 7.3

Results of the Convergence Analyses on Nonce Word Repetition and Gap Word Repetition

	Model fit measures		Likelihood ratio test		
	Deviance	No. of estimated parameters	χ^2	No. of *df*	*p*
Nonce word repetition					
Reduced model	980.200	8			
Full model	976.060	10	4.14	2	>.05
Gap word repetition					
Reduced model	998.355	8	4.96	2	>.05
Full model	993.394	10			

7.4 The Developmental Trajectory of NWR and the Role of Phonological Capacities and Vocabulary Knowledge

A preliminary way to examine the contribution of phonological capacities and vocabulary knowledge to NWR from age 2 to age 5 was to conduct regression analyses. The results were demonstrated in Table 7.4. Both productive phonology and PPVT-R were found to contribute significantly to nonce word repetition when the nonverbal intelligence was controlled. Comparing these two involving factors, PPVT-R was of higher predictability than productive phonology, as reflected by its higher β value (PPVT-R: $\beta = .50$; productive phonology: $\beta = .33$). Similar results were found for gap word repetition. Children's gap word repetition was associated with their PPVT-R and productive phonology. Again, of these two factors, PPVT-R showed higher predictability (PPVT-R: $\beta = .43$; productive phonology: $\beta = .21$). However, different from the repetition of nonce words, the repetition of gap words involved the influence from nonverbal intelligence ($\beta = .19$, $p < .005$).

The regression analyses might provide us preliminary insights of the relationships. However, our study was a cross-sequential study, which incorporated not only cross-sectional information from different age cohorts, but also longitudinal information from individual child participants. Our data showed that there was considerable variation across participants; therefore the ordinary regression methods might suffer from correlated error.

The hierarchical linear model (hereafter abbreviated as HLM) provided a more proper approach to explore our data, because longitudinal data can be viewed as multilevel data with repeated measurements nested within individuals (Hox, 2010, p. 79). In HLM, there would be models at two levels. In the *level 1 model*, each child's development was represented by an individual growth trajectory that depended on a unique set of person-specific parameters. These individual growth parameters became the outcome variables in a *level 2 model*, in which they might depend on some person-level characteristics.

Table 7.4

Hierarchical Regression Analyses on Nonword Repetition

Dependent variable	Independent variable	△R^2	△F	p	β	p
Nonce repetition	1. Leiter-R	.06	11.85	= .001	.06	>.05
	2. Productive phonology	.37	114.81	< .001	.33	< .001
	3. PPVT-R	.14	58.86	< .001	.50	< .001
	2. PPVT-R	.45	158.93	< .001	.50	< .001
	3. Productive phonology	.07	28.09	< .001	.33	< .001
Gap repetition	1. Leiter-R	.12	23.80	< .001	.19	< .005
	2. Productive phonology	.21	54.09	<.001	.21	< .005
	3. PPVT-R	.11	32.53	< .001	.43	< .001
	2. PPVT-R	.29	83.72	< .001	.43	< .001
	3. Productive phonology	.03	8.88	< .005	.21	< .005

The outcome variables for our analysis were children's performances in nonce word repetition and gap word repetition. The variables PPVT-R and productive phonology would be considered in the HLM models. Since the growth in vocabulary knowledge and phonological capacities varied with time (i.e. time-varying covariates), they would be incorporated at Level 1. The time-invariant covariate, such as nonverbal intelligence (or Leiter-R), would be incorporated at Level 2.

First we created a null growth model in which we only considered the variable of age. The purpose of the null model was to see whether HLM was needed to examine our data. Also it could serve as a baseline for evaluating the explanatory power of other models. Equation 7.1 presented the level 1 model for the null model. $\text{Nonword repetition}_{ti}$ was the nonword repetition performance of the child *i* measured at measurement occasion *t*. Since we modelled children's growth pattern from age 2.5 (30 months) to age 5, the component $(age - 30)$ represented the linear slope of the growth curve with age 2.5 as the onset.

Level 1

$$\text{Nonword repetition}_{ti} = \pi_{0i} + \pi_{1i}(age - 30) + e_{ti} \qquad (7.1)$$

The value of π_{0i} represented the level 1 intercept, or the expected score of person *i* at age 2.5. The value of π_{1i} was the expected rate of increase in the score for person *i* at age 2.5. Finally, e_{ti} was the random within-person error of prediction for person *i* at time *t*.

The coefficients in the level 1 models were the outcomes of the level 2 models (Equation 7.2). The coefficients signaled by βs represented between-participants parameters. For example, β_{00} represented the estimate of the average intercept across children. Similarly, β_{10} stood for the estimate of the average slope across participants. u_{0i} and u_{1i} were between-participants variance.

Level 2

$$\pi_{0i} = \beta_{00} + u_{0i}$$

$$\pi_{1i} = \beta_{10} + u_{1i} \qquad (7.2)$$

The analysis results on nonce word repetition were presented in Table 7.5. The model predicted a value of 17.76 at age 2.5, which increased by 0.45 with the increase in age (month). The variance components for the intercept and the regression slope for the time variable were both significant, which meant that individuals had different initial states, and different rate of change. The estimate of the level 1 variance was 7.24, while the subject-level (level 2) variance was 42.86 (42.82 + 0.04). The intra-class correlation (ICC) was estimated as 42.86/(42.86 + 7.24) = 85.55%. About 86% of the variance was variance between individuals. The finding qualified our study for the HLM analysis.

The second model added the time-varying covariate productive phonology to the model. The effect of productive phonology was significant: better productive phonology led to superior performance in nonce word repetition. Using the first model as the baseline, productive phonology explained 23% of the variance between children's intercepts and 25% of the variance between children's rates of change. Even though productive phonology is a time-varying predictor, it explained more variation between different children than within the same children from one testing time to the next.

The third model replaced productive phonology with another time-varying covariate PPVT-R. The effect of PPVT-R was significant, suggesting that the larger a child's receptive vocabulary knowledge, the higher the nonce word repetition score. Using the first model as the baseline, PPVT-R explained 9.60% of the variance between children's intercepts. However, it did not explain additional variance in children's rates of change.

In the fourth model, we entered both time-varying covariates PPVT-R and productive phonology. Both effects were significant. However, age became a non-significant factor (p = .07), suggesting that lexical and phonological knowledge could account for the NWR growth. High nonce word repetition performance correlates with high receptive vocabulary knowledge, and better productive phonology. The inclusion of the two variables to the model had explained 27.60% of the variance between children's intercepts and 25% of the variance between children's rates of change.

In the fifth model, we added the time-invariant covariate Leiter-R score to account for the differences in children's initial state and rates of change. However, the effect was not significant.

The likelihood tests were conducted to compare all the models. In general, models with a lower deviance fit better than models with a higher deviance. We found that Model 2 and Model 3 were significantly better than Model 1 (Model 2: $\chi^2(1) = 13.78$, $p < .001$; Model 3: $\chi^2(1) = 7.34$, $p < .01$), while Model 2 and Model 3 were not significantly different in model fit ($p > .05$). Comparing Model 4 with Model 2 and with Model 3, Model 4 showed significantly better fit (Model 2: $\chi^2(1) = 9.30$, $p < .005$; Model 3: $\chi^2(1) = 15.73$, $p < .001$). The difference between Model 4 and Model 5 revealed no significant difference ($\chi^2(2) = 0.12$, $p > .50$).

As a result, the fourth model was preferred to account for the developmental trajectory of nonce word repetition. Both children's productive phonology and receptive vocabulary knowledge could account for children's initial performance in nonce word repetition, even when the factor of age was controlled. Though the fourth model also appeared to account for children's variances in rates of change, the results of the second model and the third model suggested that the effect might come from productive phonology.

Table 7.5

Hierarchical Linear Models on Nonce Word Repetition

Model	M1: null model	M2: + production	M3: + PPVT-R	M4: + production + PPVT-R	M5: + production + PPVT-R + Leiter-R
Fixed part					
Intercept	17.76 (0.97)**	19.44 (0.93)**	20.60 (1.30)**	22.34 (1.20)**	22.37 (1.35)**
Child age	0.45 (0.04)**	0.33 (0.04)**	0.25 (0.08)**	0.13 (0.07)	0.13 (0.08)
Production		0.37 (0.07)**		0.38 (0.07)**	0.38 (0.07)**
PPVT-R			0.09 (0.03)**	0.09 (0.03)**	0.09 (0.03)**
Leiter*intercept					-0.02 (0.06)
Leiter*age					0.00 (0.00)
Random part					
u_{0i}	42.82**	32.97**	38.71**	31.00**	31.10**
u_{1i}	0.04**	0.03*	0.04*	0.03*	0.03**
e	7.24	7.84	7.64	8.03	7.99
Deviance	995.16	981.38	987.82	972.09	971.97
Parameters No.	6	7	7	8	10

Parallel analyses were conducted on gap word repetition, and the results were presented in Table 7.6. The model predicted a value of 13.78 at age 2.5, which increased by 0.35 with the increase in age (month). The variance component for the intercept was significant, which meant that children had different initial states. However, children appeared to demonstrate similar rates of change, as reflected by the non-significant variance component for the regression slope for the time variable. The estimate of the level 1 variance was 8.48, while the subject-level (level 2) variance was 19.99 (19.99 + 0.00). The intra-class correlation (ICC) was estimated at 19.99/(19.99+8.48) = 70.21%. About 70% of the variance was variance between individuals. Thus, the finding justified our application of HLM analysis on gap word repetition.

The second model added the time-varying covariate productive phonology score to the model. The effect of productive phonology was significant: better productive phonology led to superior performance in nonce word repetition. Using the first model as the baseline, productive phonology explained 9.85% of the variance between children's intercepts. Again, we found that even though productive phonology is a time-varying predictor, it explained more variation between different children than within the same children from one testing time to the next.

The third model replaced productive phonology with another time-varying covariate PPVT-R. The effect of PPVT-R only approached significant (p = .072), suggesting only a tendency for children with larger receptive vocabulary knowledge to have higher gap word repetition score. Comparing with the first model, PPVT-R explained 11.26% of the variance between children's intercepts.

In the fourth model, we entered both time-varying covariates PPVT-R and productive phonology. While the effect of productive phonology was significant, the effect of PPVT-R was only at the borderline of significance (p = .078). The inclusion of the two variables to the model had explained 17.06% of the variance between children's intercepts.

Table 7.6

Hierarchical Linear Models on Gap Word Repetition

Model	M1: null model	M2: + production	M3: + PPVT-R	M4: + production + PPVT-R	M5: + production + PPVT-R + Leiter-R
Fixed part					
Intercept	13.78 (0.74)**	14.66 (0.74)**	16.01 (1.33)**	16.69 (1.34)**	16.03 (1.43)**
Child age	0.35 (0.03)**	0.29 (0.04)**	0.19 (0.08)*	0.15 (0.08) (p = .079)	0.19 (0.09)*
Production		0.18 (0.08)*		0.17 (0.07)*	0.16 (0.07)*
PPVT-R			0.07 (0.04)	0.07 (0.04)	0.05 (0.04)
Leiter*intercept					0.04 (0.05)
Leiter*slope					0.00 (0.00)
Random part					
u_{0i}	19.99**	18.02**	17.74**	16.58**	16.40**
u_{1i}	0.00	0.00	0.00	0.00	0.00
e	8.48	8.63	8.80	8.88	8.70
Deviance	1001.25	997.39	997.24	993.69	991.94
Parameters No.	6	7	7	8	10

In the fifth model, we added the time-invariant covariate Leiter-R score to account for the differences in children's initial state and rates of change. However, the effect was not significant.

The likelihood tests were conducted to compare all the models. We found that Model 2 and Model 3 had better fit than Model 1 (Model 2: $\chi^2(1) = 3.86$, $p < .05$; Model 3: $\chi^2(1) = 4.01$, $p < .05$). Model 2 and Model 3 were not significantly different in model fit ($p > .05$). Comparing Model 4 with Model 2 and with Model 3, Model 4 showed the better fit; however, the differences were only at borderline significance (Model 2: $\chi^2(1) = 3.70$, $p = .051$; Model 3: $\chi^2(1) = 3.55$, $p = .056$). The difference between Model 4 and Model 5 revealed no significant difference ($\chi^2(2) = 1.75$, $p > .50$).

It seemed that Model 3 was the best model due to its significant better fit than Model 1, and its parsimoniousness compared with Model 4 and Model 5. However, it should be noted that the variable that added to Model 3, i.e. PPVT-R, did not exert a significant effect. On the other hand, Model 2 included a variable that exerted significant effect (i.e. productive phonology), and had the same advantage of parsimoniousness. Therefore, Model 2 might be the best model that account for the developmental trajectory of gap word repetition. Children's performance in gap word repetition was best predicted by their productive phonology, while receptive vocabulary knowledge played a less important role.

We could compare children's performances in nonce word repetition and gap word repetition based on the findings in HLM. The results showed that nonce word repetition performance was associated with both receptive vocabulary knowledge and productive phonology, while the gap word repetition performance was more associated with productive phonology.

The null model on nonce word repetition revealed that there were significant differences in the intercepts and the growth rates between participants in nonce word repetition. The between-participant variation could be accounted for by children's PPVT-R scores and productive phonology. In other words, children's performance in nonce word repetition was correlated with their receptive vocabulary knowledge and phonological productive ability. We also found that productive phonology could account for individual differences in the growth rates of nonce word repetition. On the other hand, the null model on gap word repetition revealed significant difference in children's intercepts, but not in their growth rates. Different from the findings in nonce word repetition, the growth in gap word repetition was more related to children's phonological productive ability.

Chapter 8 General Discussion and Conclusion

The aim of this dissertation was to explore the dynamic relationship among language knowledge (vocabulary size and phonological capacities) and NWR performance. Specifically, we investigated the effects of vocabulary growth and phonological development on the improvement of NWR. A cross-sequential design was conducted, so that we could inspect not only cross-cohort differences, but also the individual variation in the interaction among these capacities. In addition, the longitudinal observations of the age 2 cohort, the age 3 cohort, and the age 4 cohort allowed us to portrait the developmental trajectory of NWR in relation to phonological capacities and vocabulary knowledge from age 2 to age 5.

In the following, we will summarize the main findings of this study in response to the research questions.

8.1 The Contributions of Receptive Vocabulary and Productive Phonology to NWR Performance

We proposed a three-phase working model to account for the interactions among phonological capacities, vocabulary knowledge and NWR along the course of development. It was believed that different factors dominate children's NWR performances in different phases. This hypothesis is born out in the current study. We observed that the main contributor to children's NWR differed across ages, and was associated with children's levels of phonological capacities and vocabulary knowledge. However, our statistical findings also suggested revisions to the original model. The findings and their significance to the original model will be discussed below. A revised model will then be provided at the end of this section.

In our model, children's productive phonology and vocabulary knowledge are supposed to play important roles in the first two phases. It was claimed that at the beginning stage of language development, children's NWR is mainly constrained by their ability to accurately encode the sound forms. This assumption was made based on the idea that children at this age still had little vocabulary size. As a result, their encoding of nonwords would have little lexical support. Due to little involvement of vocabulary knowledge in NWR in this phase, there would not be lexicality effect or wordlikeness effect. In the case of our study, there would be no difference between nonce word repetition and gap word repetition, though both performances tend to be low. However, as children acquire a considerable amount of vocabulary, the intervention of lexical knowledge to NWR would gradually become prominent. In this second phase, we would be able to observe an advantage in children's repetition of nonwords that contain lexical syllables (e.g. nonce word repetition) when compared with their repetition of nonwords that contain non-attested syllables (e.g. gap word repetition). Also, children with a larger vocabulary size would show an advantage in the NWR task. On the other hand, productive phonology plays a minor role in this phase.

Statistical analyses were conducted to examine our hypotheses. The hierarchical linear model analyses on the data revealed that productive phonology and receptive vocabulary accounted for significant cross-individual variation in NWR, including the initial performances and the growth rates. Moreover, regression analyses showed that children's productive phonology and receptive vocabulary at an earlier time usually made significant contributions to their subsequent NWR performance. Therefore, it is verified that productive phonology and vocabulary knowledge contribute to children's NWR development.

Though both productive phonology and vocabulary knowledge played roles in NWR, the extent to which they were associated with NWR differed in different cohorts and in different testing occasions. Based on the original model, it was expected that productive phonology, rather than vocabulary knowledge, would play the dominant role in the NWR of the children in the age 2 cohort. Our hypothesis for the first phase was supported by the finding that children in this cohort showed similar performances in repeating nonce words and gap words. The lexicality effect was not prominent until age 3, when children had acquired a larger amount of vocabulary and might had established a preliminary lexical-phonological network that allowed them to recognize lexical items when they encountered a new word. While children in this age range did not reveal a general lexicality effect, the regression analyses demonstrated that their receptive vocabulary size, which was measured by $\text{MCDI-}_{\text{RECEPTIVE}}$, accounted for their variation in NWR performance. Children who had larger receptive vocabulary knowledge tended to have better NWR, either in nonce word repetition or in gap word repetition, at the subsequent ages. The contribution of productive phonology to NWR was found only in nonce word repetition, with less predictive power than receptive vocabulary.

On the other hand, the original model assumed that vocabulary knowledge would play a crucial role when children grow older. This assumption was supported by the finding that among children above age 3, there was always a prominent and robust lexicality effect in children's NWR performances. Superior performance was found in nonce word repetition than in gap word repetition. Also, a recall advantage was found in children with greater vocabulary size than children with smaller vocabulary size. These findings provided evidence of the mediation of vocabulary knowledge to NWR. However, the regression analyses revealed that the individual variation in NWR performance in this age range was more associated with their productive phonology. Children who had better productive phonology tended to showed better NWR, specifically nonce word repetition.

The findings call for a new perspective on the role of vocabulary and its interaction with productive phonology in novel word encoding in development. The findings suggest that when young children run into a novel sound form, they would make use of the lexical knowledge they have to support their repetition. At this stage, the receptive vocabulary knowledge is still not able to mediate NWR at the lexical level, because it is limited in size, and the phonological forms of the lexical items are mostly underspecified (Ferguson & Farwell, 1975). However, it could mediate NWR at the phonological level by providing a bank of phonological forms that are mostly- or partially-specified, which children could access to for phonological representation that are similar to or matched with the target form. As a result, children with larger receptive vocabulary bank would be at an advantage for the greater resources they have. On the other hand, for children with similar size of vocabulary bank, their NWR could be determined by their phonological capacities, that is, their abilities to process and produce the sound form.

The finding of the association between receptive vocabulary and NWR in the age 2 cohort suggested that children use their existing lexical knowledge, however little it is, to mediate their encoding of novel words, regardless of the lexical status of the novel words. The mediation of productive phonology occurs after receptive vocabulary once children have acquired some vocabulary.

In our study, children began to show more prominent lexicality effect at age 3, and it was robustly present until age 5. Previous studies also showed that this effect is consistently found in older children and adults (Lee, 2005). The emergence of lexicality effect has its significance in phonological and lexical development. Literature on child language development demonstrates that young language learners at the early stage of language development tend to perceive/acquire words as a whole, without much sublexical details (Ferguson & Farwell, 1975). However, as their vocabulary size increases, they begin to recognize similarities across phonological forms and add sublexical information to underlying representation (Metsala & Walley, 1998). This also leads to the formation of a network of lexical-phonological neighbors (i.e. neighborhood), and the more-specified phonemic content at the syllabic level. The growth in lexicon size and the specification of the lexical units raise the chance for children to run into familiar sound combinations in newly-encountered words and also allow them to recognize these forms more efficiently. As a result, their encoding of nonwords would be greatly benefited by the lexical knowledge. In other words, the emergence of lexicality effect may signal the formation of a network of lexical-phonological neighbors, which is important for lexical and phonological processing.

While lexical knowledge was found to mediate NWR in the age 3 cohort and the age 4 cohort, as evidenced by the lexicality effect, the regression analyses showed that variation in NWR was predominantly determined by productive phonology, but not receptive vocabulary. The findings could be explained following the account that we propose. When encoding nonwords, vocabulary knowledge comes in first to help, which leads to the lexicality effect. Children with different levels of vocabulary knowledge showed different NWR performances due to the size of lexical-phonological form bank that they can access to. However, it might have smaller explanatory power to children's NWR variation than productive phonology. This is due to the fact that syllable structures in Mandarin are relatively simple. Because of that, children can have been exposed to most of the possible syllables or constituents in Mandarin when they have acquired a certain amount of vocabulary. Then, what determines their performance in repeating these sound units in the NWR task is their phonological ability.

The results of HLM analyses also demonstrated that phonological production ability could account for significant cross-individual variation in their initial NWR performance, and also the growth rates in nonce word repetition. Therefore, children's phonological representation, as reflected by their phonological output ability, greatly determined their performances in processing nonwords. Children with refined, well-specified phonological representation would have better phonological analysis ability, which allows them to swiftly and accurately construct phonological representations of good quality for the newly-encountered words that can be robustly registered and produced.

Our account appears to assume that there is a close relationship between vocabulary and phonological development. Specifically, we made our argument based on the assumption that phonological representations are shaped through the dynamics of the production-perception loop in the process of learning the forms of lexical items (Munson, Beckman, & Edwards, 2012; Walley, 1993), which is also known as the lexical restructuring hypothesis. This hypothesis was not supported in our age 2 cohort. In this cohort, we found significant correlations between children's onset production at age 2 and age 2.5 to their receptive vocabulary at age 3, which seemed suggest that the developing phonological system affects lexical acquisition to a greater extent than the reverse (Sosa & Stoel-Gammon, 2006; Stoel-Gammon, 2011). However, we observed significant correlations between onset production and receptive vocabulary in older children. Therefore, it is likely that the effect of vocabulary on phonological capacities emerge when considerable amount of lexical items are accumulated. However, this hypothesis may require further exploration.

The nonword repetition task in our study included the nonce word repetition and the gap word repetition. The plots on children's NWR performances and the analysis results of the HLM on children data showed that children manifested growth in both nonce word repetition and gap word repetition. However, the two subtests appeared to have different growth patterns due to the nature of the stimuli. These two subtests both consist of nonsense disyllabic words; however, they differ in the lexical status of the syllables that form the nonwords.

The HLM results suggested that different factors were involved in nonce word repetition and gap word repetition. Children at the same age differed considerably in their initial performances and growth rates in the repetition of nonce words, even when the factor of age was considered. The individual variation in their initial performances was shown to be related to children's productive phonology and receptive vocabulary knowledge, while the variation in growth rates could be associated with productive phonology. With regards to gap word repetition, children showed variation in their initial performance. However, their growth rates were more or less similar. Different from what we found in nonce word repetition, the individual variation in gap word repetition could be accounted for by productive phonology, whereas receptive vocabulary played little role here. The findings appear to confirm our hypothesis regarding the mediating factors involved in nonce word repetition and gap word repetition. While the repetition of nonwords is mediated by the productive phonology, or phonological representation, the repetition of nonce words is additionally supported by lexical knowledge.

In the regression analyses, we particularly observed a role of nonverbal intelligence to gap word repetition in the age 2 cohort and the age 3 cohort. We believe that it represented a cognitive factor in NWR. Nonword repetition is a complex task in which children have to process and store sound stimuli at the same time. When the novel sounds children encounter are familiar to them, it would be easier for them to efficiently and accurately encode and storage the sound for recall. However, if the novel sound forms are strange to them, there will be greater processing load. Then, it would challenge children's ability to allocate attention between processing and storage, which is crucial to the success of a memory task.

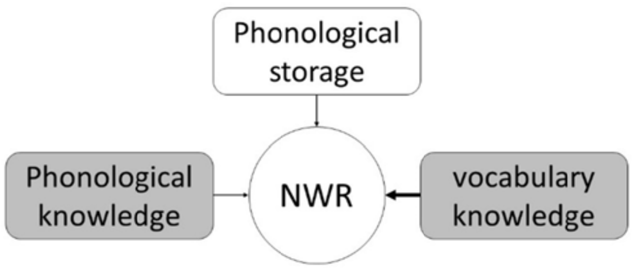
The regression analyses on gap word repetition in the age 4 cohort demonstrated a different patterns from the findings in the other two age cohorts. It was found that the gap word repetition of the age 5 children could not be accounted for by their nonverbal intelligence, productive phonology or vocabulary knowledge. We still do not have an explanation for this. Nonetheless, it is likely that gap word repetition is determined by different factors when the length of gap word stimuli is considered. Further explorations should be conducted to explore this issue.

In the original model, we has hypothesized that there should be a stage in which productive phonology and vocabulary knowledge play no or little role in nonword repetition, and their mediating role has been taken over by storage factors. This stage might occur when children have acquired a great amount of vocabulary and well-developed phonological representations which allow manipulation as abstract discrete units. However, this stage was not observed in this study. This could be due to the fact that we investigated children from age 2 to age 5, who were still developing their vocabulary and phonological system. The emergence of this stage might be observed in much older children or adults.

Based on our findings in this study, we could reformulate the original three-phase model, particularly the first phase and the second phase. First, we observed an unexpected earlier role of lexical knowledge in the youngest cohort. In this age range, children's breadth of vocabulary knowledge appeared to be a crucial factor that determined children's NWR performance. This finding may challenge our proposal of the presence of phase I, in which productive phonology is the dominant factor to NWR. We had expected to observe a mediating role of productive phonology to NWR at such a young age because these children were still at the early stage of vocabulary development. However, it is possible that vocabulary knowledge begins to mediate the processing of novel sound forms when a certain but small amount of vocabulary size is acquired. In fact, when we examined the NWR performances of the children with smaller vocabulary size, it was found that their NWR tended to be more related with productive phonology, rather than vocabulary knowledge. For that reason, we would like to retain phase I for now, since it is likely that children in this cohort is at the transition to phase II, as they showed dramatic vocabulary growth in this age range. Also, it may take a smaller vocabulary for the development of a phonological system in Chinese to support the analysis of nonwords. According to the analysis in the previous section, a smaller vocabulary less than 200 words may do. We may have to look into younger children to see a primary effect of productive phonology to NWR.

Figure 8.1

Modification on the Second Phase of the Three-phase Model



Some changes should be made to the second phase with the finding that productive phonology played an important role to NWR performance in addition to vocabulary knowledge. In phase II, both vocabulary knowledge and phonological knowledge mediate NWR (Figure 8.1). To what extent they contribute to the variance in NWR performance is determined by the increase in vocabulary size and the nature of the stimuli. The effect of lexical knowledge is persistent because when children encounter a novel sound form, they tend to rely on their existing lexical knowledge to process the form. The mediation of lexical knowledge is even more prominent when children acquire a considerable amount of vocabulary, and they demonstrate the lexicality effect in NWR. However, it is likely that when children's receptive vocabulary size reaches a certain level, all the children learn to use their lexicality knowledge to support their encoding of nonwords whenever necessary. Therefore, the individual variation that we observed in children at this time is more related to their productive phonology.

The current study does not allow us to examine the third phase. Future study should be conducted on much older children or adults to verify the hypothesis.

The findings of our study provide some insights to the debate on the nature of NWR, particularly the debate between the phonological storage account and the phonological analysis account. The phonological storage account propose that nonword repetition mainly reflects phonological memory capacity, because its repetition process is heavily influenced by one's phonological memory capacity (Gathercole, 2006). We think this proposal is only partially correct in that it points out that memory capacity could be the fundamental constraint to any human brain processing activities. For example, our study found that children's NWR performance declined dramatically with the increase of the length of the nonword stimuli. As suggested in previous studies, when the length of nonwords increases, the repetition performance is taxed by the complex interaction of several storage factors, including memory decay (Baddeley, 1986), interferences (Nairne et al., 1997), and the allocation of attentional resources (Cowan, 1999; Engle, Kane, & Tuholski, 1999). Also, we observed the involvement of cognitive factors when children encounter gap words, which are more difficult to process.

Even though storage factors greatly affect NWR performance, there are studies pointing out that storage capacity is closely tied to analysis speed, or efficiency (Case, Kurland, & Goldberg, 1982; Dempster, 1981). Developmental progress in processing speed, which could be affected by the familiarity or knowledge of the processed item, would release processing resources for the maintenance of the item and would thus lead to a better performance in the memory task. This is supported by the finding of the lexicality effect in our study. In addition, we found an interaction effect between the length and the lexicality of nonwords in this study. Nonce word repetition, with the mediation of lexical knowledge, showed smaller drop in score than gap word repetition in the three-word lists, suggesting that lexical knowledge mediated the processing and repetition performance in the memory processes.

Contrary to the proposal that phonological storage is the major constraint to nonword repetition and word learning, the phonological analysis account proposes that performances in NWR is constrained by the ability to efficiently process the novel verbal forms into accurate phonological representation (Bowey, 1996, 2001; Metsala, 1999). This proposal has received support from the findings that young children's phonological analysis, as measured by phonological awareness (Metsala, 1999) or output production (Li & Cheung, 2014), accounts for major proportion of variance when the effect of short-term memory is controlled. The finding of our study was in support of this account by demonstrating the role of productive phonology and its pervasive influence on NWR from age 2.5 to age 5, though the assumption that vocabulary increase is the driving force of phonological development (Fowler, 1991; Metsala, 1999) was not fully supported in this study. However, different from this account, our study demonstrated that vocabulary knowledge also plays a crucial role to NWR.

8.2 The Predictability of NWR to Expressive Vocabulary Development

It seems to be a well-established fact that NWR performance is associated with vocabulary knowledge (see reviews in Gathercole, 2006). Several studies have tried to disentangle the causal relationship between vocabulary knowledge and NWR, and have found that NWR predicts children's vocabulary knowledge in the earlier years. However, this proposal is not without debate. For example, with data derived from a three-year longitudinal study, Melby-Lervåg et al. (2012) assert that phonological working memory is not associated with vocabulary development.

Our study found that NWR, either nonce word repetition or gap word repetition, could predict children's subsequent expressive vocabulary knowledge, but not their subsequent receptive vocabulary knowledge. This finding was replicated in all the three age cohorts in our study.

The relationship between NWR and expressive vocabulary has been found in several studies. For example, Hoff, Core and Bridge (2008) found that NWR performances were significantly correlated with the expressive vocabulary percentile in children at age 2 ($r = .53 \sim .72$, $p < .05$). Moreover, Stokes and Klee (2009) found that NWR was the unique predictor to the expressive vocabulary knowledge of children aged 24-30 months. While these studies demonstrated the relationship only in children at around age 2, we extended the relationship to older ages.

Our findings also contradicted with other studies which have found a causal relationship between NWR and receptive vocabulary. For example, Gathercole, Willis, Emslie, and Baddeley (1992) showed that NWR performance at age 4 could predict receptive vocabulary knowledge at age 5, while our study showed that NWR at age 4 did not account for significant variance in the subsequent receptive vocabulary knowledge when the variables of age and nonverbal intelligence were controlled.

One possible reason to this incongruent finding is that different standardize tests were used to assess children's receptive vocabulary. For example, while Gathercole et al. (1992) used the short form of the British Picture Vocabulary Scale (Dunn & Dunn, 1982), we used the Peabody Picture Vocabulary Test-Revised in our study. However, if NWR is robustly associated with receptive vocabulary knowledge, the correlation should be observed even when a different assessment for the same construct is used.

While we found that NWR failed to predict children's subsequent receptive vocabulary development, we discovered that children's receptive vocabulary could predict the subsequent NWR performances. Therefore we saw the interactive relationships among receptive vocabulary, expressive vocabulary, productive phonology and NWR. That is, the growth in receptive vocabulary and productive phonology would lead to better performance in the repetition of nonwords. Then, the performances in NWR could further support the subsequent development of expressive vocabulary. Compared with receptive vocabulary, expressive vocabulary requires the complete form of a word, not only in the phonological aspect but also the lexical aspect. Receptive vocabulary could serve as the foundation for the development of expressive vocabulary, because it provides at least partial information of a lexicon. Productive phonology also supports the development expressive vocabulary, because it allows the child to construct a well-specified phonological representation of a word, and also allows the child to produce the word accurately. While NWR taps children's receptive vocabulary and productive phonology, it can serve to predict the expressive vocabulary development.

8.3 Conclusion

NWR has been found to be a powerful indicator to not only children's language development, but also to children with language disorders. Looking into the nature of this task would shed light on the mechanisms that support language development. The cross-sequential study on the relationships among vocabulary knowledge, phonological capacities and NWR allowed us not only to examine and compare children from different age cohorts, but also to delineate the developmental trajectory of NWR from age 2 to age 5. Though the number of children recruited in each cohort was relatively small compared with other large-scale studies, the longitudinal observations on children's developments in phonology, vocabulary and NWR have generated several valuable data and findings. First of all, the study provided a profile of Mandarin-speaking children's phonological, vocabulary and NWR development from as young as age 2 to age 5. The data demonstrate the average growth trend, and individual variation in their onset abilities and growth rates. Secondly, with the HLM approach, we were able to examine the strength of the participating factors, such as receptive vocabulary and productive phonology to NWR, and the loci of their effects. Particularly, we found that while the repetition of nonwords relies on the productive phonology, the repetition of nonce words is additionally supported by receptive lexical knowledge. Given the close relationship between NWR and language development reported in the literature, the distinction of the nonce words and the gap words have the potential to detect the locus of children's problem when they showed deteriorating performance in NWR and impairments in language. Particularly, we propose that the emergence of the lexicality effect signal the formation of a network of lexical-phonological neighbors, which is important for lexical and phonological processing. Our study also revealed that children's NWR performances could predict their subsequent development in expressive vocabulary. Therefore, the association between NWR and vocabulary knowledge, which has been well-established in the literature on Indo-European languages, could be extended to Mandarin-speaking children.

This study examined how NWR was associated with vocabulary knowledge and phonological capacities, which are the established long-term linguistic knowledge. However, it is of interest whether NWR could predict children's learning of novel words. Moreover, NWR is also found to be related to children's syntactic development and reading development in studies of Indo-European languages. Yet, whether this association could be replicated in Mandarin, a typologically different language, remains unexplored. It is worth exploring how NWR is related to other language aspects in

Mandarin-speaking children.

This study presented typically-developing children's profile of NWR development, which could serve as the baseline for the identification of children who perform far below the average NWR performance and show a risk for language impairment. Regarding the utility of NWR for clinical purpose, several questions should be explored. For example, how do children with different types of language impairments or language difficulties perform in the NWR task? Recently, some studies have been undertaken to examine the NWR performances of Mandarin-speaking children with language impairments, such as children with articulation disorders (Hsieh, 2007), children with stuttering (Chen, 2011), and children with specific language impairment (Chen, 2012; Zhang, 2011). These studies provide profiles of the NWR performances in different clinical groups, and revealed that children in these clinical groups (except for the children with articulation disorders) had poorer NWR performance than their typically-developing peers. However, it is very likely that these clinical groups showed similarly poor NWR performance for different underlying factors (Riches et al., 2011). Therefore, more analyses could be done to examine children's error patterns in the NWR task, which might provide insights to the underlying deficits of these language impairments. Also, all these studies used nonce words as the stimuli in the NWR task. Based on the findings of our study, the repetition of nonce words could be mediated by lexical knowledge in addition to phonological capacities. In other words, the failure in repeating nonce words could be attributed to either processing problems at the lexical level, or processing problems at the phonological level. However, if we could also examine these children's performance in dealing with gap words, we might be able to identify the specific locus of impairments by comparing their performance in the nonce word repetition and gap word repetition. Last but not the least, we should also examine whether NWR can serve as a reliable indicator to language impairment in

Mandarin-speaking children. It has been well-established that NWR could serve the diagnostic purpose to identify children with language impairments in Italian, Spanish, and English (D'odorico et al., 2007; Girbau & Schwartz, 2007; Stokes & Klee, 2009). However, the NWR study on language-impaired children speaking Cantonese, a language that is structurally more similar to Mandarin, demonstrated contradicting results (Stokes et al., 2006). With regards to the study on Mandarin-speaking children with specific language impairment, it is now known that they had poor performance in NWR. Nevertheless, the data on the diagnostic accuracy of NWR to Mandarin-speaking SLI is still lacking. Further exploration into this issue could specify the clinical value of NWR in Mandarin. To sum up, future studies on NWR and language impairment groups could add to our understanding of the language learning mechanism. More importantly, they could help to pinpoint the processing problems that these children may suffer, which further enable the therapists to construct more efficient and effective intervention strategies for these clinical groups.

Reference

Abel, A. D, & Schuele, M. (2014). The influence of two cognitive-linguistic variables on incidental word learning in 5-year-olds. *Journal of Psycholinguistic Research, 43,* 477-463.

Adams, A.-M., & Gathercole, S.E. (1995). Phonological working memory and speech production in preschool children. *Journal of Speech and Hearing Research, 38*, 403-414.

Adams, A-M., & Gathercole, S. E. (1996). Phonological working memory and spoken language development in young children. *The Quarterly Journal of Experimental Psychology, 49A*, 216-233.

Anthony, J. L., Williams, J. M., Aghara, R. G., Dunkelberger, M., Novak, B., & Mukherjee, A. D. (2010). Assessment of individual differences in phonological representation. *Reading and Writing, 23*, 969-994.

Archibald, L. M. D., & Gathercole, S. E. (2006). Nonword repetition: A comparison of tests. *Journal of Speech, Language, and Hearing Research, 49*, 970-983.

Archibald, L. M. D., & Gathercole, S. E. (2007). Nonword repetition and serial recall: Equivalent measures of verbal short-term memory? *Applied Psycholinguistics, 28*, 587-606.

Baddeley, A. D. (1986). *Working memory.* Oxford: Oxford University Press.

Baddeley, A. D. (1993). Short-term phonological memory and long-term learning: A single case study. *European Journal of Cognitive Psychology, 5*, 129-148.

Baddeley, A. D., Gathercole, S. E., & Papagno, C. (1998). The phonological loop as a language learning device. *Psychological Review, 105*, 158-173.

Baddeley, A. D., & Hitch, G. J. (1974). Working memory. In G. Bower (Ed.), *The psychology of learning and motivation* (Vol. 8, pp. 47-90). New York: Academic Press.

Bailey, A., Palferman, S., Heavey, L., & Le Couteur, A. (1998). Autism: The phenotype in relatives. *Journal of Autism and Developmental Disorders, 28*, 369-389.

Barrouillet, P., Gavens, N., Vergauwe, E., Gaillard, V., & Camos, V. (2009). Working memory span development: A time-based resource-sharing model account. *Developmental Psychology, 45,* 477-490.

Barrouillet, P., Bernardin, S., & Camons, V. (2004). Time constraints and resource sharing in adults' working memory spans. *Journal of Experimental Psychology-Genera, 133*, 83-100.

Bell, R. Q. (1953). Convergence: An accelerated longitudinal approach. *Child Development, 24*, 145-152.

Best, C. T. (1984). Discovering messages in the medium: Speech and the prelinguistic infant. In H. E. Fitzgerald, B. Lester, & M. Yogman (Eds.), *Advances in pediatric psychology* (pp. 97-145). New York: Plenum.

Bishop, D. V., North, T., & Donlan, C. (1996). Nonword repetition as a behavioral marker for inherited language impairment: Evidence from a twin study. *Journal of Child Psychology and Psychiatry, and Allied Disciplines, 37,* 391-403.

Bowey, J. A. (1996). On the association between phonological memory and receptive vocabulary in five-year-olds. *Journal of Experimental Child Psychology, 63*, 44-78.

Bowey, J. A. (2001). Nonword repetition and young children's receptive vocabulary: A longitudinal study. *Applied Psycholinguistics, 22*, 441-469.

Case, R., Kurland, D. M., & Goldberg, J. (1982). Operational efficiency and the growth of short-term memory span. *Journal of Experimental Child Psychology, 33*, 386-404.

Chang, Cheng-feng, & Chung, Yu-Mei. (1986). The Chinese version of Preschool Language Scale and its related studies. *Bulletin of Special Education, 2,* 37-52.

Chen, Li-mei. (2005). Development of vocalic production in a Mandarin-learning infant: A longitudinal case study. *Providence Forum: Language and Humanities, 1,* 169-192.

Chen, W. (2011). *Comparing the phonological and articulation abilities between stuttering and nonstuttering Mandarin-speaking children* (Unpublished master's thesis). National Pingtung University, Pingtung, Taiwan.

Chen, Y. (2012). *Novel word learning and phonological processing abilities in preschool children with specific language impairment* (Unpublished doctoral dissertation). National Taiwan Normal University, Taipei, Taiwan.

Cheung, H. (1998). The application of MLU in Chinese. *Journal of Speech-Language-Hearing Association of ROC, 13*, 36-48.

Cheung, H. (2000). Three to four-year-old children's perception and production of Mandarin consonants. *Language and Linguistics*, *1*, 19-38.

Chiat, S., & Roy, P. (2007). The preschool repetition test: An evaluation of performance in typically developing and clinically referred children. *Journal of Speech, Language, and Hearing Research, 50,* 429-443.

Cho, Shih-Chieh. (2008). *The phonological development of 3 to 6 year-old preschool children in Taiwan* (Unpublished master's thesis). National Taipei College of Nursing, Taipei, Taiwan.

Clark, E. V. (2003). *First language acquisition*. Cambridge: Cambridge University Press.

Coady, J. A, & Evans, J. L. (2008). Uses and interpretations of non-word repetition tasks in children with and without specific language impairments (SLI). *International Journal of Language and Communication Disorders, 43*, 1-40.

Coady, J, Evans, J. L., & Kluender, K. R. (2010). Role of phonotactic frequency in nonword repetition by children with specific language impairments. *International*

Journal of Language and Communication Disorders, 4, 494-509.

Conti-Ramsden, G. (2003). Processing and linguistic markers in young children with specific language impairment (SLI). *Journal of Speech, Language, and Hearing Research, 46*, 1029-1037.

Cowan, N. (1999). An embedded-processes model of working memory. In A. Miyake & P. Shah (Eds.), *Models of working memory: Mechanisms of active maintenance and executive control* (pp. 62-101). Cambridge, NY: Cambridge University Press.

D'odorico, L., Assanelli, A., Franco, F., & Jacob, V. (2007). A follow-up study on Italian late talkers: Development of language, short-term memory, phonological awareness, impulsiveness, and attention. *Applied Psycholinguistics, 28*, 157-169.

Dollaghan, C. A., Biber, M. E., & Campbell, T. F. (1995). Lexical influences on nonword repetition. *Applied Psycholinguistics, 16*, 211-222.

Dollaghan, C. A., & Campbell, T. F. (1998). Nonword repetition and child language impairment. *Journal of Speech, Language, and Hearing Research, 41*, 1136-1146.

Edwards, J., Beckman, M. E., & Munson, B. (2004). The interaction between vocabulary size and phonotactic probability effects on children's production accuracy and fluency in novel word repetition. *Journal of Speech, Language, and Hearing Research, 47*, 421-436.

Edwards, J., & Lahey, M. (1998). Nonword repetitions of children with specific language impairment: Exploration of some explanations for their inaccuracies. *Applied Psycholinguistics, 19*, 279-309.

Edwards, J., Munson, B., & Beckman, M. E. (2011). Lexicon-phonology relationships and dynamics of early language development. *Journal of Child Language, 38*, 35-40.

Engle, R. W., Kane, M. J., & Tuholski, S. W. (1999). Individual differences in working memory capacity and what they can tell us about controlled attention, general fluid intelligence, and functions of the prefrontal cortex. In A. Miyake & P. Shah (Eds.), *Models of working memory: Mechanisms of active maintenance and executive control* (pp. 102-134). Cambridge, NY: Cambridge University Press.

Estes, G., Evans, J. L., & Else-Quest, N. M. (2007). Differences in the nonword repetition performance of children with and without specific language impairment: A meta-analysis. *Journal of Speech, Language, and Hearing Research, 50*, 177-195.

Ferguson, C. A., & Farwell, C. B. (1975). Words and sounds in early language acquisition. *Language, 51*, 419-439.

Fowler, A. (1991). How early phonological development might set the stage for phoneme awareness. In S. Brady & D. Shankweiler (Eds.), *Phonological processes in literacy: A tribute to I. Y. Liberman* (pp. 97-117). Hillsdale, NJ: Erlbaum.

Foy, J. G., & Mann, V. (2001). Does strength of phonological representations predict phonological awareness in preschool children? *Applied Psycholinguistics, 22*, 301-325.

Frisch, S. A., Large, N. R., & Pisoni, D. B. (2000). Perception of wordlikeness: Effects of segment probability and length on the processing of nonwords. *Journal of Memory and Language, 42,* 481-496.

Gallon, N., Harris, J., & van der Lely, H. K. J. (2007). Non-word repetition: An investigation of phonological complexity in children with Grammatical SLI. *Clinical Linguistics and Phonetics, 21*, 435-455.

Gathercole, S. E. (2006). Nonword repetition and word learning: The nature of the relationship. *Applied Psycholinguistics, 27*, 513-543.

Gathercole, S. E., & Adams, A-M. (1994). Children's phonological working memory: Contributions of long-term knowledge and rehearsal. *Journal of Memory and Language, 33*, 672-688.

Gathercole, S. E., Adams, A-M., & Hitch, G. J. (1994). Do young children rehearse? An individual difference analysis. *Memory and Cognition, 22*, 201-207.

Gathercole, S. E., & Baddeley, A. D. (1989). Evaluation of the role of phonological STM in the development of vocabulary in children: A longitudinal study. *Journal of Memory and Language, 28*, 200-213.

Gathercole, S. E., & Baddeley, A. D. (1990a). The role of phonological memory in vocabulary acquisition: A study of young children learning new names. *British Journal of Psychology, 81*, 439-454.

Gathercole, S. E., & Baddeley, A. D. (1990b). Phonological memory deficits in language disordered children: Is there a causal connection? *Journal of Memory and Language, 29*, 336-360.

Gathercole, S. E., Frankish, C. R., Pickering, S. J., & Peaker, S. (1999). Phonotactic influences on short-term memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 25*, 84-95.

Gathercole, S. E., Hitch, G. J., Service, E., & Martin, A. J. (1997). Phonological short-term memory and new word learning in children. *Developmental Psychology, 33,* 966-979.

Gathercole, S. E., Willis, C., & Baddeley, A. D. (1991). Differentiating phonological memory and awareness of rhyme: Reading and vocabulary development in children. *British Journal of Psychology, 82*, 387-406.

Gathercole, S. E., Willis, C., Emslie, H., & Baddeley, A. D. (1991). The influences of number of syllables and wordlikeness on children's repetition of nonwords. *Applied Psycholinguistics, 12*, 349-367.

Gathercole, S. E., Willis, C. S., Emslie, H., & Baddeley, A. D. (1992). Phonological memory and vocabulary development during the early school years: A longitudinal study. *Developmental psychology*, *28*(5), 887.

Gathercole, S. E., Willis, C., Baddeley, A. D, & Emslie, H. (1994). The children's test of nonword repetition: A test of phonological working memory. *Memory, 2*, 103-127.

Girbau, D., & Schwartz, R. G. (2007). Non-word repetition in Spanish-speaking children with specific language impairment (SLI). *International Journal of Language and Communication Disorders, 42*, 59-75.

Guiberson, M., & Rodriguez, B. L. (2013). Classification accuracy of nonword repetition when used with preschool-age Spanish-speaking children. *Language, Speech, and Hearing Services in Schools, 44*, 121-132.

Gupta, P. (2009). A computational model of nonword repetition, immediate serial recall, and nonword learning. In A. Thorn & M. Page (Eds.), *Interactions between short-term and long-term memory in the verbal domain* (pp. 108-135). Hove & NY: Psychology Press.

Henry, L. A. (1991). Development of auditory memory span: The role of rehearsal. *British Journal of Developmental Psychology, 9*, 493-511.

Hoff, E., Core, C., & Bridges, K. (2008). Non-word repetition assesses phonological memory and is related to vocabulary development in 20- to 24-month-olds. *Journal of Child Language, 35*, 903-916.

Hox, J. (2010). *Multilevel analysis, techniques and applications.* Mahwah, N. J.: Lawrence Erlbaum Associates.

Hsieh, Y-S. (2007). *Nonword repetition in children with phonological disorders* (Unpublished master's thesis). National Taipei College of Nursing, Taipei, Taiwan.

Hsu, H. (2005). *The association of phonological disorders and syntactic disorders: A study of Mandarin-speaking children* (Unpublished master's thesis). National Taiwan University, Taipei, Taiwan.

Hsu, Joseph, H. (1987). *A study of the stages of development and acquisition of Mandarin Chinese by children in Taiwan.* Taipei: Crane.

Hu, C-F. (2004). The development of phonological representations among Chinese-speaking children. *Taiwan Journal of Linguistics, 2*, 103-130.

Hu, C. F., & Catts, H. W. (1998). The role of phonological processing in early reading ability: What we can learn from Chinese. *Scientific Studies of Reading, 2*, 55-79.

Huang, R., Jian, H., Zhu, L., & Lu. L. (2011). *Receptive and expressive vocabulary test.* Taipei: Psychological Publishing.

Hulme, C., Maughan, S., & Brown, G. D. A. (1991). Memory for familiar and unfamiliar words: Evidence for a long-term memory contribution to short-term memory span. *Journal of Memory and Language, 30*, 685-701.

Hulme, C., Thomson, N., Muir, C., & Lawrence, A. (1984). Speech rate and the development of short-term memory span. *Journal of Experimental Child Psychology, 38*, 241-253.

Ingram. D. (1974). The relationship between comprehension and production. In R. Schiefelbusch & L. Lloyd (Eds.), *Language perspective: Acquisition, retardation, and intervention* (pp. 313-334). Baltimore, Md.: University Park Press.

Ingram, D. (1989). *First language acquisition: Method, description and explanation.* NY: Cambridge University Press.

Jeng, Heng-hsiung. (1985). A developmentalist view of child phonology. *Studies in Language and Literature, 1,* 1-29.

Jones, G, & Witherstone, H. L. (2010). Lexical and sublexical knowledge influences the encoding, storage, and articulation of nonwords. *Memory and Cognition, 39,* 588-599.

Jusczyk, P. W., Luce, P. A., & Charles-Luce, J. (1994). Infants' sensitivity to phonotactic patterns in the native language. *Journal of Memory and Language, 33*, 630-645.

Kovács, G., & Racsmány, M. (2008). Handling L2 input in phonological STM: The effect of non-L1 phonetic segments and non-L1 phonotactics on nonword repetition. *Language Learning, 58*, 597-624.

Lee, J. C. (2005). *Nonword repetition and phonological processing in Mandarin-speaking preschool children* (Unpublished master's thesis). National Taiwan University, Taipei, Taiwan.

Li, C. N., & Thompson, S. A. (1977). The acquisition of tone in Mandarin-speaking children. *Journal of Child Language, 4,* 185-199.

Li, N., & Cheung, H. (2014). Phonological analysis facilitates nonword repetition in young children. *Chinese Journal of Psychology, 56*, 359-378.

Lin, Grace Bao-Guey, & Lin, Mei-hsiu. (1994). A study on the development of language abilities in preschool children in Taiwan, R.O.C. *Bulletin of Special Education, 10,* 259-281.

Lindgren, K. A., Folstein, S., Tomblin, J. B., & Tager-Flusberg, H. (2009). Language and reading abilities of children with autism spectrum disorders and specific language impairment and their first-degree relatives. *Autism Research, 2*, 22-38.

Liu, Y-C. (2007). *Articulation development in final nasal consonants in 2- to 5-year-old Mandarin-speaking children* (Unpublished master's thesis). Taipei Municipal University of Education, Taipei, Taiwan.. c

Liu, H-M., & Tsao, F-M. (2010). *Mandarin-Chinese Communicative Development Inventory (Taiwan).* Taipei: Psychological Publishing.

Lu, L., & Liu, H. X. (1994). *PPVT-revised, Chinese version*. Taipei: Psychology.

Mann, V. A., & Foy, J. G. (2007). Speech development patterns and phonological awareness in preschool children. *Annals of Dyslexia, 57*, 51-74.

Marchman, V., & Bates, E. (1994). Continuity in lexical and morphological development: A test of the critical mass hypothesis. *Journal of Child Language, 21*, 339-366.

Maridaki-Kassotaki, Katerina. (2002). The relation between phonological memory skills and reading ability in Greek-speaking children: Can training of phonological memory contribute to reading development? *European Journal of Psychology of Education, 17,* 63-73.

Masoura, E. V., & Gathercole, S. E. (2005). Phonological short-term memory skills and new word learning in young Greek children. *Memory, 13*, 422-429

Melby-Lervåg, M., Lervåg, A., Lyster, S. H., Klem, M., Hagtvet, B. & Hulme, C. (2012). Nonword-repetition ability does not appear to be a causal influence on children's vocabulary development. *Psychological Science, 23*, 1092-1098.

Melka, F. (1997). Receptive vs. productive aspects of vocabulary. In N. Schmitt & M. McCarthy (Eds.), *Vocabulary description, acquisition and pedagogy* (pp. 84-102). Cambridge: Cambridge University Press.

Messer, M. H., Leseman, P. P. M., Boom, J., & Mayo, A. Y. (2010). Phonotactic probability effect in nonword recall and its relationship with vocabulary in monolingual and bilingual preschoolers. *Journal of Experimental Child Psychology, 105*, 306-323.

Metsala, J. (1999). Young children's phonological awareness and nonword repetition as a function of vocabulary development. *Journal of Education Psychology, 91*, 3-19.

Metsala, J., & Chisholm, G. M. (2010). The influence of lexical status and neighborhood density on children's nonword repetition. *Applied Psycholinguistics, 31*, 489-506.

Metsala, J., & Walley, A. C. (1998). Spoken vocabulary growth and the segmental restructuring of lexical representations: Precursors to phonemic awareness and early reading ability. In J. L. Metsala & L. C. Ehri (Eds.), *Word recognition in beginning literacy* (pp. 89-120). New York: Erlbaum.

Miyazaki, Y., & Raudenbush, S. W. (2000). Tests for linkage of multiple cohorts in an accelerated longitudinal design. *Psychological Methods, 5*, 44-63.

Morra, S., & Camba, R. (2009). Vocabulary learning in primary school children: Working memory and long-term memory components. *Journal of Experimental Child Psychology, 104*, 156-178.

Munson, B. (2006). Nonword repetition and levels of abstraction in phonological knowledge. *Applied Psycholinguistics, 27*, 577-581.

Munson, B., Beckman, M., & Edwards, J. (2012). Abstraction and specificity in early lexical representations: Climbing the ladder of abstraction. In A. C. Cohn, C. Fougeron & M. Huffman (Eds.), *The Oxford handbook of laboratory phonology* (pp. 288-309). Oxford: Oxford University Press.

Munson, B., Kurtz, B.A., & Windsor, J. (2005). The Influence of Vocabulary Size, Phonotactic Probability, and Wordlikeness on Nonword Repetitions of Children with and without Language Impairments. *Journal of Speech, Language, and Hearing Research, 48*, 1033-1047.

Nairne, J. S., Neath, I., & Serra, M. (1997). Proactive interference plays a role in the word-length effect. *Psychonomic Bulletin and Review, 4*, 541-545.

Nazzi, T. (2005). Use of phonetic specificity during the acquisition of new words: Differences between consonants and vowels. *Cognition, 98*, 13-30.

Nespor, M., Peña, M., & Mehler, J. (2003). On the different roles of vowels and consonants in speech processing and language acquisition. *Lingue e Linguaggio*, *2*, 221-247.

Ramachandra, V., Hewitt, L. E., & Brackenbury, T. (2011). The relationship between phonological memory, phonological sensitivity, and incidental word learning. *Journal of Psycholinguistic Research, 40*, 93-109.

Riches, N. G., Loucas, T., Baird, G., Charman, T., & Simonoff, E. (2011). Non-word repetition in adolescents with specific language impairment and autism plus language impairments: A qualitative analysis. *Journal of Communication Disorders, 44*, 23-36.

Roid, G. H., & Miller, L. J. (1997). *Leiter International Performance Scale-Revised (Leiter-R) manual.* Wood Dale, IL: Stoelting.

Roy, P., & Chiat, S. (2004). A prosodically controlled word and nonword repetition task for 2- to 4-year-olds: Evidence from typically developing children. *Journal of Speech, Language, and Hearing Research, 47*, 223-234.

Service, E., Maury, S., & Luotoniemi, E. (2007). Individual differences in phonological learning and verbal STM span. *Memory and Cognition, 35*, 1122-1135.

Shiu, Huei-shiou. (1990). *The phonological acquisition by Mandarin-speaking children: A longitudinal case study on children from 9 months through three years old* (Unpublished master's thesis). National Taiwan Normal University, Taipei, Taiwan.

Shriberg, L. D., Lohmeier, H. L., Campbell, T. F., Dollaghan, C. A., Green, J. R., & Moore, C. A. (2009). A nonword repetition task for speakers with misarticulations: The syllable repetition task (SRT). *Journal of Speech, Language, and Hearing Research, 52*, 1189-1212.

Smith, A., Goffman, L., Sasisekaran, J., Weber-Fox, C. (2012). Language and motor abilities of preschool children who stutter: Evidence from behavioral and kinematic indices of nonword repetition performance. *Journal of Fluency Disorders, 37*, 344-358.

Smith, B. L., McGregor, K. K., & DeMille, D. (2006). Phonological development in lexically precocious 2-year-olds. *Applied psycholinguistics, 27*, 355-375.

Snowling, M., Chiat, S., & Hulme, C. (1991). Words, nonwords, and phonological processes: Some comments on Gathercole, Willis, Emslie, and Baddeley. *Applied Psycholinguistics, 12,* 369-373.

Sosa, A. V., & Stoel-Gammon, C. (2006). Patterns of intra-word phonological variability during the second year of life. *Journal of Child Language, 33*, 31-50.

Stoel-Gammon, C. (2011). Relationships between lexical and phonological development in young children. *Journal of Child Language, 38*, 1-34.

Stokes, S. F., & Klee, T. (2009). The diagnostic accuracy of a new test of early nonword repetition differentiating late talking and typically developing children. *Journal of Speech, Language, and Hearing Research, 52*, 872-882.

Stokes, S. F., Wong, A. M-Y., Fletcher, P., & Leonard, L. B. (2006). Nonword repetition and sentence repetition as clinical markers of specific language impairment: The case of Cantonese. *Journal of Speech, Language, and Hearing Research, 49*, 219-236.

Storkel, H. L. (2001). Learning new words: Phonotactic probability in language development. *Journal of Speech, Language, and Hearing Research, 44,* 1321-1337.

Studdert-Kennedy, M. (1987). The phoneme as a perceptumotor structure. In A. Allport, D. MacKay, W. Prinz, & E. Scheerer (Eds.), *Language perception and production* (pp. 67-84). London: Academic Press.

Studdert-Kennedy, M. (2000). Imitation and the emergence of segments. *Phonetica*, *57*, 275-283.

Studdert-Kennedy, M., & Goodell, E. W. (1995). Gestures, features and segments in early child speech. In B. de Gelder, & J. Morais (Eds.), *Speech and reading: A comparative approach* (pp. 65-88). Hove, England: Erlbaum (UK), Taylor & Francis.

Taylor, L. J., Maybery, M. T., Grayndler, L., & Whitehouse, A. J. O. (2014). Evidence for distinct cognitive profiles in autism spectrum disorders and specific language impairment. *Journal of Autism and Developmental Disorders, 44*, 19-30.

Towse, J. N., & Hitch, G. J. (1995). Is there a relationship between task demand and storage space in tests of working memory capacity? *The Quarterly Journal of Experimental Psychology, 48*, 108-124.

Treiman, R., & Breaux, A. M. (1982). Common phoneme and overall similarity relations among spoken syllables: Their use by children and adults. *Journal of Psycholinguistic Research, 11*, 569-598.

Tseng, Shu-Chuan. (2013). Lexical coverage in Taiwan Mandarin conversation. *International Journal for Computational Linguistics and Chinese Language Processing, 18*, 1-18.

Vitevitch, M. S., & Luce, P. A. (1998). When words compete: Levels of processing in spoken word perception. *Psychological Science, 9*, 325-329.

Vitevitch, M. S. & Luce, P. A. (2005). Increases in phonotactic probability facilitate spoken nonword repetition. *Journal of Memory and Language*, 52, 193-204.

Walley, A. C. (1993). The role of vocabulary growth in children's spoken word recognition and segmentalization ability. *Developmental Review, 13*, 286-350.

Wang, F. N-M., Huang, Helen, Chen, Ching-Wen, Phillips, Penny. (1986). Articulation development in Mandarin-speaking preschool children ages 3-6 in Taiwan.

Chang Gung Medical Journal, 9, 68-75.

Weill, France. (2011). *The role of verbal working memory in new word learning in toddlers 24 to 30 months old* (Unpublished doctoral dissertation). Seton Hall University, New Jersey, US.

Windsor, J., Kohnert, K., Lobitz, K. F., & Pham, G. T. (2010). Cross-language nonword repetition by bilingual and monolingual children. *American Journal of Speech-language Pathology, 19,* 298-301.

Yang, Shu-wen. (2002). *The role of phonological capacity in children's vocabulary learning process* (Unpublished master's thesis). National Taiwan University, Taipei, Taiwan.

Yuen, P. J. (2009). *The development of phonological memory and fast-mapping in preschool children* (Unpublished master's thesis). The University of Hong Kong, Hong Kong.

Zhang, X. (2011). *The phonological processing skills in predicting word recognition in children with specific language impairment* (Unpublished master's thesis). National Taipei University of Nursing and Health Science, Taipei, Taiwan.

Zhu, H. (2006). The normal and disordered phonology of Putonghua (modern standard Chinese)-speaking children. In H. Zhu & B. Dodd (Eds.), *Phonological development and disorders in children: A multilingual perspective* (pp. 81-108). Clevedon: Multilingual Matters.

Zhu, H., & Dodd, B. (2000). The phonological acquisition of Putonghua (modern standard Chinese). *Journal of Child Language, 27*, 3-42.

錡寶香（2009）。兒童語言與溝通發展。台北市：心理出版社。

Appendix A

A Sample of the Preschool Child Developmental Checklist

臺北市學前兒童發展檢核表

2歲（1歲11個月16天~2歲5個月15天）

檢查單位：＿＿＿＿＿＿＿＿ 單位電話：＿＿＿＿＿＿＿＿
填表人姓名：＿＿＿＿＿＿ 身分：□醫療人員 □老師 □社政人員 □家長 □其他＿＿＿＿＿＿
原始國籍：父：□本國籍□大陸□泰國□印尼□越南□柬埔寨□緬甸□其他：請註明：＿＿＿＿
母：□本國籍□大陸□泰國□印尼□越南□柬埔寨□緬甸□其他：請註明：＿＿＿＿

兒童基本資料

兒童姓名：＿＿＿＿＿＿＿＿ 性別：□男 □女 檢核日期：＿＿年＿＿月＿＿日
身分證字號：□ □ □ □ □ □ □ □ □ □ 出生日期：＿＿年＿＿月＿＿日 （早產）預產日期：＿＿年＿＿月＿＿日
實足年齡：＿＿歲＿＿個月＿＿天（請務必填寫）
戶籍住址：臺北市＿＿區＿＿里＿＿鄰＿＿＿＿＿＿
聯絡住址：＿＿＿＿＿＿＿＿ 電話：（日）＿＿＿＿＿＿（夜）＿＿＿＿＿＿

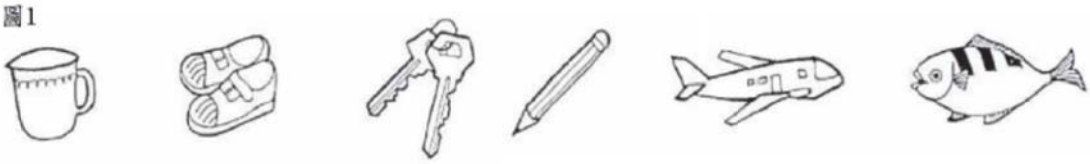
發展遲緩高危險因子

1.①□早產（懷孕期未滿36週）②□出生體重未滿2500公克③□以上皆無
2.先天性異常： ①□染色體異常（如唐氏症、特納氏症等） ②□頭顱顏面異常（如唇顎裂、外耳異常等） ③□先天性新陳代謝異常（如苯酮尿症、甲狀腺功能低下等） ④□水腦脊柱裂 ⑤□頭骨提早密合 ⑥□先天性心臟病 ⑦□手足缺損畸形 ⑧□其他＿＿＿＿＿＿ ⑨□以上皆無
3.產前、產程或產後問題： ①□孕期前3個月感染德國麻疹 ②□母親妊娠期有不正常出血安胎、糖尿病、妊娠毒血、梅毒、酗酒、抽煙 ③□產程有胎心音下降，吸入胎便、呼吸窘迫、窒息缺氧須急救、住保溫箱＿＿＿天 ④□Apgar分數過低：5分鐘後<7(或小於等於6)；請參考母子手冊的出生紀錄 ⑤□出生後有痙攣、無呼吸、反覆嘔吐、低體溫或哺乳不良等 ⑥□重度黃疸需換血者 ⑦□其他＿＿＿＿＿＿ ⑧□以上皆無
4.腦部疾病或受傷： ①□水腦 ②□出血或缺氧 ③□腦部感染 ④□癲癇 ⑤□腦瘤 ⑥□其他＿＿＿＿＿＿ ⑦□以上皆無
5.家族史或環境因素： ①□近親有視聽障礙、智能不足、精神疾病 ②□社經不利狀況 ③□孤兒或受虐兒 ④□以上皆無

發展里程檢核

兒童符合該項目描述的現象圈選「是」，若不符合或沒有該項目描述的現象圈選「否」。註記（實作）的題項表示附有圖形，請實地測試，再記錄兒童反應。		
★ 1.在少許支撐下能蹲下或彎腰撿起地上的東西，然後恢復站的姿勢	是	否
2.能夠雙手拿大東西—如搬小塑膠椅或抱大玩具向前走一小段距離（約十步左右）不會跌倒	是	否
★ 3.至少有10個穩定使用的語詞（娃娃語如「ㄇㄢㄇㄢ」為吃、汪汪為狗亦可）	是	否
★ 4.（實作）能正確指認至少一個圖形（圖1：大人依序問「哪一個是筆？鞋子？鑰匙？魚？飛機？杯子？」全部問完再從頭問一輪，必須兩次均指對的圖形才算對，以避免兒童因亂指而猜對）正確率：＿/ 6＿	是	否
5.能正確指出至少四個身體部位（大人依序問「頭、手、腳、眼、耳、鼻、嘴在哪裡？」）正確率：＿/ 7＿	是	否
★ 6.模仿做家事或使用大多數的家用器具（如掃地、用衛生紙擦東西、玩開關、玩鍋碗瓢盆等）	是	否
★ 7.有主動探索學習的動機，例如：會自己去把玩具找出來玩、或自己拿故事書出來翻看	是	否
★ 8.高興時會和別人分享喜悅：例如轉頭面對大人微笑，或把喜歡或得意的東西展示給大人看	是	否
9.無法模仿說單詞，因為（1）根本沒有仿說動機，或(2)發音困難以致難以聽懂	是	否
10.通常自顧自玩，大人反覆叫喚名字（或小名）多次仍然不理會，沒有任何抬頭、轉頭看、或回到大人身邊的反應	是	否
11.檢核過程中非常不合作，出現下列任一行為如（1）不聽說明、不看示範（2）眼睛不跟隨大人手指方向（3）不會指給大人看（4）把大人的東西搶過去自己玩（5）跑來跑去抓不住（6）似乎聽不懂指令	是	否

圖1



Appendix B

Items in the Productive Phonology Task

	Phoneme	Stimuli		Phoneme	Stimuli
1	b-z, sh-b	杯子，手錶	12	j-d, -j	剪刀，眼鏡
2	p-g, -p	蘋果，巫婆	13	q-, g-q	青蛙，國旗
3	m-z, x-m	帽子，小貓	14	x-g, d-x	西瓜，大象
4	f-j, -f	飛機，衣服	15	zh-z, sh-zh	桌子，時鐘
5	d-h, h-d	電話，蝴蝶	16	ch-sh, h-ch	廚師，火車
6	t-, sh-t	太陽，手套	17	sh-z, l-sh	獅子，老鼠
7	n-p, x-n	奶瓶，小鳥	18	r-g, sh-r	熱狗，生日
8	l-zh, x-l	蠟燭，小鹿	19	z-b, x-z	嘴巴，洗澡
9	g-, x-g	國王，小狗	20	c-m, q-c	草莓，青菜
10	k-z, d-k	筷子，短褲	21	s-b, -s	掃把，雨傘
11	h-z, l-h	猴子，老虎			

Appendix C

Stimuli in the Discrimination Task

Contrast	Target	Distractor	Feature difference
One feature	*gou3* (dog)	*dou3* (dipper-like bin)	place
(*1-f difference*)	*niao3* (bird)	*liao3* (knotweed)	manner
Two features	*hu3* (tiger)	*pu3* (music score)	place + manner
(*2-f difference*)	*xi3* (wash)	*ji3* (crowded)	manner + aspiration
Three features	*mao1* (cat)	*tao1* (big waves)	aspiration + place + manner
(*3-f difference*)	*biao3* (watch)	*qiao3* (skillful)	aspiration + place + manner

Appendix D

(1) Nonce Word Repetition

Practice			
1	gu1-lai4		
2	kang1-bi3	pan4-dao1	
3	fan4-tu1	ma1-gai4	bao3-xi1

One-word List	
1	xia4-nei4
2	biao1-luo3
3	jiao1-fei1
4	fang1-li4
5	fa1-qü1
6	biao3-jia1
7	bing1-da4
8	xin1-bai3
9	huo4-jin3
10	mo4-dan1
11	xie4-lang3
12	bin4-geng3

Two-word List		
1	bao4-li3	gao1-dan4
2	guo4-xin1	ta1-liu4
3	dui4-ding3	gong1-ba1
4	huan1-lu3	xü1-kun3
5	fei3-bo4	nan4-gu3
6	xing3-duan4	gua1-jiang1

Three-word List			
1	ben3-guan1	deng3-mu4	tian1-gai1
2	qi1-dang1	ma1-ding4	dan1-bu4
3	qie4-xiang1	niu3-xi1	huang3-lei4
4	qiao1-hao4	hui3-ta4	hun1-pao1

Note: Children below age 3 only had to perform the shaded stimuli items.

(2) Gap Word Repetition

Practice			
1	mie3-bou4		
2	lang1-te3	die4-hai1	
3	nin1-lan1	pen3-mian1	diu4-qun1

One-word List	
1	le3-gun1
2	nen3-die4
3	pei3-mu1
4	ga3-mang4
5	miu1-fo4
6	nan1-hao1
7	fai4-pua1
8	tiu1-kei4
9	biang1-tua4
10	giang1-biu1
11	fin1-dua4
12	lia1-pe3

Two-word List		
1	pa3-gei4	lan1-kao1
2	diu3-hou1	qiu4-nuan1
3	ni1-xia3	ten4-kuai1
4	kiang1-bou1	fi1-duai4
5	hiang1-dua1	fao4-bun1
6	hi4-luang1	mia3-gin4

Three-word List			
1	men3-xiong4	pen3-man1	kao1-tai3
2	fou4-que3	kuo1-te3	lu1-biao4
3	dei4-muai3	hin1-tuang4	piu4-lua1
4	nua4-ki1	gia3-miang4	luai1-nia1

Note: Children below age 3 only had to perform the shaded stimuli items.

Appendix E

(1) The Descriptive Statistics of the Age 2 Cohort

		Time 1			Time 2			Time 3		
Average age (month)		24.92			31.10			37.00		
Age range		23.83-26.63			29.93-32.93			35.57-39.53		
N		24			22			21		
Measure		*N*[a]	Mean	SD	*N*[a]	Mean	SD	*N*[a]	Mean	SD
MCDI-EXPRESSIVE		24	304.87	109.89	22	567.59	85.47	22	665.91	28.83
MCDI-RECEPTIVE		24	471.95	99.77	22	638.32	60.23	22	686.05	10.99
PPVT-R		==			22	18.82	6.55	22	29.86	8.94
REVT		==			==			21	34.57	9.77
Production_onset (42)		24	23.03	7.32	22	29.59	7.22	21	32.85	4.50
Production_rhyme (6)		24	4.24	1.15	22	4.73	1.00	21	5.26	.72
Discrimination (%)		20	56.25	14.53	22	70.45	14.02	22	75.57	9.56
Nonword Repetition										
Half	Nonce (to 2w)[b]	23	7.43	4.92	21	13.76	4.82	21	17.10	4.18
	Nonce (to 3w)[c]	==			21	16.95	5.89	21	22.38	5.76
	Gap (to 2w)[b]	23	6.57	4.90	21	12.81	3.64	21	14.71	3.12
	Gap (to 3w)[c]	==			21	14.57	4.81	21	17.86	4.41
All	Nonce	==			==			21	43.57	10.58
	Gap	==			==			21	30.76	7.06
Leiter-R		==			22	53.59	7.06	==		

Note: [a] The number of participants that completed the task. [b] Children's score in half of the nonword repetition task up to the two-word lists. [c] Children's score in half of the nonword repetition task up to the three-word lists.

(2) The Descriptive Statistics of the Age 3 Cohort

	Time 1	Time 2	Time 3
Average age (month)	36.80	42.77	48.70
Age range	35.47-38.93	41.47-44.63	47.37-50.73
N	24	23	20

Measure		*N*[a]	Mean	SD	*N*[a]	Mean	SD	*N*[a]	Mean	SD
PPVT-R		24	24.92	9.25	23	32.35	7.96	20	40.55	9.60
REVT		24	32.54	13.02	23	46.65	12.16	20	60.95	12.59
Production_onset (42)		24	32.21	4.57	23	34.66	4.53	20	36.38	3.20
Production_rhyme (6)		24	5.37	0.75	23	5.39	0.77	20	5.53	0.76
Discrimination (%)		24	63.72	12.07	23	74.46	11.81	20	81.25	10.25
Nonword Repetition										
All	Nonce	24	34.88	13.68	23	42.04	9.67	20	49.00	9.15
	Gap	24	25.25	10.41	23	29.70	10.84	20	34.05	9.48
Leiter-R		==			23	51.30	7.00	==		

Note: [a] The number of participants that completed the task.

(3) The Descriptive Statistics of the Age 4 Cohort

	Time 1			Time 2			Time 3		
Average age (month)	48.83			54.97			60.89		
Age range	47.36-50.80			53.67-56.90			59.57-62.57		
N	24			24			23		
Measure	*N*[a]	Mean	SD	*N*[a]	Mean	SD	*N*[a]	Mean	SD
PPVT-R	24	34.63	11.48	24	48.58	11.32	23	58.52	15.13
REVT	24	53.50	12.52	24	65.88	11.60	23	81.91	8.76
Production_onset (42)	24	34.99	3.39	24	37.40	2.54	23	38.96	1.70
Production_rhyme (6)	24	5.72	0.41	24	5.81	0.40	23	5.81	0.46
Discrimination (%)	24	76.56	12.64	24	80.56	11.44	23	83.88	10.68
Nonword Repetition									
All Nonce	24	49.25	8.68	24	54.17	6.46	23	57.57	5.82
Gap	24	36.17	6.90	24	39.50	7.84	22	44.86	7.55
Leiter-R	==			24	56.21	10.72	==		

Note: [a] The number of participants that completed the task.

Appendix F

(1) The Correlation Matrix of All Measures in the Age 2 Cohort

	MCDI-R_T1	Onset_T1	Rhyme_T1	Discri._T1	Nonce_T1	Gap_T1	MCDI-E_T2	MCDI-R_T2	PPVT_T2	Onset_T2	Rhyme_T2	Discri._T2	Nonce_T2	Gap_T2	Leiter_T2	MCDI-E_T3	MCDI-R_T3	REVT_T3	PPVT_T3	Onset_T3	Rhyme_T3	Discri._T3	Nonce_T3	Gap_T3
MCDI-E_T1	.77**	.43*	.48*	.25	.31	.32	.59**	.35	.42	.35	.38	-.19	.43	.50*	-.08	.26	.29	.40	.20	.23	.03	.15	.51*	.50*
MCDI-R_T1		.33	.46*	.39	.33	.42*	.71**	.59**	.19	.19	.56**	-.41	.61**	.75**	-.11	.37	.49*	.44*	.10	.19	.09	.15	.71**	.46*
Onset_T1			.75**	.24	.64**	.63**	.23	.11	.37	.78**	.55**	.15	.48*	.20	.36	.38	.19	.45*	.46*	.56**	.36	-.09	.64**	.40
Rhyme_T1				.20	.76**	.76**	.49*	.27	.32	.43*	.85**	.07	.44*	.29	-.04	.46*	.27	.44*	.23	.49*	.62**	-.07	.60**	.26
Discri._T1					.04	-.01	.45	.40	.07	.12	.35	.10	.12	.27	.30	.19	.04	.43	-.09	.15	.05	.15	.43	.33
Nonce_T1						.78**	.29	.25	.36	.25	.61**	.16	.49*	.43*	-.08	.34	.18	.60**	.33	.31	.46*	-.03	.57**	.19
Gap_T1							.51*	.37	.38	.17	.69**	-.14	.58**	.55**	.05	.15	.14	.47*	.19	.31	.40	-.14	.67**	.38
MCDI-E_T2								.81**	.35	.11	.52*	-.27	.51*	.44*	-.10	.36	.41	.40	.04	.37	.11	-.20	.66**	.38
MCDI-R_T2									.39	.05	.39	-.32	.47*	.42	-.08	.34	.38	.49*	.17	.38	.15	-.06	.48*	.29
PPVT _T2										.31	.16	.04	.40	.00	.06	.23	.25	.47*	.62**	.17	.05	-.04	.31	.19
Onset_T2											.34	.10	.31	-.02	.12	.31	.22	.25	.67**	.67**	.29	-.01	.39	.23
Rhyme_T2												-.15	.54*	.47*	-.07	.49*	.30	.41	.14	.36	.80**	-.04	.67**	.24

	MCDI-$_{R}$_T1	Onset_T1	Rhyme_T1	Discri_T1	Nonce_T1	Gap_T1	MCDI-$_{E}$_T2	MCDI-$_{R}$_T2	PPVT_T2	Onset_T2	Rhyme_T2	Discri._T2	Nonce_T2	Gap_T2	Leiter_T2	MCDI-$_{E}$_T3	MCDI-$_{R}$_T3	REVT_T3	PPVT_T3	Onset_T3	Rhyme_T3	Discri._T3	Nonce_T3	Gap_T3
Discri._T2													-.44*	-.47*	-.01	.05	-.11	-.14	.06	-.04	-.01	-.12	-.39	-.53*
Nonce_T2														.62**	-.05	.39	.32	.65**	.37	.33	.31	.09	.82**	.33
Gap_T2															.03	.08	.17	.53*	.01	.17	.13	-.02	.59**	.59**
Leiter_T2																-.10	-.22	-.01	-.14	-.03	-.21	-.24	.10	.47*
MCDI-$_{E}$_T3																	.86**	.54*	.46*	.39	.43	-.06	.58**	-.01
MCDI-$_{R}$_T3																		.42	.49*	.32	.11	-.12	.52*	.06
REVT_T3																			.57**	.51*	.25	.16	.64**	.42
PPVT_T3																				.53*	.19	.09	.30	.03
Onset_T3																					.29	-.15	.39	.27
Rhyme_T3																						.11	.30	-.11
Discri._T3																							-.04	-.21
Nonce_T3																								.55*

Note: MCDI-$_{E}$ = MCDI-$_{EXPRESSIVE;}$ MCDI-$_{R}$ = MCDI-$_{RECEPTIVE;}$ Onset = Onset production; Rhyme = Rhyme production; Discri. = Discrimination; Nonce = Children's performance in half of the nonce word repetition task up to the 2-word list; Gap = Children's performance in half of the gap word repetition task up to the 2-word list.

(2) The Correlation Matrix of All Measures in the Age 3 Cohort

	PPVT _T1	Onset_ T1	Rhyme _T1	Discri. _T1	Nonce _T1	Gap_ T1	REVT _T2	PPVT _T2	Onset_ T2	Rhyme _T2	Discri. _T2	Nonce _T2	Gap_ T2	Leiter_ T2	REVT _T3	PPVT _T3	Onset_ T3	Rhyme _T3	Discri. _T3	Nonce _T3	Gap_ T3
REVT_T1	.75**	.45*	.52**	.60**	.56**	.66**	.80**	.29	.44*	.41	52*	.69**	.47*	.68**	.63**	.63**	.35	.47*	.43	.65**	.66**
PPVT_T1		.11	.29	.34	.36	.49*	.76**	.54**	.11	.24	.51*	.59**	.36	.62**	.50*	.50*	.35	.21	.39	.51*	.54*
Onset_T1			.72**	.43*	.68**	.66**	.49*	-.09	.91**	.62**	.27	.59**	.52*	.17	.63**	.42	.69**	.66**	-.03	.52*	.42
Rhyme_T1				.42*	.62**	.61**	.57**	-.03	.61**	.94**	.02	.56**	.42*	.16	.63**	.49*	.59**	.90**	.09	.39	.45*
Discri._T1					.52**	.57**	.54**	.10	.41	.33	.56**	.56**	.52*	.46*	.32	.41	.18	.32	.50*	.31	.54*
Nonce_T1						.85**	.67**	.05	.65**	.37	.28	.86**	.84**	.34	.67**	.40	.48*	.40	.41	.83**	.84**
Gap_T1							.68**	.17	.49*	.43*	.47*	.87**	.77**	.46*	.65**	.50*	.46*	.35	.36	.82**	.84**
REVT_T2								.40	.49*	.48*	.58**	.84**	.73**	.66**	.83**	.65**	.43	.46*	.43	.72**	.72**
PPVT _T2									.01	.02	.40	.20	.20	.56**	.26	.64**	.40	-.01	.07	.15	.23
Onset_T2										.48*	.22	.50*	.45*	.20	.56*	.42	.72**	.56*	-.05	.51*	.37
Rhyme_T2											.07	.39	.25	.11	.50*	.48*	.58**	.95**	-.01	.16	.25
Discri._T2												.51*	.50*	.54**	.37	.40	.24	-.03	.39	.37	.42

	PPVT_T1	Onset_T1	Rhyme_T1	Discri._T1	Nonce_T1	Gap_T1	REVT_T2	PPVT_T2	Onset_T2	Rhyme_T2	Discri._T2	Nonce_T2	Gap_T2	Leiter_T2	REVT_T3	PPVT_T3	Onset_T3	Rhyme_T3	Discri._T3	Nonce_T3	Gap_T3
Nonce_T2													.87**	.63**	.72**	.52*	.36	.43	.50*	.84**	.87**
Gap_T2														.49*	.63**	.38	.30	.20	.60**	.79**	.90**
Leiter_T2															.60**	.62**	.16	.23	.39	.54*	.61**
REVT_T3																.57**	.44	.52*	.07	.61**	.56*
PPVT_T3																	.57**	.49*	.21	.45*	.42
Onset_T3																		.48*	-.16	.34	.24
Rhyme_T3																			.00	.19	.25
Discri._T3																				.47*	.65**
Nonce_T3																					.85**

Note: Onset = Onset production; Rhyme = Rhyme production; Discri. = Discrimination; Nonce = Children's performance of the complete form of nonce word repetition task; Gap = Children's performance of the complete form of the gap word repetition task.

(3) The Correlation Matrix of All Measures in the Age 4 Cohort

	PPVT _T1	Onset_ T1	Rhyme _T1	Discri. _T1	Nonce _T1	Gap_ T1	REVT _T2	PPVT _T2	Onset_ T2	Rhyme _T2	Discri. _T2	Nonce _T2	Gap_ T2	Leiter_ T2	REVT _T3	PPVT _T3	Onset_ T3	Rhyme _T3	Discri. _T3	Nonce _T3	Gap_ T3
REVT_T1	.51*	-.08	-.08	.39	.27	.15	.62**	.54**	-.00	.15	.40	.33	.13	.49*	.55**	.57**	.06	.29	.46*	.30	.16
PPVT_T1		.03	.07	.15	.08	.30	.38	.59**	-.03	-.11	.44*	.36	.23	.25	.35	.48*	.14	-.04	.29	.26	.28
Onset_T1			.08	.16	.46*	.44*	-.15	-.07	.69**	-.08	.00	.48*	.48*	-.13	.-.18	.11	.51*	-.14	.19	.53**	.29
Rhyme_T1				.27	.05	.29	-.19	-.04	.01	.52*	.25	.09	.29	.15	.17	-.08	.15	.53**	-.11	.22	.08
Discri._T1					.22	.16	.18	.32	.15	-.07	.64**	.14	.14	.35	.43*	.36	.02	.14	.42*	.26	-.14
Nonce_T1						.77**	.43*	-.00	.47*	.16	.13	.74**	.81**	.07	.48*	.35	.44*	.11	.46*	.78**	.66**
Gap_T1							.28	.03	.45*	.27	.16	.67**	.77**	.09	.26	.17	.55**	.21	.32	.71**	.68**
REVT_T2								.52**	.15	.14	.19	.51*	.29	.30	.67**	.66**	.19	.14	.46*	.44*	.37
PPVT _T2									-.12	-.22	.42*	.29	.13	.33	.38	.68**	-.03	-.10	.39	.23	.02
Onset_T2										.09	-.03	.52**	.51*	-.05	-.12	.18	.78**	.06	.24	.51*	.26
Rhyme_T2											-.20	.12	.16	.14	.19	-.01	.22	.94**	-.14	.25	.28
Discri._T2												.18	.18	.17	.42*	.35	.03	-.01	.47*	.15	-.03

	PPVT_T1	Onset_T1	Rhyme_T1	Discri._T1	Nonce_T1	Gap_T1	REVT_T2	PPVT_T2	Onset_T2	Rhyme_T2	Discri._T2	Nonce_T2	Gap_T2	Leiter_T2	REVT_T3	PPVT_T3	Onset_T3	Rhyme_T3	Discri._T3	Nonce_T3	Gap_T3
Nonce_T2													.87**	.24	.39	.46*	.55**	.07	.28	.82**	.86**
Gap_T2														.10	.43*	.34	.52*	.10	.23	.80**	.78**
Leiter_T2															.35	.32	.07	.23	.16	.03	.17
REVT_T3																.68**	-.00	.22	.33	.44*	.30
PPVT_T3																	.17	.08	.29	.43*	.19
Onset_T3																		.13	.39	.39	.32
Rhyme_T3																			-.11	.17	.15
Discri._T3																				.25	-.01
Nonce_T3																					.70**

Note: Onset = Onset production; Rhyme = Rhyme production; Discri. = Discrimination; Nonce = Children's performance of the complete form of nonce word repetition task; Gap = Children's performance of the complete form of the gap word repetition task.

Appendix G

Convergence analysis

Full model

Level 1:

$$\text{Nonword repetition} = \pi_0 + \pi_1\left(age - \overline{cohort\ age}\right) + e_{ti}$$

Level 2:

$$\pi_{0i} = \beta_{00} + \beta_{01}(age\ 3\ cohort) + \beta_{02}(age\ 4\ cohort) + u_0$$
$$\pi_{1i} = \beta_{10} + \beta_{11}(age\ 3\ cohort) + \beta_{12}(age\ 4\ cohort) + u_1$$

Reduced model

Level 1:

$$\text{Nonword repetition} = \pi_0 + \pi_1\left(age - \overline{grand\ age}\right) + \pi_2\left(age - \overline{grand\ age}\right)^2 + \pi_3\left(age - \overline{cohort\ age}\right) + e_{ti}$$

Level 2:

$$\pi_0 = \beta_{00} + u_0$$
$$\pi_1 = \beta_{10}$$
$$\pi_2 = \beta_{20}$$
$$\pi_3 = u_3$$