PHONOLOGICAL PROCESSING SKILLS IN YOUNG LEARNERS' EFL VOCABULARY ACQUISITION

YU-CHENG SIEH

Thesis submitted for the degree of

Doctor of Philosophy

Integrated PhD in Educational and Applied Linguistics

Newcastle University

School of Education, Communication,

and Language Sciences

November 2007

ABSTRACT

Four factors were examined to investigate English vocabulary learning among 9-year-old elementary school children in Taiwan. The four factors were use of their native language, length of English instruction, and two phonological processing capabilities—phonological memory and phonological sensitivity. Apart from a series of paper-and-pencil and computerized vocabulary assessments, two nonword repetition tasks along with five detection and production tasks of rimes and phonemes were used to measure phonological memory and phonological sensitivity. The young learners' scores on all vocabulary tests were positively correlated with phonological memory and phonological sensitivity, as was also evidenced in studies by Gathercole et al. (1997) and Bowey (1996). A similar pattern of association was found between the learners' vocabulary performance and their length of English instruction. However, neither of the two phonological processing capabilities was associated with English instruction length.

The reaction times of the two online vocabulary tests suggest that an extra input of L1 gloss in explicit vocabulary teaching might have resulted in faster aural recognition of single English words. This supports Kroll and Stewart's (1994) revised hierarchical model of bilingual representation, which postulates that beginning L2 learners have their two languages interconnected at the lexical level.

Results of stepwise and hierarchical regression analyses confirmed that English phonological sensitivity was the best predictor of young learners' English vocabulary performance and contributed uniquely to their vocabulary scores after age, English instruction length, vocabulary knowledge from school textbooks, Chinese phonological sensitivity, and phonological memory were statistically controlled.

ACKNOWLEDGEMENTS

The completion of my thesis would not have been possible without help and support from many people here in England and back in Taiwan.

First and foremost I would like to say a big thank-you to my supervisor, Dr. Nikhat Shameem, for her guidance and helpful advice throughout my writing-up period. I am also grateful to Mr. Scott Windeatt who took over the supervision after Nikhat left. Special thanks go to Dr. Annabelle David for her valuable comments on my work and her continuous support since she supervised my M.A. dissertation.

Secondly, I would like to express my enormous gratitude to my family for their understanding and unconditional love over the years, without which I could never have achieved what I have accomplished today. In particular, I would like to thank my sister, Yu-Tai, who helped make arrangements for my data collection and provided help in every way.

I would also like to thank the teachers and friends back in Taiwan for their support and encouragement. I could not have reached this stage without reference letters from Professor Burt King and Professor Yuen-Mei to start with. Heartfelt thanks go to Kuen-Hung Tsai and Cheng-Chang Lee for their help in statistic analysis. Special thanks go to my friends Cheng-Yi Chang, Yu-Ting Chen, and Ling-Yu Guo for their jokes and encouragement which never failed to cheer me up.

Last but not least, I would like to thank my subjects and their teachers for their participation and cooperation. I have enjoyed the experience working with them and realized that there is still more to find out about how children learn a new language.

CONTENTS **Page**

List of Tables

NEWCASTLE UNIVERSITY

SCHOOL OF EDUCATION, COMMUNICATION, AND LANGUAGE SCIENCES

Author's Declaration

I certify that, to the best of my knowledge, all the material in this thesis represents my own work and that no material is included which has been submitted for any other award or qualification.

Signature: ______________

Date: 13/11/2007

ABBREVIATIONS

EFL	English as a foreign language
ESL	English as a second language
FL	Foreign language
IPA	International Phonetic Alphabet
L1	First language
L2	Second language
MOE	Ministry of Education
ms	millisecond
RT	reaction time
SLA	second language acquisition
SLI	specific language impairment

CHAPTER 1: INTRODUCTION

1.1 Background to the Research Problem

English is gaining more importance than ever all over the world as it is used as an international language among speakers from different first-language (L1) backgrounds. In Taiwan, English is regarded by parents as one of the crucial components of a child's future success (Tsai, 2001). In response to the growing importance of English and the demands of parents, Taiwan's Ministry of Education (MOE) reacted by advancing the age of starting to learn English from 13 to 11 in 2001 and again to nine years of age in 2005, over a span of just four years.

Among the various component skills of language learning, vocabulary is particularly important to beginning learners. Research on native child speakers of English shows that two phonological processing skills—phonological memory and phonological sensitivity—are closely associated with children's language acquisition (Bowey, 1996; Gathercole, Willis, & Baddeley, 1991). In other words, children with a larger receptive vocabulary are also found to have better phonological processing capabilities. However, for immigrant children, age of acquisition is an important determinant of ultimate attainment in their second language (Johnson & Newport, 1989; McDonald, 2000).

The aim of the present study was to investigate whether these factors which facilitate vocabulary acquisition in native child speakers of English and children learning English as a second language (ESL) would perform similar functions in Taiwanese children, who learn English as a foreign language (EFL). The disputed issue of L1 translation specific to the EFL context has also been incorporated into the study. In an EFL context where a constant target language input is absent, age and good language models are believed to contribute directly to children's vocabulary learning (Huang, 2006; Liu, 2006; Yin, 2006). Although the focus of a large majority of vocabulary studies in Taiwan has been on investigating whether particular learning strategies ease young learners' learning loads (Cheng, 2005; Wu, 2005), they have not been able to account for a wide range of individual differences among these learners. It is hoped that the perspective of this study will contribute further to the area of EFL vocabulary research.

1.2 Research Context

1.2.1 English Education and Availability in Taiwan

English is currently a mandatory subject from the third grade at elementary school level to at least the first year at university, spanning 11 years. Instruction times range from two sessions of 40 minutes per week at primary school to at least three sessions of 45 minutes per week at secondary school and at least three sessions of 50 minutes per week in the universities across the public school system. The grammar-translation teaching method is widely used in the six years of secondary school instruction to prepare students for university entrance exams. However, recent years have seen a slow but gradual shift to the communicative approach in the wake of complaints that students can hardly speak the language at all after long hours of instruction (H.-C. Chen, 2005).

Books, magazines, newspapers, and TV programmes in English are widely available in Taiwan. English TV series and films are subtitled in Chinese but not dubbed. Bilingual English-Chinese magazines in which vocabulary and grammar points are extensively explained are popular among students and the white-collar class. There are three English language newspapers produced by local publishers in circulation, and international English newspapers, such as the *International Herald Tribune*, are available from newsstands. Street name signs are in both Chinese and English scripts. It is not uncommon to see some shop signs in English too.

Evening English classes are a common phenomenon across Taiwan and cater for learners of different ages and needs. Many students, ranging from elementary school pupils to high school students, attend these evening lessons to consolidate their vocabulary and grammar knowledge for entrance exams. Such lessons, which take place after school hours and are offered by private institutions other than state schools, are consequently termed as extracurricular English lessons.

1.2.2 English Teaching at Elementary School Level

Despite a wide availability of English language resources around Taiwan, English is taught and learned as a foreign language (FL). In other words, English is rarely spoken outside the language classroom and the chances to practise it are very limited. English instruction at elementary school level, consequently, focuses mainly on strengthening basic listening and speaking skills in the third and fourth grades and on furthering communication skills in the fifth and sixth grades.

In the aspect of vocabulary teaching, the MOE (2006) suggests that English instruction should cover a wide variety of topics closely related to the students' daily lives. Various genres, such as cartoons, nursery rhymes, and short stories, should be included to stimulate interest in young learners when publishers select the contents of teaching materials. In order to provide references for material designers, the MOE compiled and published a list of 2,000 most frequently used words and suggested that teaching materials for elementary school level be written using the first 1,200 of these. The pupils are expected to be able to spell and read at least 180 words and use 300 words in their oral production by the time they finish their elementary school education, and be able to use all 1,200 words when they graduate from junior high school.

1.2.3 Chinese Education at Elementary School Level in Taiwan

Children in Taiwan are exposed to the spoken forms of one ethnic language, for example, Taiwanese, and Mandarin Chinese after they were born. But they do not learn to read and write Chinese until they are seven years old and enter elementary school where the ethnic language is taught only in the spoken form. It is necessary to point out that strategies of teaching Chinese in Taiwan are very different from those employed in China. Since Chinese characters are predominantly pictographic and each character is represented by its own distinctive shape, a separate phonetic script, known as Zhu-Yin-Fu-Hao, is introduced to Taiwanese children in the first 10 weeks of their first year before any instruction in reading and writing Chinese characters is given. Taiwanese children first learn to read each of these 37 phonetic symbols and then learn how to put two or three of them together to pronounce a Chinese character. An analogy can be made between the strategies Taiwanese children use to read Chinese and the grapheme-phoneme correspondence rules native English child speakers employ to spell English words, with the difference that all Chinese characters are monosyllabic.

All texts in the Chinese textbooks of the lower elementary grades are marked with the Zhu-Yin-Fu-Hao symbols, and the reliance on the Zhu-Yin-Fu-Hao script to learn Chinese characters continues into the fourth year. New Chinese characters continue to be marked with Zhu-Yin-Fu-Hao symbols in the fifth and sixth grades, but texts no longer have the script printed alongside them.

1.3 Motivation and Purposes of the Study

Research into how young learners learn English as a foreign language is increasing in volume in Taiwan in the wake of the inclusion of English in the elementary school curriculum. Setting out with the goal of easing learning loads on young learners, most of the researchers carried out experiments to assess which learning strategy or teaching method yielded the best learning outcome. As most of them looked at the products of young learners, they failed or were unable to explain the individual differences which might have resulted in performances of various levels among their sample of subjects.

In contrast to a large body of studies which have aimed to investigate the achievements of young learners, the present study is principally concerned with investigating the nature of individual differences among these learners in their process of vocabulary acquisition in relation to four variables. They are use of L1 translation, length of English instruction, phonological memory, and phonological sensitivity. As Miyake and Friedman (1998) pointed out, understanding the nature of individual differences in L2 proficiency is essential because "the way in which learners differ from one another can not only provide useful constraints on the theories of L2 learning but also throw some new light on how to maximize the outcome of L2 learning and instruction" (p. 339-340).

There were hence three major purposes behind the current study. The first was to investigate the relationship between vocabulary learning and the two phonological processing abilities—phonological memory and phonological sensitivity—among young learners with a first-learned nonalphabetic language background in their initial stage of English learning. In recent years, vocabulary acquisition has been largely attributed to phonological working memory, which has been examined using a nonword repetition framework (Gathercole & Baddeley, 1989, 1990). The association between nonword repetition and vocabulary development has been extended into the field of foreign language learning (Service, 1992). Meanwhile, similar patterns of association have been observed between vocabulary and phonological sensitivity (Bowey, 1996). While several recent studies have investigated the two skills and compared their predictive power with regard to young children's vocabulary acquisition, these studies were restricted to speakers of an alphabetic language (de Jong & van der Leij, 1999; Gathercole et al., 1991; Hansen & Bowey, 1994; Muter & Snowling, 1998). No research, to the best knowledge of the researcher, has been conducted to compare their association with EFL vocabulary learning among speakers of a first-learned logographic language.

As extracurricular English lessons are extremely popular in Taiwan, the second purpose was to find out whether a greater and longer input of English was associated with the subjects' phonological processing abilities as well as with their vocabulary learning outcome. Speakers of a nonalphabetic language, for example, Chinese, are said to lack phonological processing capabilities (Cheung, 1999; Read, Zhang, Nie, & Ding, 1986), and hence their vocabulary learning result should be attributed instead to their language input. The activities of vocabulary learning would reveal whether greater input could help learners to learn English words more quickly and more thoroughly. The results, in turn, would also show whether individual differences in vocabulary learning in an EFL context were caused by extrinsic factors, for example the amount of English input, or by intrinsic variables, such as the two phonological processing abilities.

The third purpose was to investigate the effect of the wide, but controversial, use of the L1 by teachers when teaching young learners. In contrast to a popular all-English teaching method, L1 glosses were found to facilitate children's retention of English vocabulary in an EFL context (Chang, 2005; Hong, 2005). Learners already equipped with a first language are destined to learn a second language (L2) differently from the way in which they approached their first language (Vygotsky, 1986). The study consequently will contribute to the research into bilingual lexico-semantic memory which investigates how L2 vocabularies are learned, stored, and accessed in L2 learners. The results have important theoretical implications for the development of models of language processing as well as pedagogical implications for second-language teaching, since an understanding of the mental lexicon is crucial to the formulation of any models of language processing (Schreuder & Weltens, 1993).

The motivation for looking at these four variables is consequently closely related to the purposes of the study. If the factors which facilitate English native child speakers' vocabulary acquisition perform similar functions on Taiwanese children, the policy to lower the age to start English instruction at elementary school level does not serve as the only answer to successful foreign language acquisition in the initial stages. While children can be given trainings to improve their phonological processing skills which in turn ease their English vocabulary learning loads, more time can be allocated to native language education in the elementary school curriculum.

1.4 Research Questions

1.4.1 Main Research Question

What is the best predictor of young learners' successful acquisition of receptive vocabulary in an EFL context?

1.4.2 Subquestions

1. Is the use of L1 a boost or a hindrance to young learners' acquisition of English vocabulary in the foreign language classroom?
2. Does a longer length of English instruction result in better acquisition of English vocabulary? Is it helpful to young learners' phonological processing skills?
3. Is nonword repetition a good measure to predict young learners' vocabulary acquisition? What is the association between nonword repetition and receptive vocabulary?
4. Is phonological sensitivity a good predictor of young learners' receptive vocabulary acquisition? What is the association between phonological sensitivity and receptive vocabulary?
5. Do these two aspects of phonological processing skills explain independent or overlapping variation in young children's acquisition of receptive vocabulary?

There are four variables under investigation in the present study. They are L1 translation, English instruction length, phonological memory, and phonological sensitivity, as listed in the first four subquestions. L1 translation and English instruction length are termed extrinsic factors because they qualitatively and quantitatively modulate the learners' English vocabulary acquisition. In contrast, phonological memory and phonological sensitivity are called intrinsic factors because they are likely to be inherent in a child's vocabulary learning ability, but vary from one learner to another. Each of the four subquestions investigates the individual relationship of each of these variables to EFL vocabulary learning. Taken together, the answers to these subquestions will explain young EFL learners' vocabulary acquisition. Taking into account the fact that the two phonological processing skills were found to be comparable in their power to predict vocabulary development in studies of native child speakers, the last subquestion attempts to weigh and determine the importance of each phonological processing skill in terms of the role each plays in the vocabulary learning of young children learning English as a foreign language. The answers to the five subquestions together provide the answer to the main research question and have important pedagogical implications for English vocabulary teaching in the EFL context.

A battery of assessments was administered to seek answers to the subquestions. They consisted of two nonword repetition tasks, five phonological sensitivity assessments in the Chinese and English languages, two online vocabulary tests, 13 written vocabulary tests, and one productive vocabulary test. The results from the nonword repetition and phonological sensitivity tasks were used to investigate whether there was a correlation between the young learners' phonological processing skills and their vocabulary learning, while the online vocabulary tests were used to examine whether instruction in the target language would facilitate faster mapping between an object and its L2 phonological form. All of the assessments were carried out on the premises of an elementary school in southern Taiwan from September to November 2005, spanning three months.

1.5 Hypotheses and Implications

1.5.1 Hypotheses

1.5.1.i Hypothesis 1: First Language Effects

It was hypothesised that young learners who were given L1 translation equivalents to the target lexical items would need longer reaction times (RTs) in response to auditory and visual cues, in contrast with those who were offered no L1 translations. It was presumed that young learners were likely to process an FL word in the order of an L2 phonological representation, L1 translation, followed by its referent, when an auditory cue was received. Subjects exposed to L1 translations would gradually come to process the L2 phonological form first in its L1 before linking it with its referent. Such language processing would cause a delay when the learners tried to retrieve the target language from their mental lexicon, and this was expected to be apparent in their longer reaction times.

At the beginning of the experiment, young learners who were exposed to L1 translations were expected to map objects with their L2 phonological representations much faster because the L1 provided a short cut linking the L2 sounds to the L1 concept in their schemata. At the same time, it was thought highly likely that those who were not exposed to L1 translations would spend more time creating the concept in the L2 in their schemata. However, it was expected that after both groups of subjects had acquired the vocabulary, the subjects with no L1 translations would process FL vocabulary more quickly in response to the auditory input in their online tasks.

1.5.1.ii Hypothesis 2: Vocabulary Acquisition and Phonological Processing Abilities

It was expected that the strong correlations between vocabulary acquisition and phonological processing abilities observed among native English-speaking children would also be exhibited among young children learning English as a foreign language who had a nonalphabetic first language. Those who performed better on nonword repetition tasks and phonological sensitivity assessments would score higher on vocabulary tests. In other words, the weakness of those who performed poorly in vocabulary tests would be reflected in their performance in nonword repetition and phonological sensitivity tasks. It was predicted that better young language learners would not only react faster but would also score higher in the online phonological sensitivity assessments owing to their superior phonological awareness.

1.5.2 Implications for English Language Teaching

English learning has become, to some young learners, an even longer hard-fought battle since the language was introduced into the primary school curriculum in Taiwan. English teachers cannot understand why some children seem still to be poorer language learners, despite the fact that a variety of teaching methods and approaches are now being incorporated into English language teaching. If the factors which underlie vocabulary acquisition in native child speakers are helpful to EFL learners, phonological processing skills should be incorporated into early school English instruction to ease learning loads on the part of learners and to improve their learning outcomes.

The research design, incorporating certain online assessments, was aimed to trace how young learners store and recall their English vocabulary, and this helps to shed light on the long-debated status of the L1 in the foreign language classroom. As these young EFL learners are equipped with a first language which is so dominant in their daily lives, the variable should not be simply dismissed. Instead, since some studies suggest that use of the L1 facilitates the retention of FL vocabulary, L1 use should be given the credit it deserves and applied to TEFL.

1.6 Organisation of the Study

This work is divided into six chapters. The first chapter explains the research background and the purposes of the present study. The second chapter presents an extensive review of the literature on vocabulary learning in an EFL context and a description of the relationship between vocabulary learning and the four variables under investigation. The third chapter, on methodology, talks about the instruments and procedures of data collection as well as containing a description of the pilot study and its result. The fourth chapter describes the results of thorough analyses of the data from the main study in order to provide answers to the research questions. The fifth chapter, which is the discussion chapter, relates the results from the previous chapter to the relevant literature review and presents a detailed examination of the links between them. The last chapter provides a conclusion and points out the implications and limitations of this study.

The examination of the extrinsic factors has been arranged to precede that of the intrinsic factors so that the association between English instruction length and the two phonological processing skills may be made clear and separated from the discussions of the link between these two skills and vocabulary learning. This arrangement has also been applied to the chapters containing the literature review, the results, and the discussion.

CHAPTER 2: LITERATURE REVIEW

This chapter begins with a description of the importance of vocabulary learning and how it occurs in native and EFL contexts. Four factors crucial to vocabulary learning in EFL children were explored in this study; the chapter is hence organised around them. The first two factors—L1 translation and English instruction length—are termed *extrinsic* factors, because they contribute to young learners' vocabulary development by qualitatively and quantitatively modulating FL instruction. In EFL vocabulary teaching, L1 translation equivalents are discouraged in an attempt to link the object and the target language conceptually. However, previous research has suggested that a provision of L1 glosses facilitates learners' immediate recall and retention of FL lexical items (Liao, 2005; Lotto & de Groot, 1998). This literature review will focus on bilingual memory research in order to provide an explanation of the role of L1 translation in influencing how FL words are processed and stored in learners' minds. In Taiwan, extracurricular English lessons are a common phenomenon. Popular wisdom often attributes EFL success to longer English instruction. This research investigates the reasons for the popularity of extracurricular English instruction and attempts to determine whether opinions regarding its influence on later achievement are well-founded.

In contrast, the other two factors—phonological memory and phonological sensitivity—are referred to as *intrinsic* factors, as they are related to each individual's vocabulary learning ability and thus vary from learner to learner (Bowey, 2001; Metsala, 1999). The literature review will focus on the association of these two factors with vocabulary learning in native language acquisition and on links to EFL learning.

It should be noted that English as the target language is for the most part interchangeably referred to as FL or L2 throughout this work. As the subjects of the present study are bilingual speakers of Taiwanese and Mandarin Chinese, they possess two first languages. In the elementary school curriculum, native languages comprise Mandarin Chinese and one ethnic language from Taiwanese, Hakka, and an Austronesian language of the indigenous peoples, depending on the ethnicity of a young learner (MOE, 2006). Apart from Mandarin Chinese and one ethnic language, English is the only compulsory language subject in the curriculum and hence is regarded as the students' second language, despite its status of being a foreign language in the Taiwanese context.

2.1 Vocabulary

Speaking a new language involves learning at least its three subsystems—phonology, vocabulary, and syntax. There are, however, more aspects contained within the knowledge domains of a language, for example, semantics and morphology, to name just two. Despite the fact that lexical acquisition is only a small part of attaining proficiency in a new language, vocabulary learning is inevitably viewed as a central component of verbal comprehension since it is a precondition for speaking and understanding a language (Miller, 1978). In cognitive psychology, vocabulary acquisition is seen to be one of the most important aspects of human development, as vocabulary development is likely to be part of a child's general cognitive development (Bialystok, 2001; Gupta & MacWhinney, 1997). The strong emphasis placed on such a small part of language development is consequently necessary (Zimmerman, 1997).

2.1.1 Receptive Versus Productive Vocabulary

Most models of lexical knowledge distinguish between receptive and productive vocabularies mainly because of their sizes and growth rates as well as their different input conditions (Ellis & Beaton, 1993; Laufer, 1998). The distinction between the two types of vocabulary in most cases depends on its resemblance to the distinction between the receptive skills of listening and reading and the productive skills of speaking and writing.

Receptive vocabulary refers to those that can be recognized and have their meanings recalled when they are met. An individual's use of receptive vocabulary knowledge involves perceiving the phonological or orthographical form of a word and retrieving its meaning. Productive vocabulary involves those which can be retrieved for purposes of speaking or writing at the appropriate time and its use involves retrieving and producing words phonologically or orthographically after a thought is generated (Nation, 2001). Goulden, Nation, and Read (1990) suggested that a well-educated adult native speaker of English has a receptive vocabulary of approximately 17,000 base words. It is, nevertheless, hard to estimate the size of an adult native speaker's productive vocabulary because of different habits of word use. But it is commonly agreed that receptive vocabulary size is larger than productive vocabulary size (Aitchison, 2003; Melka, 1997) and that the former develops at a faster rate than the latter (Laufer, 1998; Laufer & Paribakht, 1998).

A distinction between receptive and productive learning of vocabulary is thus recommended because the goals that teachers or learners set for learning English will affect the way vocabulary is selected and learned (Nation, 1990; Rixon 1999). In the present study, only the spoken forms of English words, accompanied by their corresponding pictures, will be used in the explicit vocabulary teaching sessions and the vocabulary assessments. No written forms of English words will be shown to the pupils so as to avoid the possibility that some young learners who had a longer length of English instruction might have an advantage in an input condition containing orthographic presentations over those who did not (de Groot & van Hell, 2005).

2.1.2 Vocabulary and Literacy Development

Vocabulary is important because it is strongly correlated with later literacy development in both L1 and L2, especially in the aspect of reading comprehension (Beck, McKeown, & Omanson, 1987; Bensoussan, 1992; Braze, Tabor, Shankweiler, & Mencl, 2007; Hazenberg & Hulstijn, 1996; Laufer, 1992; Perfetti & Lesgold, 1977; Yovanoff, Duesbery, Alonzo, & Tindal, 2005). As reading comprehension is a complex interaction between the reader and the text, unknown words can create gaps when a learner attempts to construct meaning from a text and eventually result in a breakdown in comprehension if there are too many of them.

Reading comprehension is often predetermined by an individual's vocabulary knowledge. Coady, Magoto, Hubbard, Graney, and Mokhtari (1993) found a positive and significant relationship between knowledge of high-frequency vocabulary and reading proficiency, and also found that an increase in the former led to a boost in the latter. In investigating the relationship between Taiwanese EFL learners' vocabulary size and their reading comprehension, Jiang (2004) found that those undergraduate subjects who had a larger vocabulary performed significantly better on overall reading comprehension than those with a smaller vocabulary. Chern (1993) pointed out that Taiwanese college students relied heavily on lexical information when reading in English and that they only started to apply contextual information when their English proficiency level increased. This positive relationship between vocabulary and reading comprehension could be found at the two educational levels preceding university education as well. Shih (2004) found that the English proficiency of Taiwanese junior and senior high school students was correlated with their vocabulary and reading skills, with higher vocabulary scores being linked to better reading performance.

However, a large vocabulary size contributes not only to EFL learners' reading comprehension but also to their listening competence in the target language. Tsai (2005) discovered that Taiwanese college students' receptive vocabulary size was positively associated with both their English listening and their English reading performances. Those who had scored higher in tests of Listening Vocabulary Levels had also obtained higher scores in listening and reading tests, their reading scores being higher than their listening scores. The findings of these studies together demonstrate that vocabulary is crucial to the building up of language skills.

2.1.3 Connection of Form and Meaning in Vocabulary Acquisition

Nation (2001) states that knowing a word means knowing its form (spoken and written), its position (e.g., grammatical pattern and collocation), its function (e.g., frequency and appropriateness), and its meaning (e.g., concept and associations). Each of these components can be associated with comprehension or use. Embedded in the knowledge of a word are the linguistic characteristics of the word, including its phonology, morphology, syntax, and semantics, as well as its functional use. In other words, to know a word means to understand both the concept and the meaning of the word, which can be viewed as two forms of the same process (Vygotsky, 1986).

There are several theories about how children learn to connect the form and the meaning of a word. Among them, fast-mapping is a well-established notion. Using colour words, Carey (1978) demonstrated that children as young as three years old were able to employ their existing linguistic knowledge and rapidly created lexical representations for unfamiliar words they encountered. In a situation where the explicit teaching of vocabulary was intentionally avoided, a child was told to ignore the red plate and bring the carer the *chromium* plate, which referred to the unfamiliar colour of olive. The young children were able to fast-map the new word onto a new item with only a few exposures. It was hypothesised that a child's mental lexicon, as well as his conceptual domain, was re-structured to give place to the new word when the child acquired it. The hypothesised process of fast-mapping was later replicated by Dollaghan (1985, 1987), and it was revealed that normal and language-impaired children did not differ in their ability to map a novel word to its referent. These studies demonstrate that children realise the significance of object labelling and learn to connect form and meaning from an early age.

Aitchison (2003) has divided the word-learning process into various steps and given a detailed description of the three tasks which children are faced with when they acquire their native language. These are labelling, packaging, and network-building. Labelling—the ability to symbolise—develops when children aged between one to two years of age discover that a whole situation is ritually accompanied by a name repeated by their interlocutors (Smith, 1978). Successful labelling involves the repeated matching of a name with its referent. After successfully attaching many objects to their names, children begin to classify a number of them under a particular label, based mainly on the similarity of their shape (Gelman & Ebeling, 1998). As their vocabulary increases, they gradually integrate words into a network and build links between coordinates. It is through network-building that fast word-searching is achieved for the purpose of language comprehension. The three tasks sum up the gradual process of children's vocabulary acquisition.

2.1.4 English Vocabulary Learning in an EFL Context

Learning English vocabulary in a context where English is spoken as a foreign language presents a different picture. To begin with, the learners are equipped with a first language and have passed the stage of learning where they match the form and meaning of words. Furthermore, they have learned from their L1 experience how the world works and hence they are unlikely to retrace their L1 learning route and apply it to their foreign language learning. Vygotsky (1986) has argued that foreign language learning does not repeat the course of native language acquisition but instead employs the native language as a mediator. The FL learners use the semantics of the native language as a foundation and only translate the word meanings that are already well developed in the native language.

It is thus not difficult to see that the mother tongue exerts a great influence over the way a foreign language is learned and used (Corder, 1994; Swan, 1997). Coady et al. (1993) proposed that it was very likely that FL learners, in the case of English vocabulary learning, would attach an English label to an already existing native-language schema rather than building an entirely new schema for frequently occurring concepts which were universal. Hsieh (2005) conducted a study to investigate the impact of English vocabulary learning with and without L1 translation equivalents for the target lexical items on sixth-graders. It was revealed that the group taught using Chinese glosses for the English vocabulary performed significantly better in both an immediate vocabulary recall and a reading comprehension test than the other group who had been taught without using any translation. In another similar study administered to sixth-graders, Liao (2005) found that the provision of L1 translation facilitated a longer retention of English word recognition than the provision of picture clues and written contexts.

In addition to investigating how the first language of learners affects their EFL vocabulary learning, a large body of EFL vocabulary research has been devoted to discovering the most effective strategy with respect to vocabulary learning. P.-C. Chen (2005) revealed that Taiwanese junior high school students used visual, auditory, and kinaesthetic strategies and also individual strategies which were matched to their perceptual learning styles to learn the vocabulary of the target language. In another study, designed to investigate whether instruction in vocabulary learning strategies helped senior high school students, J.-H. Chen (2004) found that his three groups of subjects, each using one of the three strategies of translation, transliteration, and image, did not differ significantly from each other in either the immediate recall test or the follow-up test administered one week later. The fact that such studies failed to converge on one single strategy led researchers to claim that EFL learners who were able to employ a variety of strategies when trying to memorise words were more likely to possess a larger vocabulary (Shih, 2004; Wu, 2005).

In contrast, language instruction length has been found to be closely associated with the English proficiency of EFL learners, especially in the aspect of their listening skills (Y.-Y. Chen, 2004). Intensive FL instruction also contributed to the differences between first- and fourth-year students of English in their vocabulary size and listening comprehension (Tsai, 2005). The fourth-year college students performed significantly better than the first-year students in all tests. The issue of instruction length and its association with language abilities has been under scrutiny lately since the age to start learning a foreign language was lowered to allow young learners more time to acquire the necessary skills. The literature dealing with the two aforementioned variables—L1 translation and length of English instruction—will be reviewed in the next two sections.

2.1.5 Storytelling and Literacy Development

Research on teacher-directed vocabulary teaching has shown that reading books orally to children several times and explaining word meanings while reading is effective with elementary school children (Biemiller, 2003). Very young children successfully construct word meanings from listening to storybooks read aloud (Sénéchal, Thomas, & Monker, 1995). It has been suggested that both frequency of exposure and teacher explanation of the target words enhance vocabulary learning when children listen to stories (Penno, Wilkinson, & Moore, 2002; Robbins & Ehri, 1994).

Stories also conveniently serve to be the starting-point for a variety of language learning activities (Ellis & Brewster, 1991). In the imaginary world created by stories, children are encouraged to exercise their imagination and share their experience. Storytelling activities provide an educational vehicle for authentic language construction because it is effective in facilitating child language development (Block, 1997) and promotes greater language fluency (Wellhousen, 1993).

Storybooks are consequently a good replacement for textbooks because vocabulary is "contextualized, giving the learner a richer sense of a word's use and meaning" (Huckin & Coady, 1999, p.182). The illustrations in the storybooks are useful in providing contextual cues and bringing comprehension to the young learners when they have a small vocabulary size and little prior knowledge of English. While the children develop their listening and concentrating skills, language teachers have the opportunity to introduce and recycle language items and sentence structures through the use of storytelling. Examples of target vocabulary in use are readily given in a story context appealing to young learners. In comparison, new words in elementary school textbooks are often grouped in a lesson for convenience and short of an appropriate context.

2.2 Extrinsic Factor I—First Language

Vygotsky (1986) stipulates that a person's acquisition of both their native tongue and a foreign language belongs to one general class of speech development, even though each language has a process of development entirely different from the other. Unlike the learning of the native language, the acquisition of the foreign language is based on systematic instruction and its process is conscious and deliberate from the start. MacWhinney (2005), in the unified competition model, states that the L2 is parasitic on the L1 in terms of lexical learning because of the extensive amount of transfer from L1 to L2. In the initial stages of learning the L2 system, learners do not formulate a separate conceptual structure. According to the prediction of the competition model, the type of code-based separation occurs only for mappings at the levels of lexicon, phonology and parts of speech but not for underlying distributed conceptual representations (Hernandez, Li, & MacWhinney, 2005). They do not access L2 lexical meanings directly but instead rely on L1 forms. It is hence fair to suggest that FL learners transfer to the new language what they possess in the native language and cut short the process of making sense of the world by mapping an FL word directly onto a word in their native language.

2.2.1 First Language Effects

There are two ways of proving that the L1 inevitably plays an important role in learning a second or foreign language. One is by demonstrating the privileged status of the first language in various learning activities and the other is by examining the context of bilingual development.

2.2.1.i Priority Status of L1

Weinreich (1953) predicted that bilinguals would have their preferred language in which they learned to perform a certain task and that it would be difficult for them to switch to the other language. Owing to the priority in learning, this preferred language is usually their native language. Marsh and Maki (1976) demonstrated that bilingual subjects indeed computed simple addition problems more rapidly in their preferred language, i.e., the language in which these bilingual subjects had learned arithmetic.

With respect to language learning, those learners whose native language is linguistically close to the target language tend to derive an enormous advantage from language transfer. The learning of some L2 lexical items seems to be made easier by the fact that the two languages share cognates (Lotto & de Groot, 1998; Meara, 1993; Nation, 1982). Laufer and Paribakht (1998) took into consideration knowledge of L1 in their comparison of English vocabulary acquisition between a group of Israeli and a group of Canadian-French speakers. It was found that the Canadian subjects who had French as their L1 derived more benefit from cognates while learning English than did the Israeli learners.

Word structure appears to be one area where logographic and alphabetic languages vary greatly. It is hence expected that the variances between them might produce widely different strategies of word-handling in speakers of different languages and that strategies of L1 learning carried over to learning an L2 often cause unexpected impediments (Meara, 1984). Rickard Liow, Green, and Tam (1999) conducted a study to explore the development of visual search strategies employed by Singaporean secondary school students while they read Chinese characters and English letters. The adolescent biscriptal subjects with poor English proficiency showed U-shaped quadratic trends while doing an English letter search task, which are a feature of the search function for reading Chinese characters. An exposure to the English alphabetic script did not displace the pedagogical influence of the whole-word method in English instruction. It is hence cautioned that the factor of mother tongue should be taken into account from the start in bilingual situations because "the importance of priority in learning is likely to be so great in comparison with other psychological factors governing interference" (Weinreich, 1953, p. 88).

2.2.1.ii Context of Bilingual Development

Although it is debatable whether bilingual speakers have a shared or separate memory, it is commonly agreed that context plays a crucial role in the development of bilingualism. Ervin and Osgood (1954) proposed a compound-coordinate model to describe two different types of bilingual speakers who are distinguished by context during their course of learning a second language. A *compound* bilingual learns a foreign language in the school situation and grows up in a home where the foreign language and a home language are interchangeably spoken. As this group of bilinguals are fostered by learning vocabulary lists, they have one language compounded onto the other and their two languages are hence associated with the same set of representational mediation processes or meanings. On the other hand, a *coordinate* bilingual has a home language which is different from that of the community he or she lives in and hence the coordinate bilingual more or less speaks the two languages with the same people and in the same situations. In bilinguals of this group, who are termed 'true' bilinguals by the authors, each set of linguistic signs and responses is appropriate to one particular language and each language is associated with an individual set of representational processes, the two sets of representational processes being different.

Despite the fact that the compound-coordinate model was not supported by subsequent studies on semantic differential measures (Jakobovits & Lambert, 1961; Kolers, 1963), recent research on Singaporean bilingual children and adolescents has demonstrated the effect of different contexts on learners. Rickard Liow and Poon (1998) recruited three groups of third-graders, aged nine to 10 years of age, who had received five years of formal instruction in written English and Chinese. Among the three groups, one group spoke Chinese as their L1 and home language, a second group spoke English, and a third group, Bahasa Indonesian. In a homophone decision task, the Bahasa Indonesian-speaking children, who have a language consisting of a shallow written system with transparent grapheme-phoneme correspondence rules, performed significantly better than the other two groups on nonword matching performance. This shows that a learner's L1 does implicitly affect his L2 learning.

The results of the aforementioned studies point to the possibility that the native language has a crucial role to play during the process of second language learning. This, in turn, spells out the necessity of a connection between psycholinguistic studies investigating the mental representation of the bilingual lexicon and pedagogically motivated vocabulary studies (Meara, 1992, 1997). The next section will describe what an L2 learner's mental lexicon looks like and hence explain how L1 translation comes about in the process of second language learning.

2.2.2 Representations of Bilingual Lexicon

2.2.2.i Bilingual Lexico-semantic Memory

An interesting question when examining the organisation of a bilingual's lexico-semantic representations is in what manner his lexical forms are mapped onto their respective meanings when one more language is added. That is, in other words, how or where are his two languages connected? Researchers devoting themselves to this line of psycholinguistic research have employed similar tasks—word translation in conjunction with word- or picture-naming—which makes comparisons of these research findings easier. The research, by probing how translation comes about in a bilingual's L2, is important in that it offers an interpretation of how a bilingual stores and accesses the lexical and conceptual representations in his memory, and hence sheds light on L2 vocabulary pedagogy in particular (Kroll & Sunderman, 2003).

Potter, So, von Eckardt, and Feldman (1984) proposed two hypotheses regarding the association between equivalent words in a bilingual's two languages. The *word association hypothesis* stipulates that a direct association is established between newly learned L2 words and their corresponding translation equivalents in L1 and hence that the L2 is always mediated through the L1. The *concept mediation hypothesis*, in contrast, posits that L2 words are directly associated with the nonlinguistic concept which is common to the two languages and therefore that words in the two languages are not directly associated. In other words, concepts can be accessed directly by and for L2 words, without L1 activation. An intermediate model is proposed to combine the two hypotheses and stipulates that learners start out only with lexical associations, but the word association link is eventually replaced by the concept mediation link and hence develops direct links between the L2 lexicon and concepts.

The two hypotheses assume a distinction between word representations and their concepts and make different predictions for the relative time required to perform in the naming and translation tasks. The word association hypothesis predicts that picture-naming in the L2 should take longer than translating because two extra steps are needed for the former task. It is assumed that L2 picture-naming requires a retrieval of the concept as well as a retrieval of the L1 word before the name can be given in its L2. Translation of a word from L1 to L2 is direct because the L2 is always mediated through the L1. In contrast, the concept mediation hypothesis claims that there should be little or no difference in the time required for picture-naming in the L2 and that required to translate words from L1 to L2, because concepts can be directly accessed for L2 words. Potter et al. (1984) recruited a group of highly proficient bilingual Chinese-English speakers and compared their performance in picture-naming and L1-L2 translation. It was found that the proficient bilinguals took no longer to name a picture in the L2 than to translate a written L1 word into its L2. This result was found to support the concept mediation model. Surprisingly, the same pattern of results was found in a group of novice learners of French. The authors consequently dismissed the word association hypothesis and concluded that direct conceptual processing of L2 was in place very early in second language acquisition.

However, this finding was subsequently challenged by other research (Chen & Leung, 1989; Kroll & Curley, 1988; Kroll & Sholl, 1992; Kroll & Stewart, 1994). It was pointed out that the less proficient English-French learners in Potter et al. (1984) might in fact have passed the stage of accessing L2 words via the L1 link because they knew enough French to enable them to spend a summer in France, despite the great difference between their amount of training in the L2 and that of the proficient Chinese-English learners (Chen & Leung, 1989; Kroll & Stewart, 1994). On the basis of this argument, Chen and Leung recruited three different groups of subjects further to explore lexical processing for beginning and proficient adult learners as well as beginning child learners. The beginning adult learners had learned French for about two years, while the proficient adult learners and beginning child learners had learned English for 12 and two years respectively. All three groups had Cantonese Chinese as their L1. The results showed a consistently better performance in word-naming over picture-naming when subjects responded in their native language (Potter & Faulconer, 1975; Potter et al., 1984). For proficient subjects, picture-naming in L2 and translating L1 words into L2 were equally efficient, as was predicted by the concept mediation hypothesis. For the adult beginners, translating L1 into L2 was faster than picture-naming in L2, consistent with the word association hypothesis. The differences in patterns of lexical processing in the nonnative language between beginning and proficient learners showed that a developmental change occurred as their L2 proficiency increased.

However, the pattern of results for child beginners was unexpected and inconsistent with either of the tested hypotheses. They named pictures in the L2 440 ms faster than they translated L1 words into L2. It was suspected that these child beginners, whose mean age was about seven, still had decoding problems with L1 words because they made more errors with regard to L1 words than to pictures. Two more experiments were conducted where second- and fourth-graders were recruited. Both experiments replicated the result of the first experiment and showed that child learners were significantly faster in picture-naming than in translating words from L1 to L2. Chen and Leung (1989) suggested that the results might have occurred because the subjects were not yet proficient in reading L1 words and that pictures were typically used in Hong Kong as media for teaching English vocabulary to children. The authors concluded that child beginners showed different patterns from adult beginners and used pictorial representations rather than L1 words to access L2 words.

Kroll and Curley (1988) also questioned the finding by Potter et al. (1984) that all L2 learners could conceptually mediate the new language regardless of their proficiency. They recruited English-speaking adult subjects who were relatively fluent in German and those who had studied the language for less than two years. The data replicated the word-naming advantage over picture-naming in their L1.

When the response was given in L2, the less proficient group took significantly longer than the proficient group in both picture-naming and translation tasks. For the less fluent subjects, translation reaction time was shorter than picture-naming time, as was predicted by the word association hypothesis. For the more fluent subjects, translation and picture-naming times did not differ statistically, just as the concept mediation hypothesis predicted. These data proved that L2 learners not only accessed their first language in order to understand their second language in the early stages of language acquisition but also shifted their reliance on lexical association to a reliance on conceptual mediation in later stages as their proficiency increased.

Kroll and her colleagues further investigated the role of the two hypotheses in the two directions of translation. It was found that translation from L2 into L1 was significantly faster than from L1 to L2 in both fluent and less fluent speakers (Kroll & Sholl, 1992). The results of observed asymmetry in translation directions confirmed their hypothesis that translation from L1 to L2 is conceptually mediated and susceptible to distractors but that translation from L2 to L1 is mediated by lexical connections when categorization tasks are employed in conjunction with the translation tasks (Kroll & Stewart, 1994). It is clear that neither the word association route nor the conceptual mediation route is sufficient effectively to capture the memory representations in a bilingual speaker. The Revised Hierarchical Model is therefore proposed to account for the asymmetric mappings of words to concepts in bilingual memory as well as for the developmental shift in a bilingual's strategies of accessing L2 words (Kroll & Stewart, 1994). The strength of the connections between the lexical and conceptual representations in a bilingual's two languages differs primarily as a function of L2 fluency and the relative dominance of the L1 over the L2.

2.2.2.ii Dissociation Between Lexical and Semantic Representations

When exploring the bilingual lexico-semantic lexicon, it is necessary to bring into discussion the view of a dissociation between lexical and semantic representations. This view has contributed to clearing up the confusion regarding a shared or a separate memory system, and has provided supporting evidence for explanations in research into bilingual memory.

Collins and Loftus (1975) suggested a two-level model of knowledge representation in an individual's memory. In the model, the semantic and the lexical levels of representation are distinguished from each other. The names of concepts are stored in a lexical network. The semantic network, also called the conceptual network, is organised along the lines of semantic similarity, while the lexical network is organised along lines of phonemic and orthographic similarity. A name node in the lexical network is connected to one concept node, and often to more than one concept node, in the semantic network.

Smith and Magee (1980) conducted a series of experiments, the results of which supported the notion that representations in memory should be divided into two levels—semantic and lexical—and which showed that the information contained in each of these two levels would arrive at different rates when called upon. In a naming task, words and pictures were presented either in a control situation without distracting stimuli or in the presence of incongruent distractor items. The result showed that naming of words was rapid and minimally influenced by the presence of distractor pictures. However, the categorisation of words took longer than word-naming and the categorisation of pictures. At the same time, when the task involved picture processing, the amount of interference generated by the presence of an incongruent word was significantly greater when subjects were required to generate the verbal label for the picture than when only a yes-no category decision was required. It is assumed that the simultaneous presentation of incongruent words disrupted the process of seeking a name for a picture, indicating that picture-naming took longer than word-naming and the categorisation of pictures.

However, in a surprise recall task, words in the naming task were recalled considerably less well than when they were in the categorisation task or when it was pictures that were being recalled. The implications from the results are clear with respect to a dissociation between lexical and semantic representations. Firstly, naming and categorisation tasks involve types of different processing in the memory which are available at different rates. Secondly, the involvement of different types of processing indicates that there are different types of representations in the memory.

Accordingly, studies of bilingual memory representations which have used tasks that reflect primarily lexical processes have reported evidence consistent with the separate memory model, because a bilingual's two languages are interconnected at the lexical level. In contrast, studies which have used tasks that reflect primarily semantic processes have reported evidence consistent with the common memory model because a bilingual accesses the semantic memory to which both languages have the same access (Kroll, 1993).

2.2.3 Employment of L1 in Foreign Language Learning

The previous review of bilingual memory representations sums up why L1 translations come naturally to EFL learners. As a foreign word can only be learned if it is understood, an L1 translation corresponding to the FL word seems to be the safest and most economical way of learning it. Results from studies of EFL vocabulary learning have indeed demonstrated that L1 translation equivalents are useful supplements because learners showed better immediate recall and long-term retention of the FL words when their L1 translation equivalents had been provided

(Hulstijn, Hollander, & Greidanus, 1996; Lotto & de Groot, 1998). Laufer and Shmueli (1997) provided vocabulary glosses to their subjects; half of these glosses were explained in English and the other half were translated into their L1 equivalents. Meanwhile, the subjects were assigned to four different teaching modes—(a) isolation, i.e., word lists, (b) minimal context, i.e., in one meaningful sentence, (c) text-context, and (d) elaborated text-context, i.e., original text with clarifying phrases and sentences—in order to have the outcome of their memorisation of new words examined. The results showed that learners had better retention of these FL words when their L1 equivalents were provided, irrespective of in which mode the words had been learned.

Experienced foreign language learners appear to prefer L1 glosses to picture aids in learning new vocabulary. Lotto and de Groot (1998) recruited Dutch undergraduates who had years of experience in foreign language learning to compare the two teaching methods—word-association and picture-association. The results from two recall tests indicated that the presentation of L1-L2 word pairs during learning provided a better opportunity for acquiring L2 words than did the presentation of picture-L2 pairs. van Hell and Candia Mahn (1997) suggest that experienced learners, through their increasing experience in FL learning, would choose to associate the new vocabulary with the corresponding L1 words to achieve the most efficacy.

Similar findings have been observed in Taiwanese learners of English at different educational levels. High school students agreed that Chinese glosses were more effective than English glosses in promoting their incidental vocabulary learning (Chang, 2005; Y.-M. Chen, 2004; Chuang, 2004; Huang, 2002). Among the elementary school pupils, a provision of L1 glosses had a significantly positive effect on subjects' performance in an immediate vocabulary recall test (Hsieh, 2005).

The beneficial effect of L1 use is even more conspicuous among learners of low English proficiency, because they are able to employ top-down processing with the aid of L1 glosses to read texts beyond their current level (Hong, 2005; Huang, 2002). Bensoussan (1992) employed a word translation task to investigate learners' vocabulary knowledge in a reading comprehension test. Students of low proficiency were found to have looked up in the dictionary and written down a large number of translation equivalents in their L1 to assist them with their comprehension of the reading text, in comparison to students of high and medium proficiency. The results indicate the importance of the native language, to which learners attach the target language labels when attending to new words.

This section has reviewed the reasons why the L1 should be taken into consideration in the investigation of EFL vocabulary learning. The assumption that the privileged status of the L1 qualitatively affects EFL vocabulary learning is supported by the research into bilingual memory representations which in turn explains the results in pedagogically motivated vocabulary studies in an EFL context. The next section will examine the other extrinsic factor which quantitatively modulates EFL vocabulary learning.

2.3 Extrinsic Factor II—Length of English Instruction

Studies of language minority students have indicated that it might take limited-English-proficient students three to five years to develop English oral proficiency and four to seven years to develop academic proficiency (Cummins, 1991; Hakuta, Butler, & Witt, 2000). Even among L2 learners, proficiency is largely determined by the number of years of language learning. Favreau and Segalowitz (1983) found that their group of bilingual subjects who had equal reading rates in their L1 and L2 had been schooled longer in their L2 than the other group of bilingual subjects, who had a difference of more than 10% in their reading rates between L1 and L2. As it is true that differing amounts of vocabulary knowledge account for the most significant differences among L2 readers and that years of training in the target language are necessary to cross a linguistic threshold (Alderson, 1984; Haynes & Baker, 1993), a longer length of instruction in and a greater exposure to the target language are often resorted to as a solution in an EFL context in an attempt to bridge the gap. This solution is connected to the notion of earlier-is-better, which in turn inevitably carries the implication that there is a critical period when children should learn a second or foreign language.

2.3.1 The Notion of Earlier-Is-Better

Lenneberg (1967) proposed that the primary acquisition of language, i.e., L1 acquisition, very much hinged upon a certain developmental stage between the ages of two and around 13 and was quickly outgrown when puberty was reached. In terms of second language acquisition, Oyama (1978/1982b) suggested that human beings appeared to be better able to analyse, integrate, and fully utilise a new language if they approached it early in life than if they did so after the early teens. The identification of a critical period for successful language acquisition in L1 and L2 inevitably leads to the popular notion of 'earlier-is-better' when considering at what age to start learning a foreign language.

The notion of earlier-is-better appears to be confirmed by the difference in vocabulary size between native and ESL speakers. Native child speakers of English are estimated to know around 5,000 words when they are six years old and continue to learn an average of 1,000 words or more every year until they are 18 years of age (Nagy & Herman, 1987). When ESL children attend school with native English-speaking children, the preexisting gap in language proficiency between them is shown continuing to widen, because while the ESL children have to acquire oral and academic English, they have to keep pace with the native children, who continue to develop their language skills (Hakuta et al., 2000). However, the gap between the two groups of children is also found to exist in the aspects of phonology and syntax, to name just two. In most ESL phonology and syntax research, the subjects, many of whom were immigrants to an English-speaking country, only acquired at best a near-native English pronunciation or syntactic knowledge if they arrived in their host countries after teenage years (DeKeyser, 2000; Guion, Flege, Liu, & Yeni-Komshian, 2000; Johnson & Newport, 1989; Moyer, 1999; Oyama, 1976/1982a).

Age of acquisition has hence been labelled an important determinant in respect of L2 learners' ultimate attainment. The age at which the subjects started learning the target language was found to account for a large variance of significance in the aspect of phonology (Asher & Garcia, 1969; Flege, Munro, & MacKay, 1995; Guion et al., 2000; Seliger, Krashen, & Ladefoged, 1975/1982) and syntax (DeKeyser, 2000; Johnson & Newport, 1989; McDonald, 2000).

2.3.2 Length of Instruction as Predictor of EFL Proficiency

If there is a wide gap between the English skills of native English-speaking children and those learning English as a second language, this gap is naturally wider when it comes to children learning English as a foreign language, since they receive significantly less exposure to the target language. As the notion of earlier-is-better posits that more time allows children more contact with the target language, an early start in English offers young learners opportunities to learn more words and language skills (Burstall, Jamieson, Cohen, & Hargreaves, 1974; Carroll, 1969; Macnamara, 1966). This notion hence appeals to EFL learners and is highly influential in the EFL context, to the extent that it is used to argue in favour of advancing the age of language learning. The report at the completion of 10 years' research into the teaching of French in primary schools in England and Wales shows that FL achievement was primarily a function of the amount of time spent studying that language (Burstall et al., 1974). In the study, each group of pupils reached higher levels of achievement in those aspects of learning French to which the greater share of the available teaching time had been devoted.

Several other empirical studies have also proved that a longer length of instruction is associated with better performance in the target language. In his study of more than 3,000 subjects who had recently graduated from elementary schools, Nieh (2004) found that the third year was a dividing line in terms of determining their future language proficiency. Those who started English lessons before their third grade performed better than those who started in or after it. Yin (2006) observed a highly significant relationship between elementary school graduates' length of

English instruction and their listening and reading proficiency. The longer these graduates had learned English during their six years of elementary school, the better their English proficiency when they were admitted to the junior high school.

Discrepancies in English proficiency have not only begun to appear when pupils graduate from elementary school (Y.-P. Chen, 2004; Nieh, 2004; Yin, 2006), but have also continued to grow at high school level. Liu (2006) revealed that an early start had a great influence on the English-speaking skills of senior high school students. The younger the age they started English instruction, the better they performed in the aspects of pronunciation, fluency, grammar, and vocabulary size when they were at senior high school. This association with an early start, on the other hand, was not extended to students' reading abilities (Y.-Y. Chen, 2004).

However, the advantage possessed by early starters in an EFL context seems to disappear in the wake of intensive exposure in higher education. While investigating the effects of vocabulary level and syntactic competence on the English reading comprehension of Taiwanese college students of English, Chao (2004) discovered that there was no significant relationship between these EFL learners' length of English learning and their performances in vocabulary, syntax, or reading comprehension. Coincidentally, C.-Y. Chen (2005) revealed that first-year students from different senior high schools all around Taiwan admitted to a private university between 1998 and 2005 did not differ among themselves in overall English proficiency from year to year. If length of instruction in or exposure to the target language is the major determinant of language proficiency in an EFL context, the grades should have indicated a steady yearly improvement, because many pupils were sent to extracurricular English lessons by their parents in response to an earlier launch of an English programme in elementary schools in several cities starting in 1997 (Y.-P. Chen, 2004). On the contrary, these first-year students did not differ significantly in the aspects of grammar and reading comprehension year by year, despite the observed trend of better listening comprehension year by year. Furthermore, a late 1980s study on the association between English language skills and an early start indicates that the group who started their English education at elementary school level did not differ in their English proficiency from the other group, who started at junior high school level, by the time both groups graduated from the junior high school (Chou, 1989).

Carroll has argued that FL competence is largely "a matter of the amount of time spent in learning, rather than the actual age of starting" (1969, p. 63). It is hence understandable that early starters possess better language skills than later starters because they have had a longer length of English instruction. However, it is probable that the factor of instruction length is not the determinant of EFL learners'

English proficiency, because intensive exposure to the target language at a later stage of language learning is likely to overcome the difference between early and late starters.

These results reflect the need for a shift of focus in SLA research. Bialystok (1991) noted that research into children's cognitive and linguistic development has shifted in recent years from product descriptions of their accomplishments to process analyses of the causes of development. In the area of second language vocabulary acquisition, a processing-oriented approach enables researchers to turn their attention to psycholinguistic factors which account for successful second language acquisition among child learners. The following two sections of this literature review will focus on the two phonological processing capabilities which are empirically proved to account for individual differences in the processes of children's acquisition of L2 vocabulary: phonological memory and phonological sensitivity.

2.4 Intrinsic Factor I—Phonological Memory

Phonological processing capabilities are essential for learning phonologically-based alphabetic languages where learning the sound patterns of new words is a crucial component of vocabulary acquisition. Without a stored specification of the phonological structure of a word, an individual can neither recognise that word spoken by others nor utter that word himself in spontaneous speech. Infants as young as four months old were shown to be able to discern nonnative speech contrasts without relevant experience, but this ability was found to shift from a language-general towards a language-specific pattern during the first year of life (Fernald, Swingley, & Pinto, 2001; Polka & Werker, 1994; Werker & Tees, 2002). Words within a particular language have a highly characteristic distribution of phonological structures so that native speakers are able to accumulate phonotactic knowledge and develop a perceptual system during the course of native language acquisition (Speciale, Ellis, & Bywater, 2004). The gradual development of the perceptual system enables young children to detect foreign accents and to repeat the phonetic presentation of foreign words containing phonemes which are totally absent from their mother tongue by selecting from among those phonetic elements in their mother tongue which are closest to the foreign words (Lengyel, 1995).

2.4.1 Working Memory

Baddeley and Hitch (as cited in Gathercole & Baddeley, 1993) used the term *working memory* to refer to the short-term memory system which is concerned with the dynamic activities of processing and storage of information. The working memory model has three components: the central executive, the phonological loop, and the visuo-spatial sketchpad. The central executive, which has a finite capacity but is the most important component, regulates information flow and coordinates activity within the working memory. The phonological loop, which is responsible for temporarily storing unfamiliar sound patterns, contributes immensely to language learning. The visuo-spatial sketchpad is mainly responsible for material which has a strong visual or spatial component and hence has an insignificant role to play in the language domain (Gathercole & Baddeley, 1993).

One important assumption of the working memory theory is that some individuals have more resources available to them than do others when performing a cognitive task, such resource availability accounting for differences among individuals in task performance. Perfetti and Lesgold (1977) suggested that less-skilled L1 readers need more time than skilled readers in order to reach the same level of text comprehension because of differences in their speed of verbal coding. In a reading experiment, they gave subjects as much time as they needed to read a passage. The less-skilled readers took longer than the skilled readers, but were able to display the same level of recall patterns. The authors interpreted the results as indicating differences in the two groups' speed of "access and retrieval of a word name and the retrieval of its contextually constrained semantic properties" (p. 151). In the same study, they conducted a listening experiment in which the rate of presentation was controlled by the speaker, and expected different performances from their subjects. The less-skilled subjects were not able to comprehend the listening passage fast enough and consequently did not recall as much as the skilled ones. The results show that when more attention is required, more resources are consumed and a person's processing of information becomes consequently slower (McLaughlin, Rossman, & McLeod, 1983). If a learner is slow in word recognition or in the access and retrieval of word meanings, too many attentional resources will be allocated for that purpose, while the comprehension and development of the text of the incoming message may be hampered because insufficient attentional resources are available (Segalowitz, Watson, & Segalowitz, 1995).

Miyake and Friedman (1998) hence propose that working memory may be one central component of language aptitude which contributes to individual differences in L2 proficiency, since working memory has been proved to be an important determinant of language proficiency in the L1 (Adams & Gathercole, 2000; Papagno & Vallar, 1995). Their proposal corresponds to the notion of automaticity in L2 acquisition (McLeod & McLaughlin, 1986). In the cognitive process of learning, when learners consistently map the same input to the same pattern of activation successfully over many trials, familiarity is established and automatic processing is achieved. Automaticity has important consequences in L2 acquisition because more resources are freed to process higher-level information, and this inadvertently distinguishes fluent from nonfluent abilities among learners (Segalowitz & Hulstijn, 2005). It is widely accepted that many bilingual speakers who are normally fluent in their L2 nevertheless read more slowly in that language than in their native language (Favreau & Segalowitz, 1983).

2.4.2 Mechanism of Phonological Memory

Being part of the working memory system, the phonological loop is the mechanism of phonological memory which maintains verbally coded information and comprises two components—the phonological short-term store and the subvocal rehearsal (Gathercole & Baddeley, 1993). It represents material in a phonological code which decays with time, where a process of subvocal rehearsal serves to refresh the decaying representations in the phonological store. The subvocal rehearsal is also used to recode nonphonological inputs, such as printed words or pictures, into their phonological form, to be held in the phonological store. Spoken speech information will bypass the articulatory rehearsal and gain direct access to the phonological store. Baddeley, Gathercole, and Papagno (1998) propose that the primary purpose of the two-component phonological loop is to store unfamiliar sound patterns while more permanent memory records are being constructed. Its use in retaining sequences of familiar words is secondary, as was evidenced in young children's performance in learning phonologically unfamiliar names and familiar names for toy animals (Gathercole & Baddeley, 1990).

The contribution of the phonological loop to the learning of the sound patterns of new words is demonstrated among children as well as adults. An examination of infants' early verbal imitation during the first year of life indicates that the ability to copy words not yet in their repertoires predicts their lexical development in their second year of life and might facilitate their future lexical development (Masur, 1995). Ellis and Beaton (1993) conducted an experiment which demonstrated that learning the sounds of foreign vocabulary words is mediated by the phonological loop. Adult learners were given German words with English translations to learn under four different learning conditions. One group was instructed to adopt the rote repetition method. It was found that rote repetition significantly enhanced learners' performance in a recall task which required them to give the German words when cued by their English translation equivalents. The authors suggest that the subjects' articulatory representations were established through repeating the unfamiliar phonological representations, and hence rehearsing in the phonological loop.

An experiment involving articulatory suppression conducted by Ellis and Sinclair (1996) demonstrated the disruption of language learning that occurred when the subcomponent of subvocal rehearsal was unavailable. Participants were instructed to learn Welsh words with English translation equivalents under the three conditions: silent, repetition, and articulatory suppression. Subjects in the articulatory suppression condition were required to count a whispered one to five in a continuous cycle. Simple counting from one to five was effective in preventing the participants silently or loudly from rehearsing the new phonological codes, but was not disruptive enough to prevent the learning of new words. The results showed that the participants in the repetition condition were significantly superior to their counterparts in the articulatory suppression condition in receptive assessments to comprehend and translate words, as well as in productive tasks where they were asked to pronounce the foreign vocabulary accurately. These studies have demonstrated that the long-term learning of foreign vocabulary is enhanced by repetition but disrupted by articulatory suppression, which in turn proves the importance of the phonological loop in learning new sound patterns.

An individual's capacity to temporarily retain verbal material increases with age (Gathercole & Baddeley, 1993). Roodenrys, Hulme, and Brown (1993) explored the mechanisms responsible for short-term memory span and its development by examining the relationship between memory span and speech rate for words and nonwords of different spoken lengths. They found significant differences in speech rate between one- and two-syllable items as well as between two- and three-syllable items, while the differences interacted significantly with age. This age factor was seen to be responsible for the fact that younger children have slower speech rates and hence recall words more slowly than older children. Brown and Hulme (1996) hypothesised that short-term memory limitations were more likely to constrain the acquisition of long vocabulary items than short ones and hence suggested that long words would be acquired at least a year later than short words of an equivalent frequency. Word length is indeed another critical factor affecting memory span across languages. Roodenrys et al. (1993) found among their subjects of different age groups that memory span increased as word length decreased and that the pattern held true for both words and nonwords. Ellis and Hennelley (1980) eliminated the possibilities of individual differences by recruiting fluent bilingual speakers of Welsh and English and tested their digit span in the two languages. A significant difference was observed in their reading time of long Welsh and comparatively shorter English digits. It took on average 385 ms to read a Welsh digit and 321 ms to read an English digit.

Despite the constraints imposed by age and word length on phonological memory, it appears that rote repetition helps to enhance learners' memorisation of new sound patterns, and a provision of L1 translation successfully creates long-term memory representations for new L2 words. Hulme, Maughan, and Brown (1991) found that the memory span for Italian words was shorter than for English words among a group of native English-speaking subjects, but that memory span would increase substantially after English translations to the Italian words were given. The authors interpreted this finding as indicating that the memory span increased because long-term memory representations were created for the Italian words after the subjects had learned the English translations. This result has demonstrated that long-term memory created in another language mode can in turn assist short-term memory.

2.4.3 Phonological Memory and Vocabulary Acquisition

An individual's capacity to hold correctly a novel sound pattern in temporary memory constrains his vocabulary acquisition. A clear piece of evidence that the phonological loop component of working memory is directly related to the learning of new words was provided by the case study of a neuropsychological patient (Vallar & Baddeley, 1984; Baddeley, Papagno, & Vallar, 1988). The patient, P. V., who acquired a very specific deficit interpreted as a damaged phonological store in the wake of a stroke, had a very poor retention of auditorily presented materials. In a later word-pair associate learning task, P. V. was found to have a normal capacity for learning to associate pairs of meaningful words, but she was completely unable to learn word associations when the second member of a word pair was changed into a Russian word, i.e., a nonsense word. This result suggests that long-term phonological learning may depend largely upon the short-term phonological store component of the phonological loop. In addition to neuropsychological patients, a direct association between phonological short-term memory and vocabulary acquisition has been established by research on adults, and on normal and language-impaired children in their learning of native as well as foreign languages.

2.4.3.i Vocabulary Acquisition in Native Language

Gathercole and Baddeley (1989) conducted a longitudinal study to explore the hypothesis that the short-term phonological memory might play a role in the acquisition of vocabulary by young children. A total of 150 children, aged between 4.0 and 5.2 years at initial testing, were recruited and given a battery of five tests and then retested a year later. The tests included a nonverbal test, a receptive vocabulary test, a single-word reading test, a sound mimicry test, and a nonword repetition test. Results obtained at each of the two years of testing were referred to as Age 4 and Age 5 scores. Age 4 scores showed that each of the four measures was significantly correlated with vocabulary scores while the greatest amount of variance in the vocabulary scores was accounted for by the nonword repetition test. When a series of fixed-order stepwise regression analyses were further computed, and the variance accounted for by chronological age and nonverbal intelligence was partialled out, nonword repetition still explained a significant 15% of the variance in vocabulary scores, but the sound mimicry test did not account for any further variance beyond that of nonword repetition.

The same analyses were conducted to estimate the extent to which these four measures were associated with vocabulary performance at Age 5. Neither chronological age nor nonverbal score was significantly associated with vocabulary. However, strong associations were found between the vocabulary scores and performance on the nonword repetition and the sound mimicry tests. After the variance explained by chronological age and nonverbal intelligence was partialled out, the nonword repetition results still accounted for a substantial 21 % of the variance in the vocabulary scores. The results indicated a stable association between vocabulary knowledge and repetition performance which could not be attributed to the more general cognitive factors of intelligence or chronological age. Furthermore, the nonword repetition performance at Age 4 was a significant predictor of vocabulary skills one year later, suggesting that phonological short-term memory might mediate the long-term storage of phonological information in vocabulary development. Interestingly, the group with high vocabulary scores were also better at repeating nonwords in general than the group with low scores, suggesting an association between nonword repetition and existing vocabulary knowledge. This revealed that the repetition accuracy of a nonsense word depended largely upon the level of its wordlikeness and also that it interacted with phonological short-term memory or long-term lexical knowledge to a varying degree (Gathercole, 1995).

Gathercole, Hitch, Service, and Martin (1997) conducted a study in which the traditional digit span measure was added alongside the nonword repetition task to assess phonological memory. A total of 11 tests were administered to child subjects who had just entered the second year of full-time education. Apart from three measures of vocabulary knowledge and two tests of nonverbal ability, four experimental word learning tasks which contained varying amounts of phonological information were used in conjunction with the two phonological memory tests. The design provided the opportunity to compare the links between young children's word learning ability and the two measures of phonological short-term memory. Both phonological memory measures were consistently found to be significantly linked with young children's vocabulary knowledge, but nonword repetition was more strongly associated with vocabulary than was digit span. In terms of the association between the two phonological memory tasks and word learning ability, the results were consistent with the hypothesis that immediate memory processes were directly involved in the learning of new vocabulary items, but not with known words (Gathercole & Baddeley, 1990). Performances in the two phonological memory tasks were *un*related to children's abilities to learn word-word pairs but were significantly associated with performance on the word-nonword pair learning and recall of new names. However, they were distinguished by their links with the recall of word definitions, with which nonword repetition but not digit span scores were significantly associated. It was concluded that young children's abilities to learn the sound patterns of new words are related both to their current knowledge of the native vocabulary and to their capacity to hold phonological representations for brief periods in short-term memory.

In an attempt to investigate whether the relationship between vocabulary knowledge and phonological short-term memory remains constant throughout childhood, Gathercole, Service, Hitch, Adams, and Martin (1999) compared the performance of two groups of children aged approximately five and 13 years of age in terms of their phonological memory skills and vocabulary knowledge. A more difficult nonword memory task involving the recall of nonword pairs was designed to test the older children, while the younger group was given the usual single nonword repetition, and a digit span was also administered to both groups. Results showed that the scores of the two phonological memory tasks were significantly associated with vocabulary knowledge at both five and 13 years of age, indicating continuity in the contribution of the phonological loop to vocabulary development from the early to late childhood period. The research on native vocabulary acquisition confirmed its direct association with phonological short-term memory within the nonword repetition paradigm.

2.4.3.ii Vocabulary Acquisition in a Foreign Language

The close association between phonological short-term memory and vocabulary development has been discovered to exist not only in native but also in foreign vocabulary acquisition. Superior phonological memory function is associated with greater facility in acquiring a foreign language among children as well as adults (Dufva & Voeten, 1999; Papagno & Vallar, 1995; Service, 1992; Service & Kohonen, 1995).

Child learners

Service (1992) found in a three-year longitudinal study that the ability to repeat English-sounding pseudowords was a good predictor of learning English as a foreign language during the first two to three years of EFL teaching. Aged from nine to 10 at the initial testing, the Finnish pupils who had a higher accuracy rate in their pseudoword repetition task also had a higher English grade two and a half years after the first testing. Service concluded that a young language learner's ability to form unfamiliar phonological representations in working memory underlay the acquisition of new vocabulary items in foreign language learning. Service and Kohonen (1995) continued Service's study and collected data from the same group of subjects for the fourth year. The participants' pseudoword repetition performance was again positively correlated with both their overall English grade and their English vocabulary scores, measured by the school communicative test, respectively. There were even significant correlations between English learning and the earliest pseudoword repetition scores obtained before and at the very beginning of their English instruction. These correlations remained significant even after the measure of their general academic achievement was statistically controlled.

Adult learners

Papagno and Vallar (1995) recruited 10 nonpolyglots and another 10 polyglots who could speak two other languages fluently in addition to their native language of Italian. A battery of tests was administered to explore the relationship between phonological short-term memory and vocabulary learning. It was found that the polyglots produced a superior level of performance to that of the nonpolyglots in verbal short-term memory tasks—auditory digit span and nonword repetition—as well as in a paired-associate learning test of Russian nonsense words. The authors suggested that the construct measured by the two phonological memory tests along with the nonword learning task corresponded to the capacity of phonological short-term memory. The data suggested that the successful acquisition of vocabulary in foreign languages by polyglots involved a relevant contribution from the phonological short-term memory, those with a greater capacity of phonological memory being likely to learn foreign language vocabulary more efficiently.

2.4.4 The Nonword Repetition Paradigm

A nonword is a string of phonologically unfamiliar letters whose combination follows the phonotactic rules of a specific language but does not constitute a real word. As learning the sound patterns of new words is one of the major tasks in foreign language acquisition, the analogy between a nonword and an unfamiliar foreign language word is hence drawn to describe the association between nonword repetition and word learning.

Based on a series of studies conducted by Gathercole and Baddeley, the nonword repetition paradigm is widely used to tap an individual's capacities to retain and repeat sequences of verbal material over short periods of time. Gathercole and Baddeley (1993, 1997) proposed that nonword repetition was a purer measure than the traditional task of digit span in terms of measuring phonological memory, since the former typically correlated highly with vocabulary knowledge. Compared with word or digit span measures, nonwords have the crucial advantage of being nonlexical material. Word or digit span traditionally employed in memory tasks is confounded by varying degrees of familiarity with the to-be-remembered material. The situation is especially susceptible to the fluency and proficiency levels of learners in an EFL context. Those who have had more L2 experience might benefit from their familiarity with the to-be-repeated words or digits, while those who have had little L2 training would struggle with the phonological representations of each digit or word. The aforementioned considerations make the employment of nonsense words a more sensible choice than that of the traditional methods in an EFL context.

The nonword repetition paradigm has its shortcomings, however. As performance on nonword repetition depends largely on an individual's phonological memory, which is in turn constrained by long-term lexical knowledge, *high-wordlike* nonwords are usually better recalled than are *low-wordlike* nonwords, because high-wordlike nonwords bear a higher degree of resemblance to real words than do low-wordlike nonwords. This result indicates the possibility of a lexicality effect, where the repetition of the unwordlike stimuli depends largely on phonological working memory, while the repetition of the wordlike nonwords is mediated by long-term lexical knowledge and is therefore less sensitive to phonological memory constraints (Gathercole, 1995; Gathercole & Pickering, 1999; Gathercole et al., 1999). Dollaghan, Biber, and Campbell (1993, 1995) used pairs of identical multisyllabic nonwords, the only difference between them being that in one of them the syllable carrying primary stress corresponded to a monosyllabic English word, to test the effect on repetition performance. The results showed that normally achieving school-age boys repeated nonwords with lexical stressed syllables significantly more accurately than those with nonlexical stressed syllables. Subsequent analyses also revealed that the majority of repetition errors occurred because the participants attempted to transform nonlexical sequences into real words, and even the transformation violated both strong acoustic cues and articulatory ease.

The other inconvenience embedded in nonword repetition is the demands it makes on spoken output. It has been suggested that nonword repetition might reflect speech perception, articulation, and mental manipulation between perception and production, in addition to phonological memory (Bowey, 1997). Caution is hence advised when tempted to make the assumption that inaccurate nonword repetition reflects deficits in working memory (Bowey, 1996, 1997; Dollaghan, 1998; Edwards & Lahey, 1998).

In an attempt to compare the differences in spoken word recognition between children with and without specific language impairment (SLI), Dollaghan (1985, 1987, 1998) presented three types of words—unfamiliar words, familiar and phonologically related words, and familiar but phonologically unrelated words—to the school-age subjects. The two groups did not differ significantly in the point at which they recognised familiar words. However, the subjects with SLI were less successful in representing the essential phonological characteristics of new words in their lexicons, i.e., speech output, so as to distinguish them from existing, phonologically related word entries. They required significantly more of the acoustic-phonetic signal than did their peers to recognise unfamiliar words. As phonetic information is the most vulnerable aspect of children's fast-mapping process, the problems of retrieving the full phonological representations occurred either early, when entering the phonetic information into memory, or later, when accessing the stored phonetic information for the purpose of production. In other words, errors in nonword repetition might suggest representational and perceptual inefficiencies and not necessarily phonological memory deficits. Similar findings were reported by Edwards and Lahey (1998) when they compared the repetitions of 54 children with SLI and those of their peers in terms of number and type of errors as well as latency and duration of response. The results suggested that the differences between the two groups might lie in either the formation or the storage of phonological representations in working memory, but that they were not connected with the SLI children's inability to hold phonological information in working memory.

2.5 Intrinsic Factor II—Phonological Sensitivity

2.5.1 The Nature of Phonological Sensitivity

2.5.1.i Definition

The term phonological sensitivity is often used interchangeably with phonological awareness, the latter term being better known and more widely used. Stahl and Murray (1998) defined phonological awareness on the basis of its form of presentation and processes. According to their view, phonological awareness is the awareness of sounds in spoken words, in contrast to phonics in written words. Such awareness of spoken words can be reflected in abilities such as rhyming, matching initial consonants, and counting the number of phonemes in words. McBride-Chang (1995) has broken down the construct of phonological awareness into three components—cognitive ability, memory, and simple speech perception, adding that the basis of phonological awareness is the manipulation of speech segments. The ability to analyse spoken language into smaller component sound units and further manipulate them is viewed as conscious and explicit, constituting a form of metalinguistic knowledge (Cheung, 1999).

Stanovich (1992) proposed to differentiate phonological awareness from phonological sensitivity on the basis of its required phonemic analysis skills. According to him, *phonological awareness* requires the explicit and intentional isolation of phonemic units and is assessed using tasks like phoneme deletion, phoneme isolation, or phoneme segmentation. *Phonological sensitivity*, in contrast, is assessed using phonological judgment tasks that do not require the explicit isolation of phonemes. Bowey (1994, 2002) preferred the term phonological sensitivity to phonological awareness to avoid notions of consciousness inherent in the use of the latter term. But the dividing line is not always so clear (Bowey, 1997, 2001).

Other researchers use phonological sensitivity to refer to the global set of phonological processing abilities that display a hierarchy of sensitivity to different levels of phonological complexity in different cognitive operations (Adams, 1990; Burgess & Lonigan, 1998). Phonological sensitivity, according to their definition, is a continuum, ranging from the analysis of rudimentary levels (i.e., rhyme and alliteration) at one end to the manipulation of higher levels (i.e., phonemic awareness) at the other.

Although there exists this variety of definitions of the term phonological sensitivity, it is used interchangeably with phonological awareness throughout this work. Although most of the phonological processing tasks described in the present study did not require the deliberate isolation of phonemes, it is very likely that the young learners were conscious of their deliberate manipulation of speech sounds while performing them.

2.5.1.ii Developmental Perspective of Phonological Sensitivity

Tasks designed to measure phonological sensitivity require subjects to analyse and break up a word into its smaller sound units. Goswami and Bryant (1990) suggested that there were three different kinds of phonological awareness—syllable awareness, onset and rime awareness, and phoneme awareness. The three kinds of phonological awareness correspond to three of the four levels of *linguistic awareness* proposed by

Adams (1990), which serve as prerequisites to gaining phonological awareness. Adams suggested that the lowest of the three levels is where children gain word awareness so that they can segment words in an utterance. She further theorised that the various tasks which are used to measure phonological awareness fall into five *levels of difficulty*. The first two levels are concerned with rhymes or alliteration in which subjects compare and contrast the large units in words. The other three levels deal with small phonological units, i.e., phonemes, in which subjects need to blend, add, delete, or move any designated phoneme.

The developmental theory of phonological processing abilities stipulates that awareness of large phonological units, i.e., syllable and onset-rime, is present very early because many words very young children encounter are monosyllabic, whereas their awareness of the links between individual sounds in monosyllabic words and alphabetic letters, i.e., phonemic awareness, appears later (Bradley & Bryant, 1983; Goswami & Bryant, 1990; Treiman & Zukowski, 1991). According to this theory, the development of phonological skills goes from large units of onsets and rimes to small units of phonemes, while onset-rime is the most significant longitudinal predictor of children's reading performance. The finding of a developmental trend of children scoring higher on large units than on small units is also observed in other studies, where phonemic awareness is regarded as the best predictor of children's reading (Hulme et al., 2002; Muter, Hulme, Snowling, & Taylor, 1998). The developmental perspective is reflected not only in the phenomenon of a progression from large units to small units but also in the increase in age. Metsala (1999) administered three phonological awareness tasks to four- to six-year-olds—onset-rime blending for word and pseudoword stimuli, isolating the initial phoneme for word and pseudoword stimuli, and phoneme blending for word stimuli. The results revealed that older children performed better than young children in each of the tasks. Lonigan, Burgess, Anthony, and Barker (1998) administered six phonological sensitivity tasks to two- to five-year-olds, the scores on each of the tasks being correlated with the children's ages. The studies support the developmental perspective of phonological processing skills which observes better performance in larger phonological units and among older subjects.

2.5.1.iii Unitary or Independent Factors

The disagreement between the two camps of researchers who agree on the developmental perspective of phonological sensitivity abilities but disagree over which type of awareness, onset-rime or phoneme, is the best predictor of children's reading, centres on whether phonemic awareness is a consequence of learning to read. Those who disagree with the claim provide evidence that shows phonemic awareness to exist among prereaders in whom letter knowledge appears to serve as phonetic cues

(Bowey, 1994; Lundberg, 1991). In contrast, Bryant and Goswami (1987) suggested that these one-to-one grapheme-phoneme correspondence tasks, such as phoneme counting and phonemic segmentation, were too difficult for prereaders. The fact that most children cannot manage these tasks until they have begun to read indicates that a more developed phonemic awareness is rather the product than the cause of learning to read (Torgesen et al., 1999). Investigations of the claim that phonemic sensitivity is restricted to alphabetically literate people were extended to specific populations. Morais, Bertelson, Cary, and Alegria (1986) recruited a group of illiterate people and another group of ex-illiterates, to whom a battery of tests—rhyme detection, syllabic vowel deletion, and initial consonant deletion—were given. While both groups performed comparatively well on the first two tasks, the illiterate group was significantly inferior to the ex-illiterate group in the third task. The results were interpreted as supporting the claim that reading instruction was required for the development of an ability to analyse into phonemic segments. Despite the positive link between literacy and phonemic awareness, this argument has underlain another issue of the component construct of the phonological processing tasks.

The other major disagreement over phonological processing abilities is whether they are best described by one factor or two independent factors. This unitary or independent construct controversy is extended from the aforementioned debate on whether onset-rime or phoneme best predict children's later reading. Stahl and Murray (1994) administered four different types of phonological awareness task—blending, isolation, segmentation, and deletion—in conjunction with four written language measures—letter recognition, graded word or passage reading, environmental sign reading, and spelling—to 113 children. They found from the factor analysis that all four subtests loaded on one factor which accounted for the most variance, while the children's ability to manipulate onsets and rimes within syllables related most strongly to reading, once an adequate level of letter recognition had been achieved. Other studies have reported similar findings (Anthony et al., 2002; de Jong & van der Leij, 1999; Schatschneider, Francis, Foorman, & Fletcher, 1999). However, Muter et al. (1998), who regard phoneme as the best concurrent and longitudinal predictor of children's reading, found a distinct rhyming ability factor independent of segmentation ability. Similar findings of a two-factor model were reported in other studies (McDougall, Hulme, Ellis, & Monk, 1994; Muter & Snowling, 1998). The strong views of both camps derived from the quantitative research have made a consensus very unlikely.

It has been suggested that little standardisation within individual phonological awareness tasks and a lack of experimental control across phonological processing tasks might partly explain different results from different studies in which phonological awareness has been defined in a number of ways (McBride-Chang, 1995). Stahl and Murray (1998) indicated that these different processes required by various tasks explained the phenomenon that children appeared to acquire phonological awareness at different ages in different studies. In an attempt to establish the validity and reliability of the phonemic awareness measures, Yopp (1988) administered 10 different phonological awareness measures to a group of kindergarteners. The factor analysis grouped these tasks into two major categories, with phoneme blending, phoneme segmentation, phoneme counting, and sound isolation all loading highly on the first factor, which accounted for nearly 59% of the variance. Stahl and Murray (1994) tried to separate task difficulty from linguistic complexity, these being interwoven in phonological processing tasks. A comparison of the two loadings from the factor analysis suggested that the notion of levels of linguistic complexity appeared to be a better way of defining phonological awareness.

2.5.2 Phonological Sensitivity and Writing Systems

2.5.2.i First-Learned Logographic Writing System

It has been suggested that readers of a logographic writing system will either be greatly lacking in phonological awareness if they have not learned the alphabet, or that they simply have less phonological awareness compared to users of alphabetic writing systems (de Gelder, Vroomen, & Bertelson, 1993; Read et al., 1986).

Cheung (1999) attributed the deficiency in phonological awareness among users of a first-learned logographic writing script to a complete lack of sound cues, as reflected by grapheme-phoneme correspondences in alphabetic languages, in the logographic script. Read et al. (1986) conducted a study to investigate whether a knowledge of alphabetic orthography affected phonological awareness. A group of adults literate only in Chinese characters and another group familiar with Hanyu Pinyin, a system of spelling Chinese words with the English alphabet, were recruited. The latter group outperformed the former group when they were asked either to add or delete a single consonant at the beginning of a spoken syllable. Interestingly, given enough instruction and practice in segmentation skills, one participant from the nonalphabetic group improved significantly without alphabetic literacy.

Despite the phonetic and the semantic components in Chinese characters, the association between the form of a character and its pronunciation is neither straightforward nor transparent. Consequently, it appears that phonological awareness is inherent in phonologically driven languages even if the language is not alphabetic. Mann (1986) compared American with Japanese children in counting and deletion tests to investigate the effects of a lack of alphabetic knowledge on phonological awareness. Japanese children are given kana and kanji writing instruction in schools. Kanji is a Chinese-logogram script system and kana is another script system in which each symbol represents a syllable. The two groups performed similarly in counting the number of syllables and deleting the initial syllable in a word, but the young Japanese learners performed much less well than their American counterparts in the phoneme tasks. However, when Mann gave the same tasks to older Japanese children, aged nine to 10, it was found that the majority of this group of children, irrespective of whether they had been instructed in the English alphabet, were able to manipulate both syllables and phonemes. The author speculated that age and the experience of learning kana might have enhanced their awareness of both syllables and phonemes, since Japanese is a phonological orthography.

These results have suggested that explicit instruction in segmentation skills is required for a nonalphabetic language such as Chinese, because the acquisition of such skills depends upon alphabetic literacy in particular and not on literacy in general. In other words, an understanding of the grapheme-phoneme conversion rules is essential to literacy development in alphabetic languages among logographic users (Cheung, 1999; Read et al., 1986).

2.5.2.ii Transfer of Phonological Awareness Across Alphabetic Languages

While phonological sensitivity is likely to be absent in users of a logographic language, a positive cross-language transfer of phonological awareness is noticed in speakers of alphabetic languages. Cisero and Royer (1995) employed three tasks to test phonological awareness—rhyme detection, initial phoneme detection, and final phoneme detection—in both Spanish and English, to examine the awareness of rhyme, onset, and phoneme in kindergarteners and first-grade students on two occasions, five months apart from each other. The multiple regression analysis in the initial phoneme task showed that both native and second language performance at Time 1 significantly contributed to the prediction of second language performance at Time 2, a finding lending support to the cross-language transfer hypothesis over a hypothesis suggesting that phonological awareness skills were restricted to the language of experience.

Similar transfer effects were found between the two alphabetic languages of English and French. Comeau, Cormier, Grandmaison, and Lacroix (1999) recruited three grades of 122 English-speaking children in French immersion classes to investigate the relationship between phonological awareness and reading achievement in both languages. A battery of tests, among which were seven phonological awareness tasks where subjects were required to delete a consonant or a syllable in different positions in a word, along with a verbal working memory task, were administered in both languages. It was found that phonological awareness in the mother tongue was as strongly related to achievement in word decoding in the L1 as phonological awareness in the L2. The authors suggested that phonological awareness, rather than being language-specific, was a general cognitive mechanism used to manipulate sounds. The cognitive ability played a central role in processing auditory-phonological information for learning to read in languages with a phonologically-based writing system.

These results are consistent with the hypothesis that second language learning is founded on native language phonological-orthographic ability, with native language decoding skills being a good predictor of second language competency (Meschyan & Hernandez, 2002). Despite the association of phonological awareness with second language learning, bilingual children do not develop the ability more easily than monolingual children. Bialystok, Majumder, and Martin (2003) conducted a series of three experiments to examine whether a bilingual advantage existed in the course of development of phonological awareness. They found that phonological awareness was centrally implicated in children's acquisition of literacy, especially for alphabetic scripts, and that bilingual children did not develop the ability more easily than monolinguals.

2.5.3 Phonological Sensitivity and Literacy Development

2.5.3.i Phonological Sensitivity and Reading Ability

When examining phonological sensitivity, reading is strictly defined at word level skills, i.e., word recognition or word/nonword reading. Reading success at the word level is the manifestation of a child's knowledge of grapheme-phoneme correspondences and his ability to recode written symbols into a sound-based representational system to get from the written word to its lexical referent (Wagner & Torgesen, 1987).

It is suggested that a child's phonological sensitivity develops before the beginning of formal reading instruction and that it predicts his reading acquisition. In a longitudinal study, Bradley and Bryant (1983) asked preschool children to detect the odd word which did not share a common phoneme in a string of words. High correlations were found between the children's initial sound categorisation scores and their reading and spelling scores over three years later. Multiple regression analyses established that these relationships remained strong even after intellectual levels and memory differences were partialled out. Similar findings were observed in other studies on preschool children (Burgess & Lonigan, 1998; Muter & Snowling, 1998; Gottardo, Stanovich, & Siegel, 1996; Stahl & Murray, 1994).

Phonological sensitivity remains a strong concurrent predictor of reading after children start learning to read. McDougall et al. (1994) administered a battery of assessments, including memory span and phonological awareness measures, to seven- to nine-year-olds to investigate the role of short-term memory and phonological skills in the processes of learning to read. Phoneme awareness was found to be a stronger predictor of differences in reading abilities among these young learners. Past research on specific populations has demonstrated that literate adults who have not learned an alphabetic writing script and illiterate adults who use an alphabetic language both show different patterns from literate alphabetic language speakers in their phonemic awareness tests (Morais et al., 1986; Read et al., 1986). It is suggested that children with higher levels of phonological sensitivity are more able to understand the significance of letters within printed words and that those with lower levels are less able. Bowey (1994) administered a battery of tasks to preschool children to investigate the relationship between their phonological sensitivity and alphabetic literacy. The children were divided into nonreaders and novice readers who could identify at least one word in a test of high frequency words from an early reading test. The nonreaders with a high letter knowledge performed differentially from those with a low letter knowledge on all four phonological oddity tasks. In addition, the novice readers scored higher than nonreaders with an equivalent letter knowledge on phonological sensitivity tasks—subsyllabic oddity and phoneme identity. Collectively, the results show that phonological sensitivity and reading are reciprocally related (Adams, 1990; Goswami & Bryant, 1990). There is no doubt that alphabetic knowledge plays a central role in facilitating phonological processing abilities.

However, the close association between phonological sensitivity and reading skills is most clearly revealed by those who demonstrate reading difficulties, among whom phonological processing deficits have gradually emerged as the underlying core problem (Siegel, 1998). Bowey, Cain, and Ryan (1992) recruited two groups of skilled and less skilled fourth-grade readers and another group of second-graders who were matched with the less skilled fourth-graders on word identification performance. They found that less skilled fourth-grade readers performed less well than both chronological age and reading-level controls on the phonological oddity and pseudoword reading tests. Correlational analyses were consistent with the view that phonological analysis skills made a greater contribution than verbal working memory skills to children's decoding abilities.

2.5.3.ii Phonological Sensitivity and Vocabulary Development

A large body of research has proved the intimate relationships between phonological sensitivity (or awareness) and the processes of learning to read (for a review, see Brady & Shankweiler, 1991; Metsala & Ehri, 1998). Very few of them, however, focused on a direct investigation of the association between phonological sensitivity and vocabulary development. Gathercole et al. (1991) used measures of both nonword repetition and rhyme awareness to investigate their link with reading and receptive vocabulary development among four- and five-year-old native speaking children. Despite the result that rhyme awareness scores were strongly related to the scores of one reading test at both Ages 4 and 5, they were not significantly associated with receptive vocabulary knowledge at either age. The result was surprising in that its factor analyses showed that measures of phonological memory and rhyme awareness did share a common phonological processing component.

In an attempt to clarify the interrelationship between phonological memory, phonological sensitivity, and receptive vocabulary in five-year-olds, Bowey (1996) replicated Gathercole et al. (1991) and found that similar patterns of association did exist between phonological sensitivity and receptive vocabulary, as well as between phonological memory and receptive vocabulary. Phonological sensitivity, in addition, explained a similar amount of independent variance in receptive vocabulary as did phonological memory after age, performance IQ effects, and the other phonological processing ability were statistically controlled. Furthermore, factor analyses revealed that all performance measures loaded on a single general ability factor, supporting the latent phonological processing factor theory. Studies by Metsala (1999) and Bowey (2001) also employed both nonword repetition and phonological sensitivity tasks to investigate their link with receptive vocabulary. Both studies revealed the same finding that nonword repetition was no more strongly correlated with receptive vocabulary than was phonological sensitivity.

While the aforementioned studies empirically established the association between phonological sensitivity and receptive vocabulary knowledge, they were restricted to explaining the specified association only. Among these studies, receptive vocabulary knowledge was tested using either the British Picture Vocabulary Scale or the Peabody Picture Vocabulary Test-Revised. This kind of vocabulary assessment asks a young child to identify, by pointing, which of four pictures corresponds to the word spoken by the test administrator, while the score interprets the child testee's receptive vocabulary knowledge, i.e., current vocabulary knowledge.

de Jong, Seveke, and van Veen (2000) decided to look beyond the empirically established association and investigate whether phonological sensitivity was associated with young children's ability to learn new words. The Dutch children were given a battery of assessments, which comprised intelligence tests, receptive vocabulary, letter knowledge, phonological sensitivity, nonword repetition, and two paired-associate word learning tasks. One of the word learning tasks required children to learn the familiar real names of four cuddly toys and the other required them to learn unfamiliar nonnames of another four cuddly toys. It was found that phonological sensitivity still contributed to the learning of novel nonnames even when effects of age, nonverbal intelligence, vocabulary, and letter knowledge were controlled, in stark contrast to its dissociation with current vocabulary knowledge or with the learning of familiar names. Phonological memory, reflected by nonword repetition, was related to both receptive vocabulary knowledge and the learning of nonnames, but the relationships disappeared when nonverbal intelligence or letter knowledge was controlled. The results overall suggested that phonological sensitivity might be more important than phonological memory in young children's learning of new words.

Since the investigation into the association between phonological processing abilities and vocabulary development began, the importance of the role played by phonological working memory has long been supported not only empirically (Gathercole & Baddeley, 1989; Gathercole et al., 1997) but also theoretically (Baddeley, 1986; Baddeley et al., 1998; Gathercole & Baddeley, 1993). In comparison, phonological sensitivity was more empirically substantiated (Bowey, 1996, 2001; de Jong et al., 2000; Metsala, 1999) until Metsala and Walley (1998) proposed the Lexical Restructuring Model, which makes a connection between vocabulary development and phonemically based representations.

According to the Lexical Restructuring Model, young children recognise words in a more holistic manner when the size of their vocabulary is still small. However, the rapidly growing vocabulary soon creates a need for them to represent lexical items in a more phonologically segmental way so that they can efficiently encode, store, and retrieve phonological information. As children's vocabulary expands, their underlying phonological representational structure is reorganised and becomes more segmental and adult-like, which enables them to recognise words from a partial speech input. In the lexical restructuring account, a child's lexical restructuring is affected by the familiarity status and sound-similarity relations among individual words in the child's lexicon. Familiar items become phonemically specified sooner than relatively unfamiliar words. Words in dense phonological neighbourhoods are also phonemically represented earlier in development than are words from sparse neighbourhoods (Metsala & Walley, 1998, p. 100-101). Consequently, the similarly sounding words *big, bag, bug, big, bit, dig, and wig* have their phonological representations phonemically developed earlier than those of the word *girl.*

The lexical restructuring account has provided a theoretical background in support of the link between vocabulary development and phonological sensitivity. On one hand, it explains that a more fine-grained representation of sounds emerges in the wake of vocabulary growth. On the other hand, it suggests the possibility that those children who are able to identify acoustic differences, i.e., who are in possession of a more segmental representation of sounds, will process speech signals more accurately, and this facilitates their word learning (Čeponienė, Service, Kurjenluoma, Cheour, & Näätänen, 1999).

2.5.4 Phonological Sensitivity Studies in an EFL Context

In contrast to the large body of phonological sensitivity research on native English-speaking children, similar studies on EFL children with a first-learned logographic language background are scarce, while those on the association between phonological sensitivity and vocabulary learning are even rarer. In studying learners' vocabulary learning strategies, Cheng (2005) discovered that junior high school students with the highest achievement level in vocabulary learning were able to apply the alphabetic rules and divide words into syllables. By contrast, the students with the lowest achievement level paid less attention to the sound of a word. Liang (2005) also indicated that college students who were instructed to enhance their phonological awareness managed to improve their vocabulary learning skills.

In young EFL learners' development of phonological sensitivity skills, Lin (2005) proposed that English speech perception benefited from longer exposure to the target language among sixth-graders, who developed the auditory discrimination skills of rhyming, blending, and segmentation, this being their order of linguistic difficulty. Yang (2006) revealed that onset and rime was the English phonological sensitivity skill that fourth-graders developed earliest. She also pointed out that the correlation between the subjects' phonological sensitivity and the results of the pseudoword tests was stronger than that between phonological sensitivity and real words, indicating that phonological sensitivity and pseudoword repetition performance were likely to underlie a similar component of phonological processing ability.

In terms of the association between phonological sensitivity and literacy development, most of the related EFL studies focused on word recognition ability, i.e., the grapheme-phoneme correspondence rules (Chen, 2003; Chien, 2002). Chien revealed that elementary school children's sound manipulation of spoken words, measured by a battery of phonemic awareness, syllable awareness, and onset-rime tasks, greatly contributed to their early literacy development in an EFL context. Among the three kinds of phonological processing task, phonemic awareness was found to be the best predictor of the children's single-word spelling and sentence reading, which was consistent with the findings of L1 studies. Sun (2002) also indicated that a significant correlation existed between phonological awareness and word recognition among the elementary school fifth-graders. However, more research is required into the association between EFL vocabulary learning and phonological processing skills.

2.6 Summary

In this chapter, the relationship between vocabulary learning and the two extrinsic factors was first examined. L1 translation equivalents have been found to facilitate L2 vocabulary learning in a large body of bilingual lexico-semantic memory research, supporting the assumption that the two languages of bilingual speakers are interconnected with each other at the lexical level in the initial stage. In terms of English instruction length, the longer the learners are instructed in the target language, the larger their vocabulary size. In comparison, the two intrinsic factors have been proved to be crucial to vocabulary learning in child learners in both native and EFL contexts. No study, however, to the best knowledge of the researcher, has attempted to compare their power to predict vocabulary learning in children learning English as a foreign language who possess a first-learned logographic language, although a few studies have tried to compare the two phonological processing skills in native English-speaking children and children learning English as a second language who have an alphabetic first-language background. The next chapter will introduce the experimental measures employed in the present study to measure phonological memory and phonological sensitivity skills and will also describe the explicit vocabulary teaching embedded in English storytelling.

CHAPTER 3: METHODOLOGY

As the purposes of the present study were to investigate individual differences among young EFL learners in their acquisition of vocabulary and to examine the association between these differences and extrinsic and intrinsic factors, assessments involving a variety of components, such as nonword repetition and phonological sensitivity tasks, were administered. The task results of the two phonological processing skills as well as the learners' length of English instruction were treated as independent variables while the scores of their vocabulary learning activities served as dependent variables. Vocabulary assessments were conducted using two different formats, since this study also aimed to trace how young learners stored and accessed L2 vocabulary.

This study is hence highly quantitative because many tests were employed to assess the subjects' vocabulary learning outcome and phonological processing skills. The choice of such assessments was made based on the accumulated knowledge on the research into relationships between vocabulary development and phonological processing skills. The paradigm of nonword repetition devised by Gathercole and Baddeley (1993) was adopted for the research on the aspect of phonological memory, while a developmental perspective of phonological awareness from large to small sound units proposed by Goswami and Bryant (1990) served as the framework in researching young learners' phonological sensitivity skills.

This chapter describes the subjects of the study, the research design and instruments, and the procedures and methods of data analysis, and also presents a summary of the pilot study.

3.1 Subjects

3.1.1 General Description

A total of 64 pupils in the fourth grade were recruited from a suburban elementary school in Tainan County, southern Taiwan, for the main study. All of them participated in the storytelling programme, but only 63 of them were able to take part in the series of assessments. One boy from the control group was deemed unsuitable for the assessments owing to his severe learning difficulties. All the other subjects were normally developing children with no reported history of hearing or speaking difficulties. It is hence likely that the results obtained from this sample of subjects can be generalised to their counterparts of the same school year in a bigger population because English is introduced to the national elementary level curriculum in the third year and only limited English is taught in the EFL context before they finish the elementary level education.

One of the two classes was labelled the experimental group and the other the control group, distinguished by the pedagogical difference when implementing the English storytelling programme. The experimental group was instructed only in

English, while the control group was given Chinese translation equivalents to the English lexical items. The experimental group included 32 participants and had an equal distribution of gender. The control group consisted of 15 boys and 16 girls. All subjects were aged from 9 years to 9 years 11 months at the beginning of the experiment and their mean chronological age was 9 years 5 months.

3.1.2 English Experience

The English curriculum started at the elementary school when they entered the third grade. The subjects had two sessions of English per week, each session lasting 40 minutes. The English textbooks for the third and fourth grades, published by a local publisher in Taiwan, were equipped with a CD supplement. Forty-two lexical items were listed in the two textbooks used in their third grade. These language beginners were encouraged to become familiar with the English alphabet but were not expected to memorise how to spell words in their third grade. The medium of instruction included both Mandarin Chinese and English, with a higher proportion of instruction being given in Chinese. The major source of language input in the classroom was from the language teacher, who is a nonnative English speaker. No official assessments were required of the pupils in their third year, but a mark was given to each pupil at the end of that school year.

According to their self-reports of English experience, up to 38% of the pupils had attended a bilingual kindergarten where English was in fact minimally taught or used. One boy from the experimental group had attended an elementary school where English was largely used as the medium of instruction in the English class before he transferred to the elementary school where the experiment for the present study was carried out. None of them had visited an English-speaking country, however.

Attending extracurricular English lessons became more common among the pupils after an English curriculum was introduced in their third year. Although less than a quarter of the 63 subjects had attended private English lessons in their first and second grades, 51% of them stated that they did so in their third year. Extracurricular English lessons among pupils in the experimental group ranged from zero to a maximum of 72 months (M = 20.97, SD = 21.64) and in the control group from zero to a maximum of 65 months (M = 22.84, SD = 22.97), indicating a very wide difference of extra English input among the subjects. Extracurricular English input included the time the pupils spent studying at a bilingual kindergarten and at private English institutions from their first to third elementary school years. The calculation was made in the unit of months because extracurricular English lessons could last from one to four hours a week, depending on the private institutions which offered the lessons.

The length of English instruction, on the other hand, included the time pupils had spent at a bilingual kindergarten, at private English institutions in their first and second years, and a 12-month count for their English curriculum in the third year of the public elementary school. English instruction length was hence either longer than or the same as the length of extracurricular input, because all pupils had English lessons in their third year. The experimental group had a mean of 29.41 months (*SD* = 17.65) of English instruction and the control group 30.26 (*SD* = 21.01) months.

3.1.3 Language Use

Apart from two subjects who understood but did not speak Taiwanese, all subjects used both Taiwanese and Mandarin Chinese in their daily lives. Up to 60% of them used a mixture of both languages when speaking to the elderly at home. When speaking to siblings and neighbours, around half of them used Mandarin Chinese. Mandarin Chinese was overwhelmingly chosen as the dominant language they used with teachers and classmates at school.

3.1.4 Attitudes Towards Four Skills

Around 40% of the participants felt that speaking was the language skill that it was most important for them to master while another 25% of them felt that listening was the second most important skill, suggesting that speaking and listening were most relevant to the young learners in the initial stage of their English learning. In terms of difficulty level, writing was chosen as the easiest skill and reading the most difficult. It is probable that writing was rated the easiest because it simply referred to copying the alphabet or new lexical items in the initial stage of their English learning. Learning to read had, however, obviously started to pose difficulties to the young learners since it was rated as the most difficult skill.

3.2 Materials and Design

In addition to the questionnaire, which was administered to elicit background information about the subjects, a battery of assessments was administered to investigate their phonological processing abilities and English vocabulary learning outcome (see Table 1).

Table 1

Measures and Purposes for the Present Study

Measure	Purpose
Questionnaire	To elicit background information about the young learners.
Nonword repetition task (sound-recorded)	To investigate the association between young learners' phonological memory and their vocabulary learning.
Phonological sensitivity assessment (two tasks sound-recorded)	To investigate the association between young learners' phonological sensitivity and their vocabulary learning.
Storytelling programme (video-recorded)	To explicitly teach lexical items and to enrich language input and to investigate the effect of vocabulary teaching with and without L1 translations.
Vocabulary assessment (one sound-recorded)	To examine young learners' vocabulary learning and to investigate its relationships with different measures.

The assessments included two nonword repetition tasks, five phonological sensitivity assessments in the Chinese and English languages, two online vocabulary tests, 13 written vocabulary tests, and one productive vocabulary test. The results from the nonword repetition and phonological sensitivity tasks were used to investigate whether there was a correlation between the young learners' phonological processing skills and their vocabulary learning, while the online vocabulary tests were used to examine whether instruction in the target language would facilitate faster mapping between an object and its L2 phonological form.

3.2.1 Questionnaire

A survey was administered to the subjects to elicit information regarding their English language experience, language use at home and at school, attitudes towards the four English language skills, and their familiarity with computer use. The questionnaire (see Appendix A) was divided into four parts. Part A was designed to find out if the participants had any sources of English input other than from the compulsory elementary education, and to determine the length of their extracurricular English lessons if they did. Information regarding the length of the subjects' English instruction and extracurricular English input was gathered from the survey. The subjects were also asked to indicate if they had lived in any English-speaking country or received education with English as the medium. In Part B, the participants had to indicate the languages they used with different speakers at home and at school respectively. This was in order to determine their habits of language use. In Part C, they were asked to choose which English language skills they thought to be the most important, the most difficult, the easiest, and which they wanted to excel at. These answers reflected their attitudes towards the four skills in the initial stage of learning English as a foreign language. At the end of the questionnaire, they were requested to specify how long they had been using a computer in order to indicate their IT skills, as several phonological sensitivity tasks and vocabulary assessments were conducted online with the mouse as the answering tool.

3.2.2 Nonword Repetition Tasks

Twenty English nonsense words (see Appendix B) from Gathercole's (1995) Children's Test of Nonword Repetition which were rated most wordlike by native English speakers were adopted as test items, while an additional five nonwords were used as practice items. This choice was made on the basis of the considerations that the level of wordlikeness would not affect young EFL learners as much as it did native child speakers of English (Gathercole, 1995; Gathercole et al., 1997) and that the number of question items, i.e., nonwords to be repeated, in this task should not far exceed that of the other phonological processing measures designed to probe phonological sensitivity. Each phonological sensitivity task contained 25 question items. In addition, the nonword repetition task was designed to give the young learners the impression that they were repeating real words they had never heard before. Consequently, only half of the original 40 nonwords were selected for use in the present study.

These 20 nonwords varied in syllable count, phoneme number, and consonant complexity. In terms of syllable count, there were five two-syllable, seven three-syllable, three four-syllable, and five five-syllable nonwords. The phoneme numbers ranged from five to 13, but there was no nonword with 10 phonemes. Eleven of the nonsense words had single consonants and the other nine had consonant clusters. Three out of the nine nonwords had more than one consonant cluster.

The English nonwords were rearranged in order according to a principle of increasing complexity in terms of syllable count and phoneme number. For example, *rubid* was placed at the beginning of the test items because it contained two syllables, five phonemes, and no consonant clusters, while *versatrationist* was at the end of the test list because it had five syllables, 13 phonemes, and two consonant clusters. This arrangement was made to avoid frustration on the part of the young learners. The list was then read at intervals of five seconds and recorded onto a Compaq Presario 2800 laptop.

3.2.3 Phonological Sensitivity Assessments

Assessments of phonological sensitivity skills were developed in both Chinese and English with five tasks for each language. They included (a) rhyme detection, (b) head detection, (c) rhyme and head detection, (d) rhyme and head production, and (e) initial sound/consonant isolation. Rhymes constitute units of a vowel and a consonant in the final position of a word while heads comprise a consonant in the initial position and a vowel (Kessler & Treiman, 2001). In the first four tasks, each question contained a pair of word items but in the initial sound/consonant isolation task there was only one word in each question.

In terms of response methods, the five tasks were divided into two categories: detection and production. The subjects were required to click the mouse to indicate their answers in the three detection tasks, but had to utter a rhyme or a head as designated by the task requirement in the two production tasks. All five assessments were constructed separately but task (c) rhyme and head detection and task (d) rhyme and head production shared the same question items. Real words and nonsense words made of pseudosounds were employed in the Chinese assessments, while only real words were used in the English ones.

3.2.3.i Construction of Chinese Phonological Sensitivity Assessments

A Mandarin Chinese pseudosound chart was constructed with reference to a

Chinese-language dictionary. Possible but nonoccurring sounds were used to form Chinese nonwords. The sounds in the pseudosound chart obey the phonological rules of the Chinese language but do not exist in the language. The syllable structure in Mandarin Chinese is CV(N) or GV(N) where C stands for consonant, G glide, V vowel, and N nasal. There are no consonant blends in Mandarin Chinese either before or after the vowel. In CVC/GVC syllables, only [n] or [ŋ] is allowed in the coda (Hanley, Tzeng, & Huang, 1999). The test items employed the high level tone only. Items were treated as nonsense words if they shared the same phoneme and syllable structure with meaningful words but carried no lexical meaning when their tone was switched to the high level. There are four tones in Mandarin Chinese and they function lexically.

The CVC/GVC syllable structure was fixed in the construction of the test items (see Appendixes C.i, C.ii, C.iii, & C.iv) and each word comprised three phonetic symbols from the Zhu-Yin-Fu-Hao script, which is used to mark the sounds of Chinese characters when Taiwanese children are taught to learn to read Chinese. The main purpose of the Chinese assessments was to act as a prelude to the English tasks, in which each English word was comprised of three phonemes, resembling the structure of three Zhu-Yin-Fu-Hao phonetic symbols for a Chinese character.

Words and nonwords were paired and developed into three different types of question item: a word-word pair, a nonword-nonword pair, and a word-nonword pair. Twenty question items were constructed as demonstration and practice items and another 20 as test items in each task. The test items were read at intervals of three seconds and recorded onto a cassette tape. A ring sounded to inform the subjects of the upcoming new question item.

3.2.3.ii Construction of English Phonological Sensitivity Assessments

To build an analogy between the Chinese items and the English items, the CVC syllable structure was adopted in the construction of the English items, where only real words were used. Sounds which do not exist in Mandarin Chinese were still included in the English assessments. For each task, 10 question items were constructed for demonstration and explanation, another five as practice items, and another 25 for each assessment (see Appendixes D.i, D.ii, D.iii, & D.iv). No words were repeated in a single task.

Owing to the different methods of administering phonological sensitivity tasks in the two languages, the English items were read and recorded onto a Compaq Presario 2800 laptop. Scripts were written using WordPad to be run in the DMDX programme, which is a display system designed to measure subjects' reaction times to visual and auditory stimuli. A picture featuring an object, such as an apple, popped up on the monitor of the laptop to give warning of the upcoming of a pair of English phonological cues before each pair was played to the subjects through the earphones. A different picture was used for each task to avoid confusion with other tasks on the part of participants. The pictures carried no meaning identical to any words used in the tests.

In the first three tasks, question items were set to play automatically at intervals of 10 seconds and subjects were required to click the mouse to indicate whether a word pair shared the rhyme or the head as designated by each task. The timing started as soon as the second member of a pair was played and stopped immediately once subjects clicked the mouse to indicate their answer. Accuracy and response time were automatically recorded on the laptop. The remaining presentation time was hence cut off and the first member of a word pair for the next question was played following the notifying picture. In the other two tasks, subjects had to give the answers, i.e., rhyme, head, or phoneme, to the questions, which were played at intervals of six to eight seconds.

3.2.4 Validity and Reliability

3.2.4.i Validity

Referring to whether the content of the manifest variables, i.e., the items of the tasks, is adequate to measure the latent concept (Muijs, 2004), the content validity of nonword repetition and phonological sensitivity tasks has been supported by an extensive body of studies. As phonological memory is reflected by an individual's ability to process and store verbally coded information, digit span or word span are often used to measure it. The two traditional phonological memory measures, however, are criticised for their similarity with the to-be-remembered lexical material. Nonword repetition is hence suggested since it has the advantage of detachment from the lexicality effects. The validity of using nonword repetition to measure phonological memory has been supported by extensive research on native child speakers (Brown & Hulme, 1996; de Jong et al., 2000; Gathercole & Baddeley, 1989; Gathercole et al., 1997; Hansen & Bowey, 1994), in children and adults learning foreign languages (Dufva & Voeten, 1999; Papagno & Vallar, 1995; Service, 1992; Service & Kohonen, 1995), and in neuropsychological patients (Vallar & Baddeley, 1984; Baddeley et al., 1988).

As phonological sensitivity refers to an individual's ability to reflect explicitly upon the sound structure of spoken words, the validity of this measure is supported by the studies which probe large-unit and small-unit phonological skills via the manipulation of speech segments (Bradley & Bryant, 1983; Lonigan et al., 1998; McBride-Chang, 1995; Shatil & Share, 2003; Stahl & Murray, 1994; Yopp, 1988).

3.2.4.ii Reliability

A series of internal reliability analyses was carried out on the two nonword repetition tasks and on all phonological sensitivity measures. In the phonological memory tasks, Cronbach α values were .55 for the first nonword repetition task and .33 for the second task, based on 20 items. In the English phonological sensitivity tasks, Cronbach α values were .72 for rhyme detection, .64 for head detection, .83 for rhyme and head detection, .93 for rhyme and head production, and .94 for initial consonant isolation, all based on 25 items in each individual task, and .95 for all 125 items. In the Chinese phonological sensitivity tasks, Cronbach α values were .49 for rhyme detection, .61 for head detection, .79 for rhyme and head detection, .83 for rhyme and head production, and .63 for initial sound isolation, all based on 20 items in each Chinese task, and .88 for all 100 items.

3.2.5 English Storytelling Programme

It has been suggested that explicit vocabulary instructions result in significant gains in learners' vocabulary knowledge (Beck, Perfetti, & McKeown, 1982; Paribakht & Wesche, 1997). With vocabulary teaching embedded within it, the storytelling programme was installed as a supplement to enrich the English input of the participants (Wellhousen, 1993). Listening to stories constitutes a significant source of vocabulary acquisition in L1 learning, whether or not it is accompanied by teacher explanation of word meanings, and the learning effects will double when brief explanations of word meanings are given (Brett, Rothlein, & Hurley, 1996; Elley, 1989). But it is necessary to point out that vocabulary learning in this present study is for young learners to learn names for things, in which case their acquisition of spoken form of the words should take place long before their acquisition of word meanings (Cameron, 2001).

The programme was carried out with both groups and three stories were told in total. The first story was titled *Silly Willy*, the second *King Big Wig*, and the third *Lots of Hearts* (see Appendixes E.i, E.ii, & E.iii). The big books, published by a local publisher, were accompanied by a CD and 24 flashcards each the size of B5 paper. The lexical items taught to the young learners were chosen mainly from the flashcards, which carry an English word on one side and the corresponding picture on the other. The target vocabulary items were taught with flashcards as teaching aids, i.e., the picture association method. It is suggested that this method closely resembles a common form of L1 vocabulary acquisition and yields better learning results because young children can associate words with environmental events (de Groot & van Hell, 2005; Wimer & Lambert, 1959).

A vocabulary pretest was administered before every story was told. The words, such as apple, book, nose, and fish, were removed from later explicit teaching when more than two thirds of the participants had matched them correctly on the pretest. A few words were then added to increase the number of words in every story to more than 20 but fewer than 24. The additional words could be found in the illustrations of the storybooks. Flashcards featuring these added-on words were made by the researcher for vocabulary teaching. The storyteller was a local English teacher who had studied TESOL in the United States and ran English storytelling workshops. She was informed of the purpose of the study in advance and given an outline of how the storytelling sessions should proceed.

3.2.6 Assessments of Vocabulary from Stories

Two different formats were developed to assess the young learners' acquisition of receptive vocabulary (see Table 2). The first of these consisted of the conventional paper-and-pencil tests which were used in most of the vocabulary assessments, including the pretests, posttests, and follow-up tests for the three stories, as well as an additional vocabulary assessment taking place near the end of the first story session for the second and third stories. This format was also applied to the word review tasks. The series of written vocabulary assessments were administered to the subjects at different times, at least one week to 10 days apart from each other, so that they measured the pupils' increase in as well as retention of new vocabulary items. The second format consisted of computerised vocabulary tests which were administered once for the second and third stories in the present study. The online tasks measured not only the participants' reaction times to visual and auditory cues but also their accuracy in matching the two cues. In addition to assessing receptive vocabulary, a productive assessment was also administered to all participants to investigate the gap between their receptive and productive vocabularies.

Table 2

Formats and Implementation Methods of Vocabulary Assessments for the Main Study

Type	Format	Assessment	Administration
Receptive vocabulary	Story 1		
	Paper-and-pencil	Pretest	Whole class
	Paper-and-pencil	Posttest	Whole class
	Paper-and-pencil	Follow-up test	Whole class
	Story 2		
	Paper-and-pencil	Pretest	Whole class
	Paper-and-pencil	In-class test	Whole class
	Paper-and-pencil	Posttest	Whole class
	Paper-and-pencil	Follow-up test	Whole class
	Computerised	Online test	Individual
	Story 3		
	Paper-and-pencil	Pretest	Whole class
	Paper-and-pencil	In-class test	Whole class
	Paper-and-pencil	Posttest	Whole class
	Paper-and-pencil	Follow-up test	Whole class
	Computerised	Online test	Individual
	All 3 stories		
	Paper-and-pencil	Word Review II	Whole class
Productive vocabulary	Computerised	Picture-naming	Individual

3.2.6.i Construction of Paper-and-Pencil Vocabulary Assessments

Words explicitly taught in the storytelling programme were selected as test items.

The 20-strong vocabulary items from each story were divided into three parts so that the young learners could quickly make their choice among six to eight pictures to match the auditory cue (see Appendixes F.i, F.ii, & F.iii). Instead of words, pictures were used in both written and computerised tests to avoid any complication created by the different levels of the participants' developing reading skills, which depended principally on their extracurricular English learning experience since they were not yet required to memorise word spelling in their school education. An auditory form which phonologically presents a picture with its L2 name is more to be recommended with young children in the initial stages of learning to read (de Groot & van Hell, 2005). Generating the correct sound structure from visually presented words would have been too great a cognitive challenge for them.

The pretests were always carried out before a new story was told. They were administered to establish a set of lexical items unknown to the subjects. The posttests were used to explore the young learners' vocabulary learning, while the follow-up tests were designed to examine their retention. In the second and third stories, an in-class paper-and-pencil assessment and a computerised assessment were added. The purpose of the in-class vocabulary tests was to investigate whether young learners could acquire vocabulary in a single storytelling session. The scripts for the vocabulary tests were written, and each test item was given a number. The items on the scripts were then read by the storyteller at intervals of three seconds and recorded onto a cassette tape. Listening to the tape, pupils were told to write down below a corresponding picture on the exam sheets the number preceding a lexical item. All vocabulary tests for a story contained the same test items.

3.2.6.ii Construction of Computerised Vocabulary Assessments

A script was written for each computerised vocabulary assessment to be run in the DMDX system (see Appendixes G.i & G.ii). The words on the script were read by the storyteller and recorded onto the Compaq Presario 2800 laptop. Most of the lexical items appearing in the computerised tests were taken from the stories, but some of them were adopted from outside the stories because they had semantic associations with or were minimal pairs to the target vocabulary. For example, the word *wag* was heard when a picture of a *wig* was shown to the participants and *ears* was heard when the picture of a *nose* popped up on the monitor. Subjects were supposed to reject the words because the auditory cue did not match the visual cue. A few other words which were selected for neither of the two reasons aforementioned were also used to see whether the subjects would reject them more quickly because they rang no bell in their vocabulary inventory.

A total of 30 questions were constructed, consisting of four practice trials and 26 question items. An auditory cue appeared simultaneously with the visual picture and the timing started as soon as the picture appeared on the monitor. The timing loop stopped as soon as subjects pressed the mouse to send a signal to the laptop. Accuracy and reaction times were automatically recorded on the laptop. A time-out feature was used in the online tests which meant that the next question would automatically appear after six seconds even if the subjects did not respond. Subjects were allowed six seconds to right-click or left-click the mouse to indicate their answer to each question item. Immediately after the assessments, participants were asked in a short interview if any Chinese translations to the English words they had just heard in the online tasks had occurred to them.

The drawback to administering computerised assessments was that they were time-consuming, because such assessments had to be conducted individually. On the other hand, the computerised tests were capable of obtaining more information from the subjects than the paper-and-pencil tests could reveal. The computerised assessments were able to measure how long it took the young learners to match an object with its phonological form in the target language. The reaction times were used to determine whether an input of Chinese translation of the explicitly taught lexical items would make any difference to their processing of these new words. Measures of reaction times indicate the processing complexity, and the time involved in coming to a decision may reflect the difficulty of the question item to the subject

(Cook, 1990; McDonald, 2000). Computerised assessments provide "a controlled form of comprehension because the choice of A and B is always the same, the situation and the time available never vary, and the scoring is automatic" (Cook, 1990, p. 577).

3.2.7 Review Assessments of Words

Three word review assessments in two different formats were carried out, which included two paper-and-pencil word review tests and one productive vocabulary task. One of the two written word review tests assessed what the children had learned from their previous school year and the other assessed how many words they learned from the storytelling programme. The productive vocabulary task measured how many words from the three stories the young learners could utter.

3.2.7.i Word Review I: Assessment for School-Adopted Textbooks

The words to be tested in the first word review assessment were taken from the two English textbooks the participants had used in their previous school year (see Appendix H.i). Divided into seven parts, the 42 words were read at three-second intervals and recorded onto a cassette tape. Listening to the tape, subjects had to write down the number preceding a lexical item below a corresponding picture on the exam sheet, all pictures having been taken from the textbooks. The assessment was administered as an achievement test to check their learning outcome from the previous year.

3.2.7.ii Word Review II: Assessment for the Three Stories

The other written review assessment was designed to measure how many words the learners had learned and retained four weeks after the storytelling programme. A total of 54 words were selected and divided into seven sections, each section containing seven to eight lexical items (see Appendix H.ii). Words from different stories were rearranged and those with semantic associations were put together in one section. Before the administration of the assessment, a five-minute word review session took place every morning for eight days to help pupils recycle the vocabulary. The words were read at three-second intervals and recorded onto a cassette tape by the researcher. The assessment took place towards the end of the experiment.

3.2.7.iii Productive Vocabulary Assessment: Picture-Naming

The productive vocabulary task required the subjects to name the pictures for the words they had learned from the storytelling programme. A different script was written for each subject to be run in the DMDX display system (see Appendix H.iii for a sample script). The number of words subjects had to name in the task ranged from three to all 54 words, depending on the result of their second word review assessment. The subjects were given six seconds to respond to each picture and were tested only on the correct answers they had given in the word review assessment.

While carrying out the task, they were asked to wear a headphone set equipped with a microphone so that they would speak their answers into the microphone. Three to four practice trials preceded the test in order for them to become acquainted with the new task. Their answers were automatically recorded onto the Compaq Presario 2800 laptop as sound files. All participants were tested individually.

3.3 Procedures

3.3.1 Questionnaire

The questionnaire was administered prior to all other measures. Each question item on the questionnaire was read out by the researcher and questions raised by the pupils were answered immediately before proceeding to the next question item, in order to avoid any misunderstanding. Each copy of the questionnaire was later checked and clarified with the pupils if any question item had been left unanswered. The survey was conducted with each group separately in their home classroom.

3.3.2 Nonword Repetition Tasks

The nonword repetition task was carried out twice during the three-month study. The first testing was carried out prior to the launch of the storytelling programme to avoid any influence from the suddenly intensive English input. A timetable was devised and posted on the bulletin board in the classrooms of the two groups to notify the pupils of their testing slots. As soon as the participants were seated, the following instructions were read to them, "You are about to hear some English words that you have never learned before. The words vary in length. The task required of you is to repeat each word after you have heard it. A ringing sound will be heard before the word is played. You are free to skip it if you have difficulty repeating it". They were then asked to put on a pair of headphones and repeat five practice items. During the break between playing the practice and test items, the participants were asked whether the volume was appropriate and if they had any questions. The volume was adjusted if a request was made. They were told in advance that there would be 20 repetitions to make and that they would not be able to stop in the middle. Each nonword was played once only.

The repetition task was conducted in the control room of the audio-visual auditorium of the primary school. The task, including the time it took to give instructions to the pupils, took about eight minutes, and was conducted during breaks between classes. Participants were tested individually and their repetition performances were recorded onto the computer and scored immediately. The participants were not corrected if they made mistakes or did not repeat successfully. At the end of the task, they were not told of their scores. A second testing was repeated five weeks after the first one to see if the participants' repetition performance had improved as a result of the intensive English input from the storytelling. The same instructions were repeated.

3.3.3 Phonological Sensitivity Assessments

3.3.3.i Chinese Phonological Sensitivity Assessments

Chinese assessments were always carried out as precursors to the English ones. The former were administered to the whole class during class time and each task, including explanations of the demonstration and practice items, lasted around 25 minutes. In each task, at least three question items were first used to demonstrate how to tackle the question items. Another three to five were used to check the participants' comprehension before moving on to the 10 practice items. Responses and corrections were given to the subjects to help them clarify their reflections when tackling the practice items. An answer sheet was then distributed and an assessment was administered immediately after the practice questions were finished. All test items were played once only.

Detection tasks

In the rhyme detection task, pupils were told to tick "Yes" on the answer sheets when the last two of the three phonemes in a pair of words were the same, and "No" when they were different. In the head detection task, they were supposed to detect the first two phonemes and tick "Yes" or "No" to indicate their answers. The rhyme and head detection task was a combination of the previous two tasks in which pupils were told to tick in one of the three columns marked "The same first two phonemes", "Not rhymed", or "The same last two phonemes", to indicate whether a pair of items shared the same rhyme, did not share the same phonemes, or shared the same head in the designated positions.

Production tasks

The subjects had to demonstrate their phonological segmentation skills in the two production tasks by writing down their answers in the Zhu-Yin-Fu-Hao phonetic symbols. In the rhyme and head production task, pupils listened to the same cassette tape played in the previous rhyme and head detection task, but were asked to write down on the answer sheets the shared phonemes or to tick in the 'Not rhymed' column. In the initial sound isolation task, the children were asked to write down the first sound of a word or a nonword in the Zhu-Yin-Fu-Hao script. Instructions on how to write down their answers in each task were written on the answer sheets and were also read out to the whole class before the assessments were administered. All of the Chinese assessments were conducted in the respective classroom of each group.

3.3.3.ii English Phonological Sensitivity Assessments

After administering each Chinese phonological sensitivity task, its English counterpart was immediately introduced to the class. Referring to the assessment they had just been given, the young learners were given at least three question items for the purposes of demonstration and explanation and another seven questions to check their comprehension. Answers and corrections were given immediately to each of the 10 questions. Such instructions as to how to detect rhymed phonemes were written down in advance to make sure that the two groups were given the same instructions.

In contrast to the Chinese phonological sensitivity tasks, the English tasks were implemented individually using the Compaq Presario 2800 laptop. All subjects had to wear a pair of headphones throughout the tests to receive a clear auditory input. Test instructions were written down in Chinese and read to subjects after they were seated for the task. They were instructed which side of a mouse to click to indicate their answers, with the mouse being placed on a piece of paper marked "Yes" and "No" in Chinese on its two corners as a reminder.

Five practice trials were first given to the subjects to check whether they understood what they were expected to do and correction was provided if they made mistakes. They were free to raise questions concerning the assessment they were about to do before proceeding to answer the 25 test questions. Explanations referring back to the Chinese assessments were sometimes made if the participants had any questions. In administering the second English phonological sensitivity task, i.e., the head detection task, the same test instructions were given, while another reminder on A4 paper was placed on the keyboard of the laptop to remind the subjects that they were to take a different task to detect if the first two phonemes were the same rather than the last two phonemes. In the third task, a different reminder about using the mouse was positioned to remind subjects that they had to press the left button of the mouse for a head, the right button for a rhyme, and the wheel when there was neither a head nor a rhyme.

When administering the production tasks, the mouse was removed out of the subjects' reach and a pair of headphones equipped with a microphone was used so that the subjects spoke their answers into the microphone. Their answers were recorded on the laptop and simultaneously transcribed by the researcher onto a score sheet using the International Phonetic Alphabet (IPA) symbols. In the initial consonant isolation task, the subjects were reminded that only one word would be heard and that they had to speak the first phoneme of the word into the microphone. The same device was used and their answers were also transcribed onto a score sheet using the IPA symbols.

It usually took three to five days for a group to finish one task. Each task, including explanations and instructions, took about eight to 10 minutes to complete. All subjects were tested individually during breaks in the control room of the audio-visual auditorium of the primary school and the scores were not disclosed to them. It was made clear to the participants that the results of the tasks would not be considered as part of their English grade.

3.3.4 English Storytelling Programme

The picture association method was adopted in the explicit teaching of vocabulary to the young learners. When introducing a word, the storyteller showed the corresponding picture on the flashcard, read the word aloud for the pupils to hear, then left the flashcard on the magnetic blackboard throughout the story sessions while illustrations in the big books were used to explain the storyline. All words were reviewed again five minutes before the class time was up. Pupils had at least two opportunities in a session to repeat the words orally, once at the beginning and once near the end of the 40-minute class. All English input was auditory and no written words were shown to the young learners in order to prevent pupils from using their extracurricular knowledge of English sound-letter correspondence to gain an advantage in their vocabulary learning (Rickard Liow & Poon, 1998).

The storytelling programme commenced after the completion of the first testing of nonword repetition with both the experimental and the control groups. Two sessions were devoted to each story and all story sessions were scheduled in the mornings, using the regular class time, with both groups hearing the same part of the stories on the same day. The first two stories were told in two successive weeks.

The first story was told on Monday and Thursday mornings. From the second story onwards, however, sessions were rescheduled for Monday and Wednesday mornings so that the computerised vocabulary assessments for subjects of the two groups could be completed within the same week. The third story followed the same schedule but was carried out a week apart from the second story owing to the interruption of a national holiday on the Monday. The same three stories were told to the two groups of young learners by the same storyteller, but the control group was explicitly given Chinese glosses for the target vocabulary items. Despite this pedagogical difference, however, the control group were allowed to repeat the new words in their L2 phonological forms only and were given no opportunity to repeat them in their L1. Classroom instructions were given in English to both groups. All story sessions took place in each group's classroom and were videotaped.

The seating arrangements were different in the two groups. The pupils of the experimental group who had a bigger classroom were divided into six groups with each group seated in a semicircle. The pupils of the control group were likewise divided into six groups but were seated in six straight lines, all facing the blackboard. During the storytelling sessions, the home teacher of each class was also in the classroom.

3.3.5 Vocabulary Assessments

3.3.5.i Paper-and-Pencil Vocabulary Assessments

The same procedures were adopted throughout all paper-and-pencil vocabulary assessments. The young learners were told to listen to the tape playing the test items and to write down the number preceding an English word on the answer sheet under the pictures of the words taught in the storytelling sessions. The first paper-and-pencil vocabulary assessment administered to the young learners was a review of words they had learned from two English textbooks in the previous school year: Word Review I. The pretest for Story 1 took place on the same day as the word review assessment. The posttest was administered the next day after the story had been completed. The follow-up test was held two weeks later after the completion of the storytelling.

The pretest for Story 2 was conducted together with the posttest for the first story. A new feature was introduced to the series of paper-and-pencil assessments commencing with the second story. An in-class vocabulary assessment was administered five minutes before the end of the first story session to investigate whether vocabulary gains could be made in a single session. The written posttest was not administered until almost two weeks after the story had been completed. The follow-up test took place 22 days after the completion of the storytelling.

The pretest for Story 3 was conducted on the same day as the follow-up test for the first story. An in-class assessment was administered at the end of the first story session. The posttest took place almost two weeks later and the follow-up test three weeks after the storytelling had been completed. One week after the follow-up test, the pupils were helped to review all the words in four days, and were given a review assessment of these words—Word Review II—on the fifth day.

All paper-and-pencil vocabulary assessments were administered to the whole class of the two groups on the same day, but were conducted in each group's respective classroom using regular class time. The participants were repeatedly reminded before every written assessment that the results would not be passed on to the school and that they should not make guesses if they did not know the answers. The answer sheets were collected immediately for scoring after each vocabulary assessment was finished.

3.3.5.ii Computerised Vocabulary Assessments

The online vocabulary assessments commenced on the same day, after a story had been completed. Although the subjects had to press either the right or the left button of the mouse to indicate their answers as in the phonological sensitivity detection tasks, they were reminded that the vocabulary assessments were different from the phonological sensitivity tasks. The online vocabulary tests had richer visual cues since a different picture appeared in each question, while the subjects had to decide whether the visual cue matched the auditory cue. Four practice trials were held before proceeding to the test items. During the break between practice and test items, the volume of the speaker was adjusted if the subjects requested it. But no feedback was given to them regarding their answers. The online tests, which started on the same day, after the second session of a story was completed, usually lasted three days because subjects were tested individually.

The computerised tests were conducted in the control room of the audio-visual auditorium of the primary school. As soon as the subjects had completed the online assessment, they were given a list on which pictures depicting the target words were displayed and asked to point out the pictures whose Chinese meanings had occurred to them while they were taking the online tests. Their answers were then recorded by the researcher. The assessment took up to 10 minutes to complete, including instructions and the after-test interview.

3.3.5.iii Productive Vocabulary Assessment: Picture-Naming Task

The picture-naming task took place the day after Word Review II had been administered. Those pupils who scored lowest on Word Review II were given the productive test first, because it was feared that they might forget more words if their tests were delayed. The testing order of this task was not made random, since the results would be used as supporting evidence in the analysis of the online vocabulary tests, but were not to be further examined.

At the beginning of the test, pupils were told how many pictures they had to name before they put on the headphone and microphone set. They were given three to four practice items before proceeding to the test items, which varied from as few as three to as many as all 54 words. Testing time varied from person to person as a result. Incorrect answers were transcribed immediately by the researcher using IPA symbols, while a tick was used to indicate correct answers. The subjects were given a second chance and asked immediately after their completion of the productive vocabulary test if any words they had failed to name during the online task had come back to them. Correct answers at this point were taken into account when scoring. All participants were tested individually in the control room of the audio-visual auditorium of the school.

3.4 Summary of Data

3.4.1 Data Collection

Data collection for the present study commenced in early September 2005 and was completed at the end of November. All assessments were administered during breaks between classes or the 40-minute nap time after lunch on school days, and were held on the premises of the primary school. The data collected were divided into qualitative and quantitative data for convenience of explanation (see Table 3). There was a much higher proportion of quantitative data.

Table 3

Types, Contents, and Sources of Data Collected for the Main Study

Type	Content	Source	
Qualitative data	Speech data	PM assessment	Nonword repetition task 1
			Nonword repetition task 2
		English PS assessment	Rhyme & head production
			Initial consonant isolation
		Vocabulary assessment	Picture-naming
Quantitative data	Written data	Vocabulary assessment	Word Reviews I & II
			3 pretests
			2 in-class tests
			3 posttests
			3 follow-up tests
		Chinese PS assessment	Rhyme detection
			Head detection
			Rhyme & head detection
			Rhyme & head production
			Initial sound isolation
	Spoken data	PM assessment	Nonword repetition task 1
			Nonword repetition task 2
		English PS assessment	Rhyme & head production
			Initial consonant isolation
		Vocabulary assessment	Picture-naming
	Online data: RT	English PS assessment	Rhyme detection
			Head detection
			Rhyme & head detection
		Vocabulary	Stories 2 & 3

(Table 3 continues)

(Table 3 continued)

	Online data: Accuracy	assessment	Stories 2 & 3
	Questionnaire	Background information about subjects	English experience, including instruction length. Language use at home and school. Attitudes towards English skills. Familiarity with mouse use.

Note. PM = phonological memory; PS = phonological sensitivity.

3.4.1.i Qualitative Data

The utterances made by the participants in the nonword repetition, English phonological sensitivity, and productive vocabulary tasks were classified as both qualitative and quantitative data. They were regarded as qualitative data because these utterances were used to check the subjects' pronunciation and hence to determine whether their oral production was a mistake or simply caused by immaturities in their developing English phonological system. All aforementioned utterances were in English. The English phonological sensitivity tasks included rhyme and head production and initial consonant isolation. These utterances, however, also served the purposes of quantitative data for analysis and were hence quantified and transferred to the scores of the assessments, with correct repetitions and utterances being treated as correct answers in calculating scores. No attempt was made to further analyse the phonological features of the utterances.

3.4.1.ii Quantitative Data

The quantitative data consisted of written data, spoken data, online accuracy counts and reaction times, and data from two surveys. Written data included those from the Chinese phonological sensitivity tasks and the English vocabulary assessments. Speech data were the same as the qualitative data, collected from two nonword repetition tasks, two English phonological sensitivity tasks, and a picture-naming task. The reaction times were generated from the online tasks, which comprised two vocabulary assessments and three English phonological sensitivity tasks. In addition to the reaction times, the online vocabulary tasks also provided the accuracy counts, which represented the number of questions the participants had answered correctly. The questionnaire data included the subjects' background information.

These quantitative data were regarded as ordinal data, in terms of their level of measurement. As these tasks were administered to measure the participants' phonological memory, phonological sensitivity, and English vocabulary knowledge, higher scores represented higher levels of a construct. However, an increase from five to seven in word knowledge did not represent the same change as an increase from 12 to 14, because word knowledge in the present study was represented by the learning outcome, which was complicated by the lexical items they had learned previously. Despite the fact that they lacked the characteristics of interval data, these ordinal data still demonstrated that higher scores represented better word knowledge and lower scores poorer knowledge.

3.4.2 Scoring Principles

In evaluating the young learners' oral production, it was important to distinguish whether the production was a mistake or whether it was caused by immaturities in their developing English phonological system. In principle, any phonological inaccuracies reflecting consistent immaturities or simplification were credited as being correct. However, a distinction between phonemes /l/ and /r/ was required, and a substitution of either phoneme was regarded as an incorrect answer since the two phonemes are distinguishable in both Taiwanese and Mandarin Chinese. Nonword repetition was scored as a whole, with mistakes in a single phoneme or syllable in the repetition leading to its being treated as an incorrect answer.

All assessment was scored manually by the researcher except for the two online vocabulary tasks and the three English phonological sensitivity assessments. The accuracy of these was automatically recorded by the computer. For those tasks which were scored manually, one point was awarded for a correctly answered question and the total of points awarded for correct answers represented the score for the task.

3.4.3 Methods of Data Analysis

3.4.3.i Comparison Between Two Groups

The first research subquestion was designed to investigate whether the input of L1 would impede or facilitate the mapping of an object with its L2 phonological form. Independent-samples *t*-tests were used for the purpose of comparing the mean scores and reaction times of the experimental (n = 32) and control (n = 31) groups under two different experimental conditions distinguished by the input of Chinese glosses to target L2 vocabulary items with the control group. The data providing information regarding the effects of L1 input included the learners' accuracy in their vocabulary assessments and their reaction times in the online vocabulary tasks.

3.4.3.ii Relationships Between Factors and Vocabulary Knowledge

The second, third, and fourth subquestions were asked in order to investigate whether vocabulary learning was related to any of the three variables—English instruction length, phonological memory, and phonological sensitivity. Prior to carrying out any computation, all data from the two groups were collapsed (N = 63) and treated as a whole for later analyses. Scatterplots were first drawn to demonstrate that most data points appeared to fall within the vicinity of other points and that there were no obvious outliers. Pearson's correlation analyses were used to investigate the relationships between vocabulary and the three variables.

Both factor analysis and regression analysis were used to examine whether phonological memory and phonological sensitivity explained independently, or overlapped in their explanation of, any variation in the participants' vocabulary scores. The method of principal component analysis in the exploratory factor analysis was employed to establish whether a common component, i.e., factor, existed within the data pertaining to vocabulary learning and the two phonological processing abilities. By default in SPSS, only factors with Eigenvalues larger than one will be obtained, where the Eigenvalue measures the substantive importance of the variables with regard to a particular factor. The larger the Eigenvalue, the more representative of the data the factor is. The factor loadings of variables on a particular factor represent their relationships to the factor. A high factor loading indicates a close relationship between the variable and the factor. As factor analysis is used to establish whether all entered variables share the same component, it can be used to explain that the two phonological processing abilities might overlap with each other in explaining the variance in vocabulary scores when they were found to share the same factor.

The hierarchical entry method was used to investigate whether one of the two phonological processing skills made unique contributions to vocabulary learning. In the hierarchical multiple regression analysis, all the known predictors relevant to vocabulary learning, for example, age, L2 instruction length, Word Review I, and

Chinese phonological sensitivity scores, were first entered into the equation and the variance accounted for by the known relevant predictors was removed from further analysis. When one of the two phonological processing abilities—phonological memory and English phonological sensitivity—was next entered into the equation, the percentage point it obtained was the variance it uniquely explained. Alternating the two abilities in the entry to the equation on one hand indicated which ability explained more variance in a single assessment. On the other hand, it also showed whether one processing ability accounted for further variance in a vocabulary assessment when the variance accounted for by the other ability was removed. In other words, when different variables contributed differently to the participants' vocabulary scores, the hierarchical multiple regression analysis helped to establish how much more variance was explained by a newly entered variable.

While the correlation analyses helped to establish which variables were associated with vocabulary learning, and the hierarchical regression analyses explained which variables contributed uniquely to vocabulary scores, the stepwise method in regression analysis was used to search for the variable which best predicted the dependent variable, i.e., the vocabulary scores. The results from all the statistical analyses collectively answered the main question of the present study.

3.5 Pilot Study

3.5.1 Subjects

A pilot study was carried out at the beginning of September in the same primary school. Equally divided into two groups, ten pupils were recruited from two fourth-year classes different from those used for the main study. The experimental group, which received no Chinese input, contained two boys and three girls. The control group, which was given Chinese translations to the target vocabulary items only, consisted of three boys and two girls. Their mean chronological age at the start of the pilot study was nine years and six months.

3.5.2 Materials and Design

Similar materials were used in the pilot study. A survey was administered to determine the subjects' habits of language use, their attitudes towards general English learning, and their IT skills. None of the subjects had lived in an English-speaking country, although one subject from the control group had for a few months attended a school where English was used as the medium of instruction in the English class. All of the subjects were Taiwanese-Mandarin bilinguals.

A number of assessments were employed to explore their phonological processing skills, including a nonword repetition task and four phonological sensitivity tasks. The same set of 20 nonsense words which were used in the main study were adopted in the pilot study. The words were read and recorded onto a cassette tape. The four English phonological sensitivity assessments comprised tasks of (a) rhyme detection, (b) head detection, (c) rhyme and head production, and (d) initial consonant isolation. The items were constructed according to the same principles applied in the main study. In each task, 10 question items were constructed for practice and another 20 for testing. The 20 test items were read and recorded onto a cassette tape. Pairs of Chinese phonemes were constructed and used as examples to help the young learners to understand the purpose of the tasks. Only English phonological sensitivity assessments were administered in the pilot study. The participants' Chinese phonological sensitivity skills were not investigated.

An English storytelling programme was introduced to investigate whether the use of L1 would make differences to the young learners in terms of their vocabulary acquisition. The story *King Big Wig*, which was Story 2 in the main study, was told to the subjects of the pilot study. A total of four story sessions were held on four mornings. The two groups alternated every other day and each group had two story sessions.

Four vocabulary assessments—a pretest, a posttest, a follow-up test, and a computerised test—for the story were carried out to investigate the subjects' vocabulary acquisition. Twenty-four pictures, the English words for which were introduced in the storytelling, were used as the test items. The words were first read by the researcher and recorded onto a cassette tape, which was used in the pretest only. The words were later read and recorded by the storyteller onto a cassette tape and the Compaq Presario 2800 laptop again for two paper-and-pencil assessments and a computerised test respectively.

3.5.3 Procedures

The first assessment administered was Word Review I, followed by the nonword repetition task and the vocabulary pretest. The nonword repetition task was administered to participants individually while the others were conducted with all 10 participants at the same time.

The pretest was conducted on the Friday prior to the week commencing the storytelling programme in order to establish which words were unknown to the learners. The posttest was conducted one day after the completion of the story. The computerised vocabulary assessment, which had a different design from the paper-and-pencil tests, took place one week after the posttest. The follow-up test was administered two weeks after the completion of the story. The paper-and-pencil assessments were administered to all participants at the same time but the computerised test was administered individually.

The surveys, storytelling, and all assessments were carried out in the audio-visual auditorium of the primary school. The storytelling took place during the 30-minute study session in the morning after pupils had finished tidying up the campus and working on exercises assigned by their home teacher. The assessments were carried out for the most part during breaks.

3.5.4 Results

The nonparametric Mann-Whitney test was computed to obtain the mean and the standard deviation of each experimental measure employed and a summary was given (see Table 4).

Table 4

Group Differences for All Experimental Measures in the Pilot Study

Group / Measure	Experimental (n = 5)		Control (n = 5)		Mann-Whitney
	M	SD	M	SD	*z*
L2 instruction length	15.80	8.50	20.20	16.71	-0.643
Phonological processing capability					
Nonword repetition	13.20	1.64	13.00	2.45	-0.532
Rhyme detection	14.60	2.41	15.20	3.63	-0.745
Head detection	15.00	3.67	15.40	1.14	-0.212
Rhyme production	7.20	6.69	10.00	4.47	-0.973
Initial consonant isolation	13.40	2.07	13.20	2.17	0.000
Vocabulary assessment					
Word Review I	30.20	7.63	34.80	5.17	-0.955
Pretest	13.40	5.64	14.80	5.12	-0.314
Posttest	19.80	2.95	20.80	3.03	-0.529
Follow-up test	18.20	4.66	20.80	3.90	-1.803
Online test	19.40	0.89	22.20	0.84	-2.660**
RT (in ms)	1992.65	518.17	1897.72	292.73	-0.104

Note. L2 instruction length was measured in the unit of month.

RT was measured in milliseconds.
$^{**}p < .01$

3.5.4.i. Vocabulary Assessments

It is obvious from Table 4 that the control group scored higher in all four written assessments. The group proved in the pretest and later in the posttest that they knew more words and had also learned more words from the story. Even in the follow-up test which was carried out two weeks after the story was told, they were able to retain more words than the experimental group, and consequently showed a lesser degree of forgetting. However, no significant differences were observed between the two groups regarding their written vocabulary assessments. The only significant difference observed between the two groups was demonstrated in the online vocabulary test. A nonparametric Mann-Whitney test showed that the control group had given more correct answers than the experimental group, $z = -2.660$, $p < .01$, (two-tailed).

3.5.4.ii Reaction Times

When the mean reaction times of the groups were compared, the control group again showed an advantage in the online task. Their reaction times were found to be faster ($M = 1897.72$, $SD = 292.73$) when their correct answers were compared with those of the experimental group ($M = 1992.65$, $SD = 518.17$). A nonparametric

Mann-Whitney test showed that the difference was nonsignificant, $z = -.104$, $p > .05$, (two-tailed).

3.5.4.iii Correlations Between Different Measures

Spearman correlation analyses were used to investigate the association between different measures because of the small sample size. A series of Spearman correlation analyses were run to investigate the link between the vocabulary assessment results and the subjects' English instruction length. A significant positive correlation was found between instruction length and the posttest results, $rho = .725$, $p < .05$, (two-tailed). No significant correlation was found between the variable and any of the phonological processing skills. Nor was any association established between phonological processing skills and the subjects' vocabulary learning.

Stepwise regression analyses were run and the four independent variables of English instruction length, phonological memory, phonological sensitivity, and Word Review I were entered into the regression model. Word Review I, which represented the subjects' existing English achievement, was chosen as the best predictor of their vocabulary pretest results. No variable was chosen as a predictor in the other vocabulary assessments.

3.5.4.iv Conclusion

It was temporarily concluded from the pilot study that a use of the L1 in the learning of the foreign language was not statistically significant in the children's processing of auditory cues, even though the group receiving the L1 glosses did exhibit faster reaction times and a higher accuracy in vocabulary memorisation. It is not surprising that a positive association was found to exist between English instruction length and better vocabulary scores in the vocabulary posttest. However, it should be noted that English instruction length was not correlated with either of the phonological processing skills. This issue would be explored further in the main study.

Chinese phonological sensitivity was not examined in the pilot study even though the question items were used to help the subjects to get a grasp of the English phonological sensitivity questions. Chinese items would be expanded and administered as part of the phonological sensitivity assessments in the main study as precursors to their English counterparts. The results would be used to investigate whether there was a cross-language transfer of phonological processing skills.

3.6 Ethical Considerations

A letter of consent addressed to the parents of the participants in both the pilot (see Appendix I.i) and the main (see Appendix I.ii) studies was sent out and collected before the beginning of the experiment since the subjects were under 16 years of age. The consent letters stated the aim of the study and asked for the permission of the parents to record the classroom and assess their children. Parents of two pupils from the main study agreed to their children's participation in the study but refused to have them videotaped. A change of seating was consequently arranged by the home teacher to avoid filming them.

The other ethical consideration related to the research design was the use of real words, instead of nonwords, as the target vocabulary items in the storytelling programme. Although nonwords are sometimes used to replace target vocabulary items to control for the possibility that subjects might previously have learned or practised them in their free time, the study used real words for two reasons. On one hand, it was unlikely that the parents of the subjects would accept a proposal to allocate regular class hours to learning nonwords. On the other hand, this was not a laboratory experiment but a study conducted in language classrooms where English input was not strictly controlled and could have come from sources other than the language teacher, as in real life.

Anonymity was maintained in the present study by allocating each subject a code number, and the scores of each participant on all assessments were entered under the designated code name. The collected data and all test results were used for the purpose of the study only. With regard to ethical considerations, the researcher has followed the British Association for Applied Linguistics' ([BAAL], 2006)

Recommendations for Good Practice in Applied Linguists.

CHAPTER 4: RESULTS

In this chapter, the results of the tests on the young learners' vocabulary acquisition are presented, the focus being on examining the relationship between vocabulary acquisition and the factors under investigation in the present study. These factors are L1 translation, length of English instruction, and two aspects of phonological processing skills—phonological memory and phonological sensitivity. The first section presents whether an extra input of L1, i.e., a provision of L1 translation equivalents to target vocabulary items, had an effect on the young pupils' FL vocabulary learning. Independent-samples *t*-tests were used in the analyses to compare the mean differences between the experimental and the control groups in their online and written vocabulary assessments. A comparison of the two groups' accuracy in their vocabulary assessments reveals their learning outcome, while a comparison of their online reaction times shows how quickly they mapped an auditory cue onto the picture of an object. These results provide an explanation of how young learners learned, stored, and accessed vocabulary in the initial stage of EFL learning.

In the second section, the data pertaining to all 63 subjects are collapsed and the relationships between assessments of the same category are examined to demonstrate the consistency of the subjects' performances and hence the reliability of the assessments. Pearson's correlation analyses are used in the third and fourth sections.

The third section contains an examination of the effects of the learners' different lengths of English instruction on both their vocabulary performance and the two phonological processing skills. In the fourth section, a closer look is taken at the relationship between vocabulary learning and each of the two phonological processing skills. A factor analysis was used to determine whether the two phonological processing capabilities per se shared a component construct. Stepwise and hierarchical regression analyses were used in the fifth section to determine which factor and which task was the best predictor of young learners' EFL vocabulary performance.

4.1 Effects of the First Language

4.1.1 Starting Point of the Two Groups

An analysis to investigate whether the two groups of subjects had the same level of proficiency at the beginning of the study was carried out before any other analysis was conducted. The means of the six experimental measures from the two groups were compared using the independent-samples *t*-tests to determine the differences between them (see Table 5).

4.1.1.i Chronological Age

As expected, the chronological ages of the two fourth-grade groups had no statistical significance, $t(61) = -.38$, $p > .05$, (two-tailed), since all 63 subjects were born between September 1, 1995 and August 31, 1996. The school had placed pupils born between those 12 months randomly in different classes.

Table 5

Group Differences for Age, English Instruction Length, and Preexisting Vocabulary Knowledge

Measure \ Group		Experimental (n = 32) M	Experimental (n = 32) SD	Control (n = 31) M	Control (n = 31) SD	df	t
Chronological age		113.16	3.67	113.52	3.83	61	-0.38
L2 instruction length		29.41	17.65	30.26	21.01	61	-0.17
Vocabulary assessment							
Word Review I		28.88	8.94	24.19	11.92	61	1.77
Pretest	Story 1	10.91	4.07	10.77	5.28	61	0.11
	Story 2	14.34	4.43	12.13	3.74	61	2.14*
	Story 3	5.19	3.08	4.32	3.75	61	1.00

Note. Both chronological age and L2 instruction length were measured in the unit of month.

*$p < .05$.

4.1.1.ii Length of English Instruction

The row showing English instruction displays the mean length of English instruction each group had received, calculated in the unit of months. It is clear from the result that the majority of the pupils started learning the foreign language at least one year before the English curriculum was implemented in their third year, because the mean instruction length in both groups exceeded 12 months. The experimental group had a mean length of 29.41 months of English instruction, in comparison to the control group's 30.26 months. Despite the fact that the control group had a slightly longer instruction length, no significance ($t[61] = -.17, p > .05$, [two-tailed]) was found when their means were compared. However, the big gap between the early starters and their peers is demonstrated in the columns of standard deviation (SD), suggesting that big differences existed between individual participants and that the mean value was a less accurate representation of the data. This result shows why English instruction length should be considered as a crucial factor in explaining the performance of young learners in an EFL context.

4.1.1.iii Vocabulary Assessments

The four assessments revealed the vocabulary knowledge of the participants at the beginning of the study. The Word Review I served as an achievement test to compare the students' previous learning results. In contrast, the other three pretests were used to check the existing vocabulary knowledge of the young learners before each story was told.

In *Word Review I*, the experimental group achieved a higher score ($M = 28.88$, $SD = 8.94$) than the control group ($M = 24.19$, $SD = 11.92$). However, when their means were compared, the difference was not found to be significant ($t[61] = 1.77, p > .05$, [two-tailed]), indicating that the two groups were similar in terms of their learning outcome from the previous year. On testing their vocabulary knowledge for the target lexical items in *Story 1*, the experimental group (M = 10.91, SD = 4.07) scored similarly to the control group (M = 10.77, SD = 5.28). A t-test revealed no significant difference in the vocabulary knowledge of the two groups, $t(61) = .11$, $p > .05$, (two-tailed). However, a different picture was presented in *Story 2*. In the pretest, the experimental group scored higher (M = 14.34, SD = 4.43) than the control group (M = 12.13, SD = 3.74), and a t-test revealed that the two groups differed significantly, $t(61) = 2.14$, $p < .05$, (two-tailed). The result from Story 2 indicates that the former group knew more words than the latter group before the story was told. In contrast to the other two pretests, in the pretest for *Story 3*, the results indicated that both groups knew fewer of the words to be taught for Story 3. However, the experimental group (M = 5.19, SD = 3.08) still scored a little higher than the control group (M = 4.32, SD = 3.75), even though a t-test clearly shows that the between-group difference was nonsignificant, $t(61) = 1.00$, $p > .05$, (two-tailed).

4.1.1.iv Summary

Despite the statistically significant differences found in the pretest for Story 2, it is temporarily concluded that the two groups had similar English proficiency at the point when the study began. Care was taken, however, to monitor the results of other written assessments related to Story 2 to see if the gap continued. In addition, judging from the results of the three pretests, it appears that Story 3 contained the highest number of words which were unknown to the participants, followed by Story 1, with Story 2 containing the lowest number. The three stories could be graded with Story 3 as the most difficult and Story 2 the easiest, based on the existing vocabulary knowledge of the young learners.

4.1.2 The Variable of First Language

The major pedagogical difference in implementing the storytelling programme was an input of L1 glosses with the control group while the experimental group received no such input. The between-group differences reflected by vocabulary scores were consequently attributed to the provision of L1 translation equivalents for target vocabulary items, considering the fact that all other variables were controlled. Independent-samples t-tests were computed to compare the mean scores of the two groups in eight written and two online vocabulary tests, as well as their mean reaction times.

4.1.2.i Written Vocabulary Assessments

Comparison of overall vocabulary scores between groups

The overall scores of all written vocabulary tests taken by the two groups were compared using the independent-samples t-tests. The mean scores and t-test results showed which group performed better in these vocabulary assessments and whether the differences were significant (see Table 6). The pretest results, which have been discussed above, have been listed to make it possible to compare and contrast them with the results of subsequent vocabulary tests.

Table 6

Overall Score Differences for Written Vocabulary Assessments Between Groups Who Were or Were Not Provided L1 Glosses

Assessment \ Group	Experimental (n = 32)		Control (n = 31)			
	M	SD	M	SD	df	t
Story 1						
Pretest	10.91	4.07	10.77	5.28	61	0.11
Posttest	15.69	5.83	16.13	5.41	61	-0.31
Follow-up test	13.25	6.44	13.10	6.42	61	0.10
Story 2						
Pretest	14.34	4.43	12.13	3.74	61	2.14*
In-class test	14.13	4.93	14.13	4.61	61	-0.003
Posttest	15.66	4.74	15.00	4.21	61	0.58
Follow-up test	15.44	4.74	14.81	4.17	61	0.56
Story 3						
Pretest	5.19	3.08	4.32	3.75	61	1.00
In-class test	8.66	4.69	8.26	4.20	61	0.36
Posttest	10.22	5.17	9.87	4.84	61	0.28
Follow-up test	11.34	5.62	11.23	5.45	61	0.09

$*p < .05$

Story 1

Although the experimental group (M = 15.69, SD = 5.83) scored lower than the control group (M = 16.13, SD = 5.41) in the posttest, a t-test comparing the means indicates that the difference was nonsignificant, $t(61) = -.31, p > .05$, (two-tailed). In the follow-up test, the experimental group ($M = 13.25$, $SD = 6.44$) scored similarly to the control group ($M = 13.10$, $SD = 6.42$) and again the difference, not surprisingly, was nonsignificant, $t(61) = .10$, $p > .05$, (two-tailed). None of the three vocabulary assessments for Story 1 indicated any statistical difference between the two groups.

Story 2

Both groups were given a surprise test five minutes before the end of the first storytelling session. This in-class test interestingly shows that the two groups had an identical mean score ($M = 14.13$). The posttest administered about two weeks after the story was told indicated that although the experimental group ($M = 15.66$, $SD = 4.74$) scored slightly higher than the control group ($M = 15.00$, $SD = 4.21$), the difference was not statistically significant, $t(61) = .58$, $p >.05$, (two-tailed). In the follow-up test, which was given one week after the posttest, the experimental group ($M = 15.44$, $SD = 4.74$) performed better than the control group ($M = 14.81$, $SD = 4.17$), with the difference being nonsignificant, $t(61) = .56$, $p > .05$, (two-tailed). The significant between-group difference in the pretest disappeared in all the other three subsequent vocabulary assessments for Story 2.

Story 3

In the in-class vocabulary test, the two groups performed similarly ($t[61] = .36$, $p > .05$, [two-tailed]), with the experimental group ($M = 8.66$, $SD = 4.69$) only slightly better than the control group (M = 8.26, SD = 4.20). In the posttest administered nearly two weeks later, the two groups again did not perform differentially from one another, $t(61) = .28$, $p > .05$, (two-tailed). The follow-up test, conducted approximately one week after the posttest, yielded a similarly nonsignificant difference, $t(61) = .09$, $p > .05$, (two-tailed). The two groups hence performed similarly in all written vocabulary assessments for Story 3.

Summary of the written vocabulary assessments

Among the eight written assessments, the experimental group achieved higher scores in six of them. Only in the posttest for Story 1 did the control group score higher than the experimental group. However, none of the differences between the two groups had statistical significance, indicating that the two groups did not perform differentially. It was noteworthy that the significant difference observed in the pretest for Story 2 was not repeated in the three subsequent written assessments for that story, and this indicated that the control group had caught up with the experimental group in just one story session.

Furthermore, a larger standard deviation was observed in the experimental group on most vocabulary tests. By contrast, the control group had a larger standard deviation than the experimental group on only two vocabulary tests which were both pretests. This finding indicates, on one hand, that the scores among individual learners in the experimental group were more widely different from each other than they were in the control group. On the other hand, the result also implies that use of the L1 was likely to have made it easier for the students to see the connection between form and meaning in the foreign language, with the result that the scores of the subjects in the control group varied less widely.

4.1.2.ii Comparison of Vocabulary Gains Between Groups

In contrast to the overall vocabulary scores, vocabulary gains made by the two groups in each story were calculated by deducting the pretest score from that of a test administered later. Pupils who scored lower in the later test and hence had a negative score had their score converted to zero, indicating that they had made no progress in learning new vocabulary. Negative scores could not be used because they would cancel out the positive scores of other learners when a group mean value was computed. Independent-samples *t*-tests were then computed to compare the vocabulary gains made between the pretest and the other paper-and-pencil tests between the two groups. The results (see Table 7) show which group learned more lexical items and whether the differences between the groups were statistically significant.

Table 7

Group Differences for Vocabulary Gains Made Between Two Written Vocabulary Assessments as a Function of L1 Effects

Group / Vocabulary gains	Experimental (n = 32)		Control (n = 31)			
	M	SD	M	SD	df	t
Story 1						
Pretest to posttest	5.22	3.52	5.55	4.06	61	-0.35
Pretest to follow-up test	3.75	3.98	3.65	3.43	61	0.11
Story 2						
Pretest to in-class test	1.19	2.01	2.71	2.92	61	-2.42*
Pretest to posttest	2.03	1.89	3.55	3.38	61	-2.21*
Pretest to follow-up test	2.13	2.37	3.32	2.99	61	-1.77
Story 3						
Pretest to in-class test	3.84	3.09	4.00	2.54	61	-0.22
Pretest to posttest	5.22	3.54	5.77	4.11	61	-0.58
Pretest to follow-up test	6.25	4.53	7.06	4.49	61	-0.72

*$p < .05$

The results relating to vocabulary gains show that both groups successfully learned new words from each story and that they were able to retain these newly learned words for a period of time. At a glance, it is clear that the control group on average learned and retained more new words than did the experimental group on seven out of the eight counts of vocabulary gains between a pretest and a test administered later. However, when a series of independent-samples *t*-tests were computed to compare their means, significant differences were found on Story 2 only. The differences in vocabulary gains from the pretest to the in-class test ($t[61] = -2.42$, $p < .05$, [two-tailed]) and from the pretest to the posttest ($t[61] = -2.21$, $p < .05$, [two-tailed]) were statistically significant, with the control group consistently making more progress than the experimental group.

The results also show that the two groups had similar patterns of vocabulary learning but that the patterns of vocabulary gains among the three stories are different. Both groups made the most vocabulary gains on Stories 1 and 3 and the fewest gains on Story 2, the easiest one, despite the assumption, based on their scores in the pretests, that the three stories were of different levels of difficulty. However, the vocabulary gains made between pretests and follow-up tests in the first two stories showed a decline compared with those observed between pretests and posttests, demonstrating that the pupils gradually forgot new words when they did not encounter such words in their environment. In contrast to the follow-up tests for Stories 1 and 2, both groups scored higher in the follow-up test than in the posttest for Story 3. This probably occurred because a review of all the words from all three stories took place during the week when the follow-up test for Story 3 was administered. This indicates that a simple review refreshed the young learners' memory of the new vocabulary items.

4.1.2.iii Online Vocabulary Assessments

In addition to the paper-and-pencil vocabulary assessments, the subjects were asked to complete two online vocabulary tests. The results of the two online vocabulary assessments for Stories 2 and 3 were divided into the categories of accuracy and reaction times. Accuracy refers to the number of question items each group answered correctly. Reaction times demonstrate how quickly subjects responded to the online questions, which in turn reflects the complexity of their processing in mapping auditory and picture cues. Independent-samples *t*-tests were computed to compare the mean scores and reaction times between the two groups (see Table 8).

Table 8

Group Differences for Results of Accuracy and Reaction Times in Online Vocabulary Tests as a Function of L1 Effects

Group / Online test	Experimental (n = 32)		Control (n = 31)			
	M	SD	M	SD	df	t
Story 2						
Accuracy	19.41	3.63	19.68	2.61	61	-0.34
Overall RT	2336.54	454.58	2107.01	356.55	61	2.23*
Correct RT	2274.32	437.56	2047.90	343.31	61	2.28*
Incorrect RT	2630.18	709.91	2276.57	528.50	61	-2.24*
Story 3						
Accuracy	17.53	3.63	18.52	3.16	61	-1.15
Overall RT	2426.91	497.91	2196.25	330.01	61	2.16*
Correct RT	2382.52	492.68	2173.30	339.27	61	1.96
Incorrect RT	2601.43	734.93	2313.03	431.20	61	-1.89

Note. RT was measured in milliseconds.

The overall RT represents the times subjects spent on answering questions, irrespective of whether their responses were right or wrong. The correct RT refers to the reaction times to those questions which subjects answered correctly. The incorrect RT is the response times of subjects to questions to which they gave wrong answers.

$*p < .05$.

Accuracy

From the results shown in Table 8, it is evident that the control group scored higher than the experimental group in both online tasks, in contrast to most of the written assessments. On Story 2, the control group (M = 19.68, SD = 2.61) scored only slightly higher than the experimental group (M = 19.41, SD = 3.63). It is not surprising that the group difference was nonsignificant, $t(61) = -.34$, $p > .05$, (two-tailed). On Story 3, the control group (M = 18.52, SD = 3.16) answered on average one more question correctly than the experimental group (M = 17.53, SD = 3.63). However, when a t-test was computed to compare the means, no significant difference was found, $t(61) = -1.15$, $p > .05$, (two-tailed). The two groups performed in-differentially in terms of accuracy in the online vocabulary tasks.

Reaction times

The control group had significantly faster reaction times than the experimental group on the three counts of RT for Story 2. The result revealed that the control group (M = 2107.01, SD = 356.55) reacted more quickly than the experimental group (M = 2336.54, SD = 454.58) in terms of their overall RT, irrespective of right or wrong answers. An independent-samples t-test showed that the group difference was significant, $t(61) = 2.23$, $p < .05$, (two-tailed). The control group (M = 2047.90, SD = 343.31) was also faster in correctly accepting matched pictures and L2 phonological forms or rejecting unmatched ones than the experimental group (M = 2274.32, SD = 437.56), as shown in the correct RT. The mean difference of 226.42 ms was statistically significant between the two groups $t(61) = 2.28$, $p < .05$, (two-tailed). Even when they made wrong decisions, the control group (M = 2276.57, SD = 528.50) was consistently faster than the experimental group (M = 2630.18, SD = 709.91), and the difference was found to be significant, $t(61) = -2.24$, $p < .05$, (two-tailed).

The control group was again consistently faster than the experimental group in all three kinds of RT in the online test for Story 3. However, a significant difference was found only in the overall RT ($t[61] = 2.16$, $p < .05$, [two-tailed]) between the control group (M = 2196.25, SD = 330.01) and the experimental group (M = 2426.91, SD = 497.91), and not in the other two types of reaction time.

Considering the fact that the control group had consistently faster reaction times, an additional independent-samples t-test was conducted to compare the two groups' experience of IT skills and the group difference was found to be significant, $t(61) = -3.32$, $p < .001$, (two-tailed). According to their self-reports in the questionnaire, the control group had a mean of 22 months (SD = 5.56) of IT experience, compared to the experimental group's 15.41 months (SD = 9.59). Despite the significant between-group difference, the IT experience was believed not to have been a factor which affected the reaction time results because the act of clicking the mouse is purely mechanical.

It is worth noting that both groups were, in the two online vocabulary assessments, quicker at accepting those question items with matched pictures and L2 sounds than at rejecting the unmatched items. An independent-samples t-test was computed to compare the correct and incorrect RTs within each group. A statistically significant difference was found in the experimental group in Story 2 ($t[31] = -4.18$, $p < .05$, [two-tailed]) and Story 3 ($t[31] = -2.11$, $p < .05$, [two-tailed]). A similar significant response pattern was found in the control group in Story 2 ($t[30] = -3.33$, $p < .05$, [two-tailed]) and Story 3 ($t[30] = -2.25$, $p < .05$, [two-tailed]). The results indicate that participants were quicker when accepting and responding to matched question items than they were when they had to reject unmatched ones (Kroll & Potter, 1984).

4.1.2.iv Interview

A brief interview was conducted immediately after the online tasks to investigate whether the subjects had tried to produce L1 translations for picture or auditory cues during the online tasks. Over half of the respondents from the experimental group replied that Chinese translation had occurred to them during both online vocabulary assessments for Stories 2 and 3. When the two online tests were compared, a higher percentage of the subjects were found to have resorted to Chinese translation in Story

3. In contrast, nearly two thirds of the control group said they had mapped the cues with Chinese translation in Story 2 and nearly three quarters of them had done so in Story 3. The data obtained from the interviews clearly indicate that most of the young learners had used their L1 to help them memorise FL words.

4.1.2.v Productive Vocabulary Assessment

The result of the productive vocabulary assessment showed that the receptive vocabulary assessments—the paper-and-pencil and online assessments—were easier on the subjects than the task which involved production. In the productive vocabulary assessment, subjects were asked to name the pictures they had matched correctly in Word Review II. The maximal and minimal numbers of pictures successfully named were 40 and 0 (M = 11.75, SD = 9.48) by the experimental group and 30 and 1 by the control group (M = 11.06, SD = 8.22). The percentage rate of correct picture-naming was 39.4% in the experimental group and 37.5% in the control group. However, it should be pointed out that these figures were not accurate representations of group performance, and were hence unsuitable for making a group comparison, since the number of pictures each subject was asked to name varied widely.

4.1.3 Summary of L1 Effects

The experimental and control groups were shown to have overall similar vocabulary knowledge at the beginning of the study, despite the result that the experimental group appeared to have known significantly more target words in Story 2 than the control group. But the gap between the two groups was soon closed when both groups obtained the same mean score in the in-class test after one story session. The result in turn shows that the control group had learned more words with the aid of L1 translation of the target words. This result was further supported when vocabulary gains between the two groups were compared and it appeared that the control group had gained and retained more new words on almost all counts of vocabulary gains between a pretest and a later administered test.

In terms of accuracy for online vocabulary assessments, the control group scored higher than the experimental group, even though the group differences were nonsignificant in both stories. Faster reaction times, nevertheless, were consistently observed in the control group. However, in terms of correct RT, a significant difference was observed in Story 2 only. The results collectively suggest that L1 translation equivalents did not hamper the processing of FL vocabulary items and that, quite on the contrary, they might be a facilitating factor in enabling the young learners to recognise new words more quickly.

4.2 Relationships Between Experimental Measures

Since the results described in the previous section suggested that the two groups had similar vocabulary knowledge and learning outcomes, i.e., the two groups were on a similar footing, all data were hence collapsed and the mean value was computed for each experimental measure (see Table 9).

Table 9

Descriptive Statistics for All Experimental Measures (N = 63)

Measure		M	%	SD	Max. possible
Chronological age		113.33		3.72	
L2 instruction length		29.83		19.23	
Vocabulary assessment					
Word Review	I	26.57		10.69	42
	II	28.51		13.02	54
Story 1	Pretest	10.84		4.67	24
	Posttest	15.90		5.59	24
	Follow-up test	13.17		6.37	24
Story 2	Pretest	13.25		4.22	24
	In-class test	14.13		4.74	24
	Posttest	15.33		4.46	24
	Follow-up test	15.13		4.44	24
	Online test	19.54		3.15	26
	RT	2162.91		407.16	
Story 3	Pretest	4.76		3.43	21
	In-class test	8.46		4.42	21
	Posttest	10.05		4.98	21
	Follow-up test	11.29		5.49	21
	Online test	18.02		3.42	26
	RT	2279.57		433.79	
Phonological processing capability					
PM	Nonword repetition 1	11.83		2.65	20
	Nonword repetition 2	13.05		2.29	20

(Table 9 continues)

(Table 9 continued)

Chinese PS	Rhyme detection	16.70	83.49	2.17	20
	Head detection	15.25	76.27	2.83	20
	Rhyme & head detection	12.19	60.95	4.31	20
	Rhyme & head production	10.30	51.51	4.29	20
	Initial sound isolation	11.63	58.17	2.94	20
English PS	Rhyme detection	17.97	71.87	3.89	25
	RT	3314.15		1090.75	
	Head detection	17.17	68.70	3.60	25
	RT	3239.62		1001.57	
	Rhyme & head detection	16.59	66.35	4.99	25
	RT	3246.99		1151.02	
	Rhyme & head production	10.27	41.08	6.67	25
	Initial consonant isolation	13.19	52.76	7.47	25

Note. Both chronological age and L2 instruction length were measured in the unit of month.

PM = phonological memory; PS = phonological sensitivity.

After the computation of the experimental measure means, a series of Pearson's correlation analyses were conducted to investigate the relationships between these measures. The mean scores of measures of the same category and the associations between them are listed in Appendixes J.i to J.iii and explained as follows.

4.2.1 Vocabulary Assessments

Pearson's correlation analyses revealed that the subjects' performances in vocabulary assessments of the same category, i.e., Word Reviews I and II, were positively correlated with each other (see Appendix J.i). The moderate correlation between the two word review assessments, $r = .631, p < .01$, (two-tailed), suggested that the young learners who had achieved good vocabulary grades in their previous school year also performed well in learning vocabulary from the three stories.

The vocabulary assessments for each story were also correlated with each other (see Appendix J.i). The three written tests for Story 1 were associated with each other, with their r values ranging from .522 to .791, $p < .01$, (two-tailed). With regard to the four written tests and one online task for Story 2, these were all moderately correlated with each other, r ranging from .459 to .824, $p < .01$, (two-tailed). The same pattern of correlations was found with the vocabulary tests for Story 3, and their Pearson's correlation coefficients ranged from .292 to .848, $p < .05$, (two-tailed). These results indicate that the subjects were consistent in their vocabulary performance, which reflects the test-retest reliability of the vocabulary assessments.

However, better learners did not necessarily respond more quickly in the online tasks. Correlation analyses were computed for the two online tasks and their reaction times, but no correlation was found for Story 2 ($r = -.090$, $p > .05$, [two-tailed]) nor for Story 3 ($r = -.110$, $p > .05$, [two-tailed]).

4.2.2 Nonword Repetition

In the first and second nonword repetition tasks, which were administered five weeks apart from one another, the participants were found to perform better in the second task (M = 13.05, SD = 2.29) than in the first (M = 11.83, SD = 2.65). A paired-samples t-test showed that the difference was significant, $t(62) = -4.624$, $p < .001$, (two-tailed), and that the two scores were positively correlated, $r = .648$, $p < .001$, (two-tailed). The moderate correlation coefficient indicates that the subjects performed consistently in the nonword repetition tasks.

4.2.3 Chinese Phonological Sensitivity

The subjects performed comparably in the five Chinese phonological sensitivity tasks, but their scores were not so consistently associated with each other (see Appendix J.ii). In terms of the two single-layer detection tasks, the participants scored slightly higher on rhyme detection ($M = 16.70$, $SD = 2.17$) than on head detection ($M = 15.25$, $SD = 2.83$), and a paired-samples t-test showed that the difference was significant, $t(62) = 4.79$, $p < .001$, (two-tailed). Tasks involving production proved to be more challenging than detection tasks. The mean score for rhyme and head *production* ($M = 10.30$, $SD = 4.29$) was lower than that for rhyme and head *detection* ($M = 12.19$, $SD = 4.31$). A paired-samples t-test showed that the difference was significant ($t[62] = 3.515$, $p < .01$, [two-tailed]) and that performances in these two tasks were moderately correlated, $r = .508$, $p < .001$, (two-tailed). The initial sound isolation task ($M = 11.63$, $SD = 2.94$) produced the second lowest score among the five tasks, and its score was not associated with that of either the rhyme detection or the head detection task.

4.2.4 English Phonological Sensitivity

The participants' performance in the English tasks followed a similar pattern to that in the Chinese tasks, in the sense that they scored higher in the first three detection tasks and lower in the last two production tasks. However, unlike the Chinese tasks, their performances in all five English phonological sensitivity tasks were correlated with each other without exception (see Appendix J.iii). Although the participants scored similarly higher on English rhyme detection (M = 17.97, SD = 3.89) and lower on head detection (M = 17.17, SD = 3.60), a paired-sample t-test revealed no significant difference between the two scores, $t(62) = 1.869, p > .05$, (two-tailed). However, the two scores were positively correlated with each other, $r = .597, p < .001$, (two-tailed). In contrast to the two previous tasks, participants performed differentially on the rhyme and head *detection* (M = 16.59, SD = 4.99) and *production* (M = 10.27, SD = 6.67) tasks. A paired-samples t-test and Pearson's correlation analysis revealed that although the two task performances differed significantly from each other ($t[62]$ = 10.377, $p < .001$, [two-tailed]), they were still positively correlated, $r = .691, p < .001$, (two-tailed). The mean score for the initial consonant isolation task was 13.19 (SD = 7.47), the second lowest score among the five tasks. In terms of reaction times and task scores, Pearson correlation analyses revealed that a statistically significant difference only existed between head detection and the reaction times associated with it ($r = -.348$, $p < .01$, [two-tailed]), indicating that the participants who scored higher on this task responded more quickly.

4.2.5 Comparison of Phonological Sensitivity Between the Two Languages

When the same task type in the two languages was compared in the aspect of percentage rate of correct answers, it is clear that the participants performed better in Chinese than in English. Among these five types of task, they scored higher in four in their native language, except in the rhyme and head detection task, where they performed better in the English task. In the aspect of linguistic level irrespective of whether the task type was detection or production, the participants on average scored higher at the rhyme level than at the phoneme level, indicating that their rhyme awareness was better developed than their phonemic awareness. However, in terms of production tasks only, the young learners scored higher in the initial sound/consonant tasks in both languages than in the rhyme tasks. It was likely that the rhyme and head production tasks were more cognitively demanding because they involved the production of two linguistic measures—rhyme and head—in one single task.

4.3 Effects of Length of English Instruction

In this section, the effects of English instruction length on vocabulary knowledge and on the two phonological processing skills are examined. Pearson's correlation analyses were used to explore these relationships.

4.3.1 Association Between EFL Instruction Length and Vocabulary Scores

A series of Pearson's correlation analyses were computed to investigate the association between L2 instruction length of the subjects ($M = 29.83$, $SD = 19.23$) and the results of various vocabulary assessments (see Table 10). The assessments included 13 written as well as the two online tasks and the associated reaction times. Only the correct RTs were examined.

Table 10

Correlations Between Young Learners' English Instruction Length and Vocabulary Assessment Performances

Vocabulary Assessment		
Word Review	I	.383**
	II	.384**
Story 1	Pretest	.303*
	Posttest	.363**
	Follow-up test	.399**
Story 2	Pretest	.249*
	In-class test	.342**
	Posttest	.355**
	Follow-up test	.292*
	Online test	.237
	RT	-.101
Story 3	Pretest	.297*
	In-class test	.380**
	Posttest	.281*
	Follow-up test	.391**
	Online test	.130
	RT	-.314*

*$p < .05$. **$p < .01$.

4.3.1.i Written Assessments

A positive correlation was consistently observed between English instruction length and the scores in all 13 written vocabulary assessments, r ranging from .249 to .399. The findings indicated that the longer the length of English instruction, the higher the vocabulary scores. The advantage of longer L2 instruction was clearly demonstrated by its association with the scores of the three pretests. Explicit vocabulary teaching did not start until after the pretests. The scores obtained in the pretests were obviously a result of what the pupils had learned from their after-school English lessons, since English is learned as a foreign language in Taiwan and most input is from formal schooling either in school or in extracurricular lessons.

4.3.1.ii Online Assessments

In contrast to the written vocabulary assessments, the online vocabulary tasks presented a different picture, with none of the scores being associated with English instruction length. In terms of reaction times, a significant but negative correlation was observed in Story 3 ($r = -.314, p < .05$, [two-tailed]), indicating that the longer the length of the subjects' English input, the shorter their response times. This result indicated that L2 instruction length might have had an impact on the speed at which the participants processed the auditory cue of a new word. However, no significant difference was found in the reaction times for Story 2 ($r = -.101, p > .05$, [two-tailed]), suggesting the association between reaction time and instruction length was inconsistent.

4.3.2 Association Between EFL Instruction Length and Vocabulary Gains

It is clear that students who were exposed to English at an earlier age scored higher overall in the written vocabulary tests. However, it is not clear if there is an association between FL instruction length and learning speed, i.e., whether more words were learned when a learner had had a longer exposure to English. Pearson's correlation analyses were computed to probe these relationships (see Table 11), with vocabulary gains being calculated by deducting the pretest score from that of a later administered test. Again, negative scores obtained from the deduction were converted to zero before Pearson's correlation analyses were conducted.

Table 11

Means, Standard Deviations, and Correlations Between Young Learners' English Instruction Length and Vocabulary Gains Made Between Two Written Vocabulary Assessments

Vocabulary gains		M	SD	
Story 1	Pretest to posttest	5.38	3.77	.143
	Pretest to follow-up test	3.70	3.69	.162
Story 2	Pretest to in-class test	1.94	2.60	.060
	Pretest to posttest	2.78	2.81	.137
	Pretest to follow-up test	2.71	2.74	.008
Story 3	Pretest to in-class test	3.92	2.81	.219
	Pretest to posttest	5.49	3.81	.102
	Pretest to follow-up test	6.65	4.49	.229

The results show that FL instruction length and vocabulary gains were consistently unassociated with each other, indicating that the pupils who had received longer instruction in English did not gain more vocabulary items. It appears that the factor of English instruction length might have had no impact on the speed of young EFL learners' vocabulary acquisition.

4.3.3 Association Between EFL Instruction Length and Phonological Processing Abilities

The association between the variable and the young learners' phonological processing abilities was examined to determine whether longer English instruction produced better phonological memory and phonological sensitivity skills in the young EFL learners (see Table 12).

Table 12

Correlations Between Young Learners' English Instruction Length and Phonological Processing Capabilities

Phonological processing capability		
PM	Task 1	.155
	Task 2	.134
Chinese PS	Rhyme detection	.319*
	Head detection	.141
	Rhyme & head detection	.080
	Rhyme & head production	.086
	Initial sound isolation	-.153

(Table 12 continues)

(Table 12 continued)

English PS	Rhyme detection	.113
	RT	.044
	Head detection	.178
	RT	-.095
	Rhyme & head detection	-.064
	RT	-.016
	Rhyme & head production	.185
	Initial consonant isolation	.132

Note. PM = phonological memory; PS = phonological sensitivity.
* $p < .05$.

In contrast to the consistent correlations between L2 instruction length and the written vocabulary test scores, little association was found between the variable and the phonological processing measures. The only correlation obtained was with the Chinese rhyme detection task, $r = .319$, $p < .05$, (two-tailed). No correlation was found between the variable and the results of any other Chinese phonological sensitivity tasks. Similarly, the scores of neither of the two nonword repetition tasks nor of any English phonological sensitivity task were correlated with English instruction length. Generally speaking, the results show that English instruction had no effect on EFL young learners' phonological processing abilities.

4.3.4 Summary of Effects of English Instruction Length

In section 4.2, it was shown how the data pertaining to all 63 participants of the present study were collapsed in order to investigate their association with the other three factors, after the use of L1 translation was found not to make significant differences to the learning results of the experimental and control groups. But the factor of English instruction length was found to be consistently associated with the results of all written vocabulary assessments but not with those of the two online tests. This may be because the young learners had opportunities to correct their answers in the written tests but not in the online tests, so that the online test results were unrelated to the variable.

However, a further investigation of the link between vocabulary gains and EFL instruction length revealed that those pupils who had received longer English instruction did not necessarily gain more words, when they were given the same amount of instruction time as were the pupils who had had less English input. Also in contrast are the associations between FL instruction length and the two phonological processing abilities. The phonological processing skills of the subjects were not affected by length of instruction in English, as was evidenced by the dissociation between these factors. It is clear from these results that young EFL learners had a better vocabulary knowledge if they had begun learning the language at an earlier age. However, the factor of instruction length was not found to contribute to the two phonological processing skills, which are believed to facilitate young learners' vocabulary learning.

4.4 Young Learners' Phonological Processing Abilities

The data analysis presented in this section is divided into three parts. The first two parts focus on the association between the young learners' two phonological processing skills—phonological memory and phonological sensitivity—and their vocabulary scores, using Pearson's correlation as the method of data analysis. In the third part, the relationship between the two phonological processing skills per se is examined using the method of factor analysis. This method helps disclose whether the two phonological processing skills underlay the same component construct as indicated by previous research.

4.4.1 Association of Phonological Memory With Vocabulary Assessment Results

Pearson's correlation analyses were computed to determine the relationships between the two nonword repetition tasks and all vocabulary assessments as well as the reaction times (see Table 13).

Table 13

Correlations Between Young Learners' Two Phonological Memory Task Performances and Results of Vocabulary Assessments

Vocabulary assessment		Nonword repetition 1	Nonword repetition 2
Word Review	I	.356**	.164
	II	.501**	.429**

(Table 13 continues)

(Table 13 continued)

Story 1	Pretest	.287*	.155
	Posttest	.373**	.369**
	Follow-up test	.413**	.361**
Story 2	Pretest	.400**	.416**
	In-class test	.472**	.383**
	Posttest	.264*	.335**
	Follow-up test	.465**	.434**
	Online test	.366**	.350**
	RT	-.034	.038
Story 3	Pretest	.311*	.262*
	In-class test	.427**	.364**
	Posttest	.500**	.392**
	Follow-up test	.461**	.338**
	Online test	.309*	.357**
	RT	.006	.053

*$p < .05$. **$p < .01$.

4.4.1.i Nonword Repetition 1

The participants' first nonword repetition performance was positively correlated with the scores of the two word review assessments, $r = .356$ and .501 respectively, $p < .01$, (two-tailed), suggesting that those who had a good phonological memory also performed well in the two vocabulary tests. In addition to the review tests, the participants' 11 written and two online vocabulary assessment scores were all consistently and positively correlated with the score of their first nonword repetition task, r ranging from .264 to .500, $p <. 05$, (two-tailed). The findings clearly indicate that phonological memory is closely associated with young learners' receptive vocabulary. However, neither of the two online reaction times was related to the score of the nonword repetition task, $r = -.034$ for Story 2 and .006 for Story 3, $p > .05$, (two-tailed), suggesting that the learners' speed of cue mapping was not reflected in their phonological memory ability.

4.4.1.ii Nonword Repetition 2

Unlike the first nonword repetition task, the participants' performance in the second task was not associated with the scores of all the written vocabulary assessments. Their performance in the second task was positively correlated with the scores from 11 written assessments and the two online tests (r range = .262 to .434, $p < .05$, [two-tailed]). However, it was unrelated to the score of Word Review I ($r = .164$, $p > .05$, [two-tailed]) and that of the pretest for Story 1 ($r = .155$, $p > .05$, [two-tailed]). With regard to its association with reaction time, the score of the second nonword repetition task was found to be unrelated to the RT for either of the online tests, $r = .038$ for Story 2 and .053 for Story 3, both $p > .05$, (two-tailed).

4.4.2 Association of Phonological Sensitivity With Vocabulary Assessment Results

Pearson's correlation analyses were computed to examine the relationships between phonological sensitivity and all vocabulary assessment results. The report of the findings of the association is presented separately with respect to the language of the task, i.e., Chinese or English. The presentation of the results for each language includes the overall and individual phonological sensitivity skills. The possibility that an interrelationship existed between the two language tasks is also investigated.

4.4.2.i English Phonological Sensitivity

The results of Pearson's correlations computed between vocabulary assessments and the subjects' (N = 63) mean scores in all and each of the five English tasks are reported to demonstrate the association between them (see Table 14).

Table 14

Correlations Between Young Learners' Overall and Individual English Phonological Sensitivity Performances and Vocabulary Assessment Results

PS task / Vocabulary assessment	Overall	Rhyme detection	Head detection	Rhyme & head detection	Rhyme & head production	Initial consonant isolation
Word Review I	.401**	.270*	.262*	.250*	.493**	.248*
Word Review II	.730**	.437**	.537**	.536**	.685**	.588**
Story 1						
Pretest	.303*	.129	.142	.168	.317*	.317*
Posttest	.617**	.356**	.491**	.444**	.557**	.514**
Follow-up test	.653**	.390**	.526**	.436**	.606**	.541**
Story 2						
Pretest	.455**	.342**	.260*	.295*	.499**	.329**
In-class test	.581**	.323**	.425**	.390**	.551**	.502**
Posttest	.607**	.327**	.481**	.464**	.612**	.442**
Follow-up test	.637**	.431**	.515**	.431**	.613**	.477**
Online test	.541**	.224	.380**	.408**	.451**	.539**
RT	-.039	-.082	-.099	-.063	-.033	.052
Story 3						
Pretest	.347**	.315*	.081	.218	.413**	.255*
In-class test	.598**	.395**	.354**	.463**	.606**	.448**
Posttest	.665**	.405**	.410**	.457**	.595**	.618**
Follow-up test	.681**	.410**	.467**	.461**	.656**	.577**
Online test	.570**	.287*	.320*	.436**	.426**	.623**
RT	-.116	-.041	-.233	-.016	-.108	-.085

Note. PS = phonological sensitivity.
*$p < .05$. **$p < .01$.

Overall English phonological sensitivity performance

The participants' overall performance in all English phonological sensitivity tasks (M = 15.04, SD = 4.19) was mildly correlated with their vocabulary knowledge, as was indicated by each individual vocabulary assessment, including both written and online tests (r range = .303 to 730, $p < .05$, [two-tailed]). The result suggests that pupils with good English phonological sensitivity also performed well in vocabulary tests.

The association between the subjects' overall English phonological sensitivity and their Word Review II score ($r = .730, p < .01$, [two-tailed]) was stronger than that between the variable and their Word Review I score ($r = .401, p < .01$, [two-tailed]), suggesting that their English phonological sensitivity was more closely related to the vocabulary knowledge they had acquired from the study's three stories than to their previous learning outcome. In addition, the association between overall English phonological sensitivity and the written vocabulary assessments within each story became stronger, as was evidenced by the increasing value of correlation coefficients. This result suggests that English phonological sensitivity played a central role in the young children's EFL vocabulary learning, especially when explicit vocabulary teaching and the testing time were far apart. However, similar to the results for phonological memory, neither of the reaction times in the online vocabulary tests was correlated with English phonological sensitivity performance, $r = -.039$ for Story 2 and -.116 for Story 3, both $p > .05$, (two-tailed).

Performance in individual English phonological sensitivity task

Pearson's correlation coefficients between the scores for each English phonological sensitivity task and the vocabulary assessment results were computed (see Table 14). Presentation of the results is divided into detection and production tasks because of the similar patterns observed separately in the two types of task. It is again noted that reaction times for the two online tasks were not related to the variable (r range = -.233 to .052, $p > .05$, [two-tailed]), as with the results for phonological memory and overall English phonological sensitivity performance.

Detection tasks

The detection tasks, which were the first three tasks, required the participants to detect rhyming pairs and click the mouse to indicate their answers without having to speak. In the rhyme detection task, the score was associated with the results of most of the vocabulary assessments (r range = .270 to .437, $p < .05$, [two-tailed]), except the pretest for Story 1 ($r = .129$, $p > .05$, [two-tailed]) and the online test for Story 2 ($r = .224$, $p > .05$, [two-tailed]). The score in the head detection task was related to the results of the majority of the vocabulary tests (r range = .260 to .537, $p < .05$, [two-tailed]), except the pretest for Story 1 ($r = .142$, $p > .05$, [two-tailed]) and the pretest for Story 3 ($r = .081$, $p > .05$, [two-tailed]). The pattern of association between the rhyme and head detection task and the vocabulary assessment results was very similar to that of the head detection task. The score for the task was correlated with almost all vocabulary tests (r range = .250 to .536, $p < .05$, [two-tailed]), except the pretest for Story 1 ($r = .168$, $p > .05$, [two-tailed]) and the pretest for Story 3 ($r = .218$, $p > .05$, [two-tailed]).

It is noteworthy that the pretest results of the three stories were not related to the scores of the three aforementioned phonological sensitivity tasks but were related to English instruction length. The finding appears to support the two claims: (a) Instruction length contributed to vocabulary knowledge and (b) Pupils with longer instruction did not necessarily have better phonological sensitivity skills.

Production tasks

The two production tasks included (a) rhyme and head production and (b) initial consonant isolation. The former task required participants to utter the shared sound units which appeared either in the CV- or -VC position of a CVC-structure word. All vocabulary assessment results were positively correlated with the score for this task, r ranging from .317 to .685, $p < .05$, (two-tailed). The latter task required the participants to utter the consonant in the initial position of a CVC word. The score for this task was again associated with all vocabulary assessment results, r range $= .248$ to $.623$, $p < .05$, (two-tailed). The findings appear to indicate that English phonological sensitivity tasks involving oral production had a closer association with young learners' vocabulary acquisition than the other tasks which required no such oral output.

4.4.2.ii Chinese Phonological Sensitivity

The presentation of the results of the Chinese phonological sensitivity tests is modelled on that of the English results to facilitate a clear cross-language comparison.

Overall Chinese phonological sensitivity

Pearson's correlation analyses were computed between the subjects' mean score for the five Chinese tasks ($M = 13.22$, $SD = 2.40$) and their results in each vocabulary assessment (see Appendix K). The mean overall score for the Chinese phonological sensitivity tasks was positively associated with the scores of all the 12 written vocabulary assessments (r range $= .250$ to $.523$, $p < .05$, [two-tailed]), except the pretest for Story 1 ($r = .193$, $p > .05$, [two-tailed]), and the two online tests, both $r = .377$, $p < .01$, (two-tailed). However, as with phonological memory and English phonological sensitivity, Chinese phonological sensitivity was associated with neither of the online reaction times, $r = -.056$ for Story 2 and $-.033$ for Story 3, both $p > .05$, (two-tailed). The correlational relationships observed between English vocabulary knowledge and Chinese phonological sensitivity performances primarily indicate that overall Chinese phonological sensitivity, i.e., L1 phonology, was of central importance to the children's L2 vocabulary acquisition.

Performance in individual Chinese phonological sensitivity task

Pearson's correlation analyses between the scores of each Chinese phonological sensitivity task and the vocabulary assessment results were computed to determine their association (see Appendix K).

Detection tasks

Among the three detection tasks, only the score for the rhyme detection task was positively correlated with the results of all the vocabulary assessments, (r range = .286 to .459, $p < .05$, (two-tailed). The score for the head detection task was correlated with the least number of vocabulary assessment results, r range = .257 to .369, $p < .05$, (two-tailed). It was not associated with any of the three pretests (r range = .069 to .189, $p > .05$, [two-tailed]), nor was it related to Word Review I score ($r = .054$, $p > .05$, [two-tailed]), the posttest for Story 1 ($r = .181$, $p > .05$, [two-tailed]), or the online test for Story 2 ($r = .154$, $p > .05$, [two-tailed]). The score of the rhyme and head detection task was associated with the results of 11 out of the 13 written tests and the two online vocabulary assessments (r range = .253 to .438, $p < .05$, [two-tailed]), but not with any of the three pretests (r range = .155 to .211, p

> .05, [two-tailed]) nor with Word Review I, $r = .234$, $p > .05$, (two-tailed). Again, none of the reaction times was related to the variable, r range = -.186 to .100, $p > .05$, (two-tailed).

Production tasks

The pattern of association between the scores for the Chinese production tasks and those of the vocabulary tests was very different from that with the English tasks. The score for the rhyme and head production task was correlated with the results of almost all the vocabulary tests (r range = .265 to .531, $p < .05$, [two-tailed]), except the pretests for Story 1 ($r = .231$, $p > .05$, [two-tailed]) and Story 3 ($r = .209$, $p > .05$, [two-tailed]). The Chinese initial sound isolation score was surprisingly associated with none of the 15 written and online vocabulary assessments, r range = -.178 to .151, $p > .05$, (two-tailed). In comparison with their English counterparts, the Chinese production tasks were remarkably less associated with L2 vocabulary learning.

4.4.2.iii Interrelationships Between English and Chinese Phonological Sensitivity

Owing to the different patterns of association displayed by the English and Chinese phonological sensitivity tasks with young learners' vocabulary learning, a series of correlation analyses were computed to examine their interrelationships (see Table 15 & Appendix L).

Almost every individual task was positively correlated with the tasks in the other language, r ranging from .280 to .868, $p < .05$, (two-tailed). The only exception was the Chinese initial sound isolation task. The score of this task was associated with that of the rhyme and head detection task among the English tasks ($r = .304$, $p < .05$, [two-tailed]), in addition to two tasks among the Chinese tasks.

Table 15

Intercorrelations Between Phonological Sensitivity Tasks in the Two Languages as a Function of Cross-Language Transfer of Phonological Processing Skills

English task / Chinese task	1	2	3	4	5	6
1. Overall PS	**.679****	.477**	.552**	.659**	.568**	.440**
2. Rhyme detection	.543**	**.403****	.344**	.500**	.480**	.383**
3. Head detection	.441**	.355**	**.380****	.352**	.282*	.382**
4. Rhyme & head detection	.571**	.335**	.542**	**.539****	.473**	.382**
5. Rhyme & head production	.613**	.423**	.459**	.606**	**.515****	.411**
6. Initial sound isolation	.211	.198	.166	.304*	.246	-.015

Note. PS = phonological sensitivity.
*$p < .05$. **$p < .01$.

With regard to using each Chinese task as a prelude to its corresponding English task, the results appear to suggest that the participants managed to transfer their phonological processing skills in four out of five tasks, as was shown by the correlation coefficients computed between two task scores of the two languages. The result that the Chinese initial sound isolation task was not consistently associated with the other tasks suggested that the participants' performance in isolating a phonetic unit smaller than a rhyme, i.e., a phoneme, was not consistent with their performance in the other phonological sensitivity tasks. However, this was not true of the English initial consonant isolation task, because the score for this task was still correlated with all the other Chinese tasks except the initial sound isolation task. The results generally indicate a successful cross-language transfer of phonological analysis skills from the native to the foreign language. The possible explanation for the lack of association between the Chinese initial sound isolation task and the other tasks will be given in the next chapter.

4.4.3 Construct of Phonological Memory and Phonological Sensitivity

Pearson's correlation analyses revealed that both phonological memory and phonological sensitivity were positively correlated with young learners' receptive vocabulary learning results (see Tables 13 & 14), indicating that a degree of commonality might exist between these tasks. Previous studies have suggested that the two kinds of phonological measure might share a common component construct (Gathercole et al., 1991). For the present study, in addition to the two nonword repetition tasks and the five English phonological sensitivity tasks, the second word review task was also included in the analysis to test not only for a common component construct but also for a possible latent vocabulary learning factor. Before a principal-component analysis was conducted for the purpose, data screening was attempted and a determinant of *R*-matrix of .014 was obtained to avoid multicollinearity, i.e., variables that are very highly correlated, and singularity, i.e., variables that are perfectly correlated.

Table 16

Factor Loadings From Principal-Components Analysis of Word Review II and Phonological Processing Tasks

Variable	Factor 1	Factor 2
Word Review II	.81	.04
Phonological memory		
Nonword repetition 1	.65	.53
Nonword repetition 2	.65	.59
English phonological sensitivity		
Rhyme detection	.73	-.36
Head detection	.72	-.43
Rhyme & head detection	.79	-.24
Rhyme & head production	.84	-.29
Initial consonant isolation	.70	.35
Eigenvalue	4.37	1.22
% of variance	54.64	15.25

Table 16 shows that two factors were extracted from the principal-component analysis. Both factors acquired an Eigenvalue larger than the conventional cut-off point of 1.0. Factor 1 had an Eigenvalue of 4.37 and accounted for 54.64% of the variance. All eight variables loaded heavily on this factor. This pattern of associations suggests that the nonword repetition and English phonological sensitivity tasks shared a common phonological processing factor and that this factor is closely linked to vocabulary learning, because Word Review II also loaded heavily on this factor.

Factor 2 had an Eigenvalue of 1.22 and explained 15.25% of the variance. Interestingly, only the two nonword repetition tasks and the initial consonant isolation task loaded moderately on this factor. A pattern of dissociation was found between the small-unit task and other large-unit tasks. In contrast to the result in Factor 1, the loading of Word Review II was very low on this factor.

4.4.4 Summary of Young Learners' Phonological Processing Abilities

Pearson's correlation analyses showed that the two phonological processing skills—phonological memory and phonological sensitivity—were largely associated with the young learners' EFL receptive vocabulary learning. The correlations between tasks of the same type in the two languages suggest a cross-language transfer of phonological processing skills. At the same time, the result of a factor analysis indicates that the two phonological processing skills might share the same component construct and underlie a vocabulary learning factor.

In a comparison of the associations between vocabulary learning and phonological sensitivity in the two different languages, English phonological sensitivity appears to be more closely associated with the children's vocabulary performance. A trend of stronger association was observed between the results of written vocabulary assessments and those of overall English phonological sensitivity skills when a longer interval elapsed between explicit vocabulary teaching and testing time. These results further suggest that the possession of keen phonological sensitivity in the native language contributes to an acquisition of phonological sensitivity in a foreign language, which in turn plays a central role in young learners' FL vocabulary learning.

4.5 Determinants of Young Learners' Vocabulary Knowledge

In this section, stepwise regression analyses were used to determine the best predictor of young learners' EFL vocabulary acquisition, while hierarchical regression analyses were used to investigate whether phonological memory or phonological sensitivity made a unique contribution to vocabulary acquisition after the variance of one of them was removed from the equation.

4.5.1 Predictors of Young Learners' EFL Vocabulary Acquisition

Although correlation analyses are a very useful tool in explaining the relationship between variables, they say nothing about the predictive power of the variables. The type of analysis which may be used to serve this purpose is the stepwise regression analysis, which will help to determine the best predictor of young learners' vocabulary acquisition in the present study. Eight variables were considered to be possible predictors at the beginning. They were (a) young learners' age, (b) length of English instruction, (c) Word Review I, (d) phonological memory (i.e., nonword repetition task), (e) Chinese phonological sensitivity, (f) English phonological sensitivity, (g) the interactional effect between phonological sensitivity performances in the two languages, and (h) the interactional effect between phonological memory and English phonological sensitivity. Phonological memory consisted of the first nonword repetition task only, because the second nonword repetition task was not correlated with all vocabulary tests.

However, only the first six variables were eventually entered into the model after correlation analyses were computed to check the association between them. It was discovered that both Chinese phonological sensitivity and English phonological sensitivity were closely related to the interaction between Chinese and English phonological sensitivity (r = .859 and .947 respectively, $p < .001$, [two-tailed]). In addition, English phonological sensitivity was again highly correlated with the interaction between phonological memory and English phonological sensitivity (r = .876, $p < .001$, [two-tailed]). A similarly high correlation was observed when the former variable was replaced by phonological memory performance (r = .831, $p < .001$, [two-tailed]). The two predictors representing the interactional effects between phonological memory and phonological sensitivity were consequently dropped from the list of predictors to avoid a problem of collinearity. When collinearity occurs, it is difficult to obtain unique estimates of the regression coefficients. It will hence be impossible to determine which variable out of the two with a strong correlation accounts for the variance.

Stepwise regression analyses were computed, with the results of each vocabulary assessment serving as the dependent variable and the six aforementioned predictors as independent variables (see Appendix M). Only those predictors whose F-ratio had statistical significance (i.e., $p < .05$) were listed since this means that the regression model fitted the set of observed data (Field, 2000). It can be seen from Appendix M that only three out of the six entered variables were chosen as good predictors of young learners' EFL vocabulary learning. These were Word Review I, English phonological sensitivity, and English instruction length.

The variable of *English Instruction Length* was chosen three times but it was never the best predictor in any of the three tests. *Word Review I* was chosen 13 times among the 14 vocabulary assessments—seven times as the best predictor and five times as the second best predictor. The seven assessments were the pretests (for Stories 1, 2, and 3), the in-class tests (for Stories 2 and 3), and the posttests (for Stories 1 and 2), which were administered either before or soon after storytelling. In comparison, the variable of *English Phonological Sensitivity* was chosen 12 times among the 14 vocabulary tests—seven times as the best predictor and five times as the second best predictor. The seven assessments in which the variable effectively predicted the young learners' vocabulary performance were a posttest (for Story 3), online tests (for Stories 2 and 3), follow-up tests (for Stories 1, 2, and 3), and Word Review II, with the online tests being conducted in the same week as the storytelling but the other written tests being administered to students at least 10 days away from the storytelling. The results suggest that English phonological sensitivity might have played an important role in accounting for the young learners' L2 vocabulary acquisition when an extended lapse of time separated the vocabulary teaching and the tests, in contrast to the predictor of Word Review I.

4.5.2 Accountability of Experimental Measures in Vocabulary Knowledge

A series of parallel hierarchical multiple regression analyses were carried out to determine whether phonological memory and English phonological sensitivity showed differential patterns of association with receptive vocabulary (see Table 17 & Appendix N). In all the parallel hierarchical regression analyses, chronological age was entered first into the regression model, followed by English instruction length, Word Review I, and Chinese phonological sensitivity. These four variables collectively accounted for variances in vocabulary scores ranging from at least 27.7% ($F[4,58] = 5.543$, $p < .001$) for the online test of Story 3 to at most 60.4% ($F[4,58] = 22.097$, $p < .001$) for the posttest of Story 1.

Table 17

Parallel Hierarchical Multiple Regression Analysis for Phonological Processing Skills Predicting Young Learners' Vocabulary Scores as a Function of Determining Unique Contributions of PM and English PS to Vocabulary Scores

Step	R^2 change	F change	Step	R^2 change	F change
Story 3 Posttest					
5 PM	.072	7.093**	5 PM	.077	7.682**
6 PS	.144	18.638***	6 PS	.131	16.719***
5 PS	.192	24.012***	5 PS	.192	24.012***
6 PM	.024	3.050	6 PM	.016	2.060

Note. PM = phonological memory; PS = phonological sensitivity.
The variation in the parallel hierarchical multiple regression analyses is the entry of the score of nonword repetition task 1 in the *left* panel and that of the mean score of the two nonword repetition tasks in the *right* panel.
**$p < .01$. ***$p < .001$.

Two kinds of phonological memory scores were used—one was the score for the first nonword repetition task only and the other was the mean score of the two nonword repetition tasks. In the *left* panel of Table 17, the first nonword repetition (i.e., PM) score and the English phonological sensitivity (PS) score were entered in Steps 5 and 6 separately and then this order was reversed in the two steps in each vocabulary assessment, in order to determine whether the variable entered in the sixth step still accounted for any variance in and hence contributed uniquely to vocabulary scores. In the *right* panel, the score for the first nonword repetition was replaced by the mean score of both nonword repetition tasks, but the same procedures as carried out for the left panel were repeated in each vocabulary assessment. The summary of results does not include the pretests for Stories 1 and 3 because neither of the two variables in the two tests had a significant *F* change (see Appendix N). The posttest for Story 3 is listed in Table 17 for demonstrating and explaining the results.

It is clear from Table 17 that when English phonological sensitivity was entered first, phonological memory, which was entered in Step 6, accounted for very little of the variance: 1.6% in the right panel or 2.4% in the left in the case of the posttest for Story 3. However, when phonological memory was entered in Step 5 and English phonological sensitivity in Step 6, the latter still accounted for 13.1% to 14.4% of the variance in this vocabulary assessment. The fact that a similar pattern was observed in the other vocabulary assessments suggests that English phonological sensitivity contributed uniquely to the variance in explaining young learners' EFL vocabulary learning after the variance from the first four variables and phonological memory was removed.

4.5.3 Best Task Predictors of Vocabulary Performance

Stepwise regression analyses were carried out to find out which task best predicted young learners' EFL vocabulary learning. The scores for all 10 phonological sensitivity tasks from both languages as well as the first nonword repetition task and the mean score of both nonword repetition tasks were entered into the regression model. The results showing the best task predictors for each vocabulary assessment are reported in Appendix O.

The results showed that six tasks were chosen a total of 30 times from the 15 vocabulary assessments. English rhyme and head production was picked 13 times and was found to be the best predictor for 12 vocabulary assessments. English initial consonant isolation was chosen eight times—twice found to be the best predictor. Chinese initial sound isolation was chosen three times as the second best or the third best predictor. Chinese rhyme and head detection was the only detection task chosen, and this was found to be the best predictor in the pretest for Story 1. The first nonword repetition task was chosen three times but was never found to be the best predictor. The mean score of the two nonword repetition tasks was chosen twice as the second best predictor. The findings indicate that English rhyme and head production was overwhelmingly the best predictor among these tasks and that production tasks were undoubtedly better predictors than detection tasks.

4.5.4 Summary of Determinants of Young Learners' Vocabulary Knowledge

The results discussed in the previous section demonstrate that English phonological sensitivity was overwhelmingly the best predictor of young EFL children's vocabulary assessment scores. Among the variables which are most likely to be associated with young learners' vocabulary development, this skill of English phonological sensitivity accounted for more variance and hence played a crucial role, especially when an extended lapse of time separated instruction and tests. This indicates that those subjects with good English phonological sensitivity were able to retain more vocabulary items as time progressed. Among the tasks, production tasks were more effective than detection tasks in predicting children's vocabulary performance. Results from the hierarchical multiple regression analyses also showed that English phonological sensitivity contributed uniquely to vocabulary performance, after the other experimental measures were statistically controlled.

4.6 Summary of the Chapter

The four factors were examined in turn in this chapter to demonstrate their individual association with vocabulary learning. The use of L1 glosses appeared to be primarily facilitatory to young learners' vocabulary learning, since the control group who were provided L1 translation performed better with regard to online accuracy and reaction times as well as learning and retaining more vocabulary items, even though the between-group differences were not consistently significant.

Longer FL instruction time resulted in better overall vocabulary performance among young children, but this advantage was not extended to their vocabulary gains, suggesting that a better overall vocabulary knowledge was primarily a result of starting to learn the target language at an earlier age. Furthermore, instruction length was not associated with young learners' phonological processing abilities, implying that a better vocabulary knowledge in the initial stage of EFL learning owing to a longer period of FL instruction was likely to be temporary.

Phonological memory and phonological sensitivity were both positively correlated with young learners' EFL vocabulary performance, supporting the claims of previous research that the two skills contributed to the vocabulary acquisition of alphabetic languages. Both stepwise and hierarchical regression analyses proved that English phonological sensitivity was the best predictor of children's EFL vocabulary learning. In the next chapter, the findings of the present study will be discussed in relation to the existing literature.

CHAPTER 5: DISCUSSION

The discussion of the findings of this study in relation to the existing literature is divided into six parts, organised mainly around the four factors under investigation in the present study. The first section explains the qualitative link between L1 translation and vocabulary learning which facilitates a longer retention of new vocabulary items. The discussion also attempts to analyse the subjects' online behaviour and to give a step-by-step account of their processing procedures. The second section describes the important factor of instruction time and its effects on the development of vocabulary knowledge as well as its association with the two phonological processing skills. The third and fourth sections contain a description of the association between vocabulary scores and each of these two skills. A time factor was found to have affected the patterns of association each skill had with the vocabulary scores. The last two sections compare the predictive power of the two phonological processing skills and determine which was the best predictor of young learners' EFL vocabulary learning.

5.1 Extrinsic Factor I—First Language

5.1.1 Facilitator of EFL Vocabulary Learning

This part begins with a discussion of the results of the written vocabulary assessments in relation to the role which L1 glosses play in young learners' immediate recall and retention of English words, and is followed by a discussion of the connection between L1 translation and psycholinguistic studies of bilingual memory representations.

5.1.1.i Facilitation of Immediate Recall and Retention of FL Words

An examination of the starting point of the two groups (see Table 5) shows that they had similar vocabulary knowledge, although the significant between-group difference in the pretest for Story 2 reveals that the experimental group knew more target vocabulary than the control group before this story was told. The subsequent written vocabulary assessments (see Table 6), however, show that the two groups continued to perform similarly, while the control group taught using L1 glosses was able to close the statistically significant gap in Story 2 in the space of just one story session, as was evidenced by a tie with the experimental group in the in-class test. When compared with that of the experimental group, the performance of the control group is consistent with the findings of previous research that a provision of L1 translation equivalents contributes to young learners' immediate recall of new FL words (Hsieh, 2005). It is likely that when L1 translation was given and attached to English words, a short-term memory was created for these FL phonological representations in connection with their L1 translation equivalents. The connection of the two languages at the lexical level enabled the control group to catch up with the experimental group in the surprise in-class vocabulary test which was added to the study. A comparison of the vocabulary scores reveals that the experimental group did not make any progress from their pretest, suggesting that the group might not have undergone the same learning processes as the control group. Despite representing group performance, the overall vocabulary scores were confounded by the learners' preexisting vocabulary knowledge and could not accurately represent the progress individual learners made.

Vocabulary gains, on the contrary, might represent more accurately the learning outcome of the explicit vocabulary teaching activity since they measured the increase in the vocabulary of individual learners between the pretest and a later administered test. A look at the vocabulary gains (see Table 7) reveals a very different picture from the overall vocabulary scores (see Table 6), since the control group, rather than the experimental group, was shown to have learned and retained more new words between the two vocabulary tests. Among the eight counts of vocabulary gains, the control group performed better on seven of them, while significant between-group difference was observed on two counts in Story 2. This result is, on one hand, consistent with previous research which shows that L1 translation has facilitated the retention of FL words (Laufer & Shmueli, 1997; Liao, 2005; Lotto & de Groot, 1998). On the other hand, it also indicates that L1 translation, rather than picture cues, might have created not only short-term memory but also long-term memory for FL words, with English words being labelled by L1 glosses in an already existing schema

(Coady et al., 1993; Hulme et al., 1991).

5.1.1.ii Interconnection of L1 and FL at the Lexical Level

As it has been suggested that foreign language learning does not repeat the course of native language acquisition, it is very likely that FL words are parasitic on the L1 and that these EFL learners attached an English label to an already existing native-language schema rather than building an entirely new schema for universal concepts (Coady et al., 1993; MacWhinney, 2005; Vygotsky, 1986). Consequently, the control group was able to employ the strong connection between English words and their L1 translation equivalents to help them memorise new target vocabulary. By contrast, as the experimental group was provided with picture cues and English phonological forms—the pedagogical method which was modelled on first language acquisition—these young learners might have eventually needed to attach labels to the pictures when they were overwhelmed by an increasing number of pictures and unfamiliar FL phonological representations, because their language proficiency was not sufficiently advanced for them to access FL meaning conceptually.

Kroll and Stewart (1994) have proposed the Revised Hierarchical Model, which posits that the strength of the connection between the lexical and conceptual representations for a bilingual's two languages differs primarily as a function of L2 proficiency and the relative dominance of the L1 over the L2. It is suggested that beginning L2 learners have their two languages interconnected at the lexical level and will not be able to access L2 meaning conceptually because of the constraints caused by the weak connection between the L2 lexicon and the conceptual representations. Since the subjects of the present study were in the very early stage of English learning and their L1 was overwhelmingly dominant, the connection between their FL lexical and conceptual representations was relatively weak in comparison to the connection between the two languages at the lexical level. The control group, which was readily provided with L1 glosses, was able to process the FL phonological forms and attach them to their L1 equivalents to enhance the strength of the connection between the two languages. When the English label was attached to the L1 equivalents, retention was made easier probably because a long-term memory was created for these EFL words (Hulme et al., 1991). By contrast, the experimental group, which could not access FL meaning conceptually owing to their limited English proficiency, could face a greater loss of new words because the connection between the picture cues and the FL phonological forms was not strengthened in their long-term memory.

Despite the explanations of the role played by L1 translation in EFL vocabulary learning, the above discussion regarding the written vocabulary assessments could not truthfully reflect the subjects' behaviour in terms of the time they took to make their online choices. As reaction times are believed to reflect participants' processing complexity (Cook, 1990; McDonald, 2000; van Hell & Candia Mahn, 1997), the investigation of the first language effects hypothesis (see 1.5.1.i) will be based on the results from the two online vocabulary assessments. The explanations of subjects' online behaviour will help to interpret how the young learners stored and accessed the English words in their minds, which will in turn reveal how they approached the target language in the initial stage of EFL learning.

5.1.2 Longer Reaction Times via Pictorial Mediation

In the present study, the control group, who were given L1 glosses in explicit vocabulary teaching, consistently responded more quickly in both the online tests than the experimental group, who were given picture cues only (see Table 8). In terms of the correct RT, in which question items were answered correctly, the between-group difference in reaction times was significant in Story 2 but not in Story 3. This failure to obtain a consistent result made it difficult to reach a definite conclusion at this point that L1 glosses should always be given in EFL vocabulary teaching as an aid to strengthen the connection between the two languages. However, the consistently shorter reaction times by the control group do suggest that L1 translation might have resulted in faster mapping between picture cues and English phonological forms, and hence lends support to the claim that the L1 is more of a facilitator than a hurdle in the initial stage of children's FL word learning.

Although the discussion in the previous section indicates that L1 translation facilitated the young learners' immediate recall and their retention of more EFL vocabulary, the finding that reaction times were longer when using a pictorial mediation of the L2 meaning is rather unexpected in two aspects. First of all, it contradicts the findings of previous research in which picture stimuli produced shorter reaction times in young learners (Chen & Leung, 1989). Secondly, it also calls into question the hypothesis concerning L1 effects, which predicted that the experimental group, which was instructed to map L2 phonological representations directly to their corresponding pictures, would process one step fewer than the control group and hence produce shorter reaction times.

The finding of this study that reaction times were longer when using pictorial mediation contradicts the results of Chen and Leung's (1989) study in particular, in which child beginners were found to have developed a better picture-to-concept link and to have used pictorial representations as a medium to access meanings in L2. In their study, the young subjects displayed quicker responses when naming pictures in L2 than when translating words from the L1 to the L2, in stark contrast to adult beginners. The child beginners' advantage in picture-naming was attributed to the method of original L2 instruction with these two different age groups. The child learners were given concrete teaching media such as real objects or pictures when taught new vocabulary, while the adult learners were given word pairs. The authors argued that the instruction method must have led to a better development of the picture-to-concept link than the word-to-concept link in the child beginners.

In the present study, flashcards containing pictures corresponding to the target L2 words—teaching aids similar to those used in Chen and Leung's (1989) study—were used, but the result obtained was dissimilar to theirs. The experimental group, which was pedagogically instructed to map the English words to the picture cues, consistently demonstrated longer reaction times. This result contradicts Chen and Leung's argument that child beginners were liable to access L2 meaning via pictorial representations if real objects or pictures were used in teaching. As the result from the present study shows the opposite, it suggests that the method of instruction might not have been as effective as it was believed to be because of the privileged status of the children's L1 (Ervin & Osgood, 1954; Weinreich, 1953). It is, however, inadequate to refute completely the impact of pedagogical instruction, since the effective use of environmental events or pictorial representations in L2 teaching is modelled on first language acquisition and is empirically supported as well (Jenkins, 1968; Wimer & Lambert, 1959). It is hence speculated that, in the present study, both groups of children employed pictorial representations as well as L1 translation while attempting to map the FL phonological representations to the cued pictures online, irrespective of the pedagogical difference. It is highly probable that the control group had shorter reaction times because they had available a readily-matched mapping of English phonological forms to L1 translation equivalents from the classroom instruction they had received, while the experimental group had to spend extra milliseconds in locating the L1 names for the pictures.

An attempt will be made next to reconstruct the processing steps which might have been employed by the young learners during the online tasks. The results from interviews conducted immediately after the online tests will be used to explain the step-by-step analysis of the online processing procedures, which in turn lend support to the claim regarding the route of L1 translation in FL vocabulary learning.

5.1.3 A Route of L1 Translation in Initial L2 Learning

5.1.3.i Interviews After Online Tests

The first piece of evidence supporting the possibility that participants employed L1 translation equivalents during online processing comes from the survey after the online vocabulary assessments. According to the result of the interviews with the subjects, more than half of them in both groups admitted that L1 translation equivalents had occurred to them during the online tests. A further question was then put forward to clarify whether some words might have been learned along with their L1 translation in their extracurricular lessons before the storytelling sessions. A mixed picture was presented by the subjects from both groups. However, the children from the experimental group expressed that L1 translation had occurred in their processing of even those words they had learned from the storytelling sessions, during which no Chinese glosses were given. The interview results indicate not only that the native language had a far greater impact on how young learners learned a new language than expected but also that the L2 instruction method had a far lesser impact than expected. The investigation into L1 effects is not complete, however, without taking a look at the participants' online processing behaviour. An analysis of the possible processing steps is described below in an attempt to retrace how differences in reaction times might have occurred.

5.1.3.ii Analysis of Processing Steps in Online Vocabulary Tasks

During the online vocabulary tasks, an auditory cue consisting of an English phonological representation and a visual cue consisting of a picture appeared simultaneously, and the subjects were required to decide whether the two cues matched each other. Since it has been established that L1 translation was employed by both groups during their online processing, irrespective of the pedagogical instruction they had received, three processing steps were expected of all subjects. These involved recognising the picture cues, processing the English phonological forms, and locating L1 translation equivalents.

Step 1: Recognition of a cued picture

It has been established that naming and categorisation tasks involve different types of processing in the memory. When the two levels of information are called upon, the information at the semantic level should arrive faster than that at the lexical level, i.e., pictures are recognised (or categorised) faster than words are (Collins & Loftus, 1975; Kroll & Potter, 1984; Potter & Faulconer, 1975; Smith & Magee, 1980). It is hence assumed that picture recognition took place before the subjects recognised the L2 phonological forms. An alternative explanation is that the connection between the L1 lexicon and images/pictures was likely to be stronger than that between an L2 word and its referent, i.e., a faster concept-to-L1 link in the beginning FL learners. Consequently, picture recognition is expected to have been the first step which took place in the online vocabulary assessments.

Step 2: Processing, retention, and recognition of an English phonological form

Even though picture recognition was likely to be the first processing step, the retention and recognition of L2 phonological forms was the most crucial step to the subjects. Rather than being prepared to recognise the cued picture and make an immediate guess as to the name of the picture in the target language, the young EFL learners had to process and retain the FL phonological form in their working memory before they were able to recognise it and give their answer. The empirical evidence in support of the previous claim that the subjects were more likely to process and retain the auditory cue rather than produce the name for the picture comes from the results of the pencil-and-paper tests and the productive vocabulary task.

In the written vocabulary tests, the children were able to map the spoken form of the FL word correctly to their corresponding pictures on the worksheet within 3000 ms, the interval at which every word was read. But in the productive vocabulary task, less than 40% of the pictures were successfully named within the given time of 6000 ms, in contrast to the average reaction times of a little more than 2000 ms in the online vocabulary tests (see Table 8). If the participants could not successfully name the pictures within 6000 ms, it was very unlikely that they would recognise a picture, conjure up its phonological representation in the L2, compare it with the auditory cue, and click the mouse to indicate their answer within an even shorter time.

However, according to the first hypothesis for this study (see section 1.5.1.i), the experimental group should have been able to map the auditory cue to the visual cue and give their answer at this point and consequently achieve shorter reaction times than the control group. The opposite results from the online reaction times indicate the likelihood that their online processing was still taking place beyond this point. Taken together, the results from the online tasks and the subsequent interviews suggest that subjects of the experimental group might have undertaken the next processing step.

Step 3: Location of L1 Translation Equivalents

As Kroll and Curley (1988) proposed that the two languages of L2 beginners are connected at the lexical level, the job remaining for the subjects at this point was to succeed in locating the L1 translation equivalent for the retained L2 phonological form. It is hence suggested that the between-group differences in reaction times occurred during this step, when the subjects of both groups were searching for the L1 translation. The control group had the advantage over the experimental group because the link between an English phonological representation and its L1 equivalent was stronger for them, since L1 translation equivalents had been provided and attached to the English words in the explicit vocabulary teaching. The experimental group, on the other hand, might have spent hundreds of milliseconds more searching in their L1 lexicon to locate the right candidate. It was only after they had completed this processing step that the subjects clicked the mouse to indicate their answer.

The above step-by-step analysis of online processing behaviour shows that the L1 is an inevitable route in the beginning phase of learning another language and that the FL beginners used the route of word association, instead of either pictorial representation or conceptual mediation, to access L2 word meaning. This is also the case in Kroll and Stewart's (1994) Revised Hierarchical Model of lexical and conceptual representation in bilingual memory. Furthermore, psycholinguistic research suggests that lexical candidates are routinely activated in L1 among proficiency L2 users when words in L2 are processed (Kroll, Michael, Tokowicz, & Dufour, 2002). If an L1 translation equivalent is activated when a proficient bilingual processes an L2 word, it is highly likely that a less proficient L2 learner will have to resort to the L1 route because the strength of the connections between lexical and conceptual representations in a bilingual's two languages is determined by L2 fluency and the relative dominance of the L1 over the L2. The present study hence lends support to their view that L2 learners in the initial phase of learning another language have their two languages interconnected at the lexical level and that they access the L2 meaning via this lexical association.

5.2 Extrinsic Factor II—Length of English Instruction

5.2.1 Contributor to a Larger Vocabulary

5.2.1.i Contribution of Time Factor to Overall Vocabulary Scores

In the present study, the length of the subjects' English instruction was positively associated with the results of all written vocabulary assessments (see Table 10). The positive association between longer instruction length and vocabulary scores is consistent with the findings of previous research that a longer exposure to the target language was beneficial to vocabulary size in the EFL context (Y.-Y. Chen, 2004; Nieh, 2004; Yin, 2006). This demonstrates that length of English instruction plays a central role in determining the level of vocabulary knowledge among young learners in an EFL context.

The advantage of greater vocabulary knowledge as a result of longer English instruction is clearly evidenced by the three vocabulary pretests. In the three pretests, the young learners who were tested on the to-be-taught vocabulary items showed differential degrees of vocabulary knowledge, their vocabulary scores being positively associated with the length of English instruction they had received. This association between instruction time and vocabulary scores suggests the importance of the time factor: As more time is allowed, more words can be learned by a child (Cameron, 2002; Macnamara, 1966). This advantage of the participants who had learned English for longer and who hence performed better in the pretests was extended to their performance in subsequent vocabulary assessments, for example, posttests and follow-up tests. The number of new words they learned from the stories was added to the number of previously learned words in the later administered assessments. In comparison, the children who knew fewer words since they had studied English for a shorter period were faced with a heavier learning load if they wanted to keep up with the early starters within the same amount of story instruction (Hakuta et al.2000).

Given that the early and late starters had the same speed of vocabulary learning, it is likely that the late starters eventually had more unknown words than the early starters because of the preexisting gap in their vocabulary knowledge.

It is thus not difficult to link the demand to lower the age of starting to learn English with the phenomenon of popular extracurricular English lessons in Taiwan, since FL competence is largely accounted for by "a matter of the amount of time spent in learning" (Carroll, 1969). This claim is justified by the attribution of a larger vocabulary size to longer instruction time because young learners in Taiwan need to be instructed, rather than trying to teach themselves how to crack the code of the language, despite the general availability of English resources from TV programmes and in print, in books and magazines. Mere exposure to books or TV programmes will not facilitate the foreign language learning of young students, as English is very different from Chinese, both phonologically and orthographically.

5.2.1.ii Dissociation Between Vocabulary Gains and Instruction Time

The discussion of the overall vocabulary scores has demonstrated that a longer period of instruction contributed to a larger vocabulary in young EFL learners. However, the close association between English instruction length and vocabulary learning disappeared when only vocabulary gains were compared against instruction length (see Table 11). Not one count of vocabulary gains between two vocabulary assessments was associated with the factor of instruction length. This result indicates that those participants who had received more English instruction did not necessarily learn new words more quickly than those who had received less. In other words, more English input might have contributed to the amount of vocabulary acquisition but did not lead to a faster speed of learning. This finding was different from those of some studies, which found that children with higher language abilities acquired new lexical items more quickly from listening to stories than those with lower abilities (Penno et al., 2002). The result of the present study, indicating a dissociation between instruction time and vocabulary gains also suggests a lesser possibility of incidental learning among children learning English as a foreign language, since they are more likely to pay attention to those words most relevant to them in the early stage of language learning when their capacity to process information is limited.

The lack of a continued association between the young EFL learners' vocabulary gains and their English instruction length also indicates that English instruction length is not sufficient to account for wide individual differences in vocabulary learning. In addition, as longer exposure did not result in faster vocabulary learning, the advantage of gaining a greater amount of vocabulary could be short-lived and easily overtaken, because the words taught in the early stages of foreign language learning are words of high frequency which will soon be covered. As White and Genesee (1996) have advised, one or two hours' instruction in an EFL context does not constitute significant exposure to the target language because the classroom instruction would not prompt students to use the language for communication. It certainly takes more than a head start for young learners to become proficient FL users (Williams, 1996).

5.2.1.iii Limitation of Vocabulary Intake

An interesting pattern of vocabulary learning was revealed by the young learners' vocabulary gains from the three stories. The mean number of word gains in Stories 1 and 3 reached up to a maximum of six, although it fluctuated from test to test (see Table 11). But the mean number of vocabulary gains in Story 2 remained at around three, only half of the number in the other two stories. When the three pretests were compared, Story 2 was regarded as the easiest because the participants scored highest in the pretest. By contrast, Story 3 seemed to impose more learning loads on the participants because they scored lowest in the pretest for that story. However, these participants did not manage to gain more or even the same number of new vocabulary items in Story 2 as in the other stories. This result seems to indicate that an average of 15 words was the maximal number of target words the young learners could manage to store in their working memory in two storytelling sessions, irrespective of the fact that some of the words had been learned previously.

5.2.2 Dissociation Between Instruction Length and Phonological Processing Skills

In contrast to the consistent correlation between instruction time and their overall vocabulary scores, the participants' instruction length was associated with neither phonological memory nor phonological sensitivity (see Table 12). In addition to the aforementioned lack of association between English instruction length and vocabulary learning speed, this result demonstrates that the factor of instruction time contributed positively to the amount of learned vocabulary but did not make a direct contribution to the children's capabilities in vocabulary learning. This does, however, justify the attribution of young children's EFL vocabulary learning to intrinsic factors such as phonological memory and phonological sensitivity, since the phonological processing skills are categorised as intrinsic factors and are not shaped by the external factor of instruction length. That is to say, the young learners' ability to learn EFL vocabulary depends on their L2 phonological processing skills which are likely to be inherent but vary from one to another, as was shown by the correlation between their vocabulary scores and phonological processing performances.

Two points associated with phonological memory and phonological sensitivity may be extrapolated from this dissociation. The first is the claim made in the methodology chapter (see section 3.2.2) that the wordlikeness of nonwords has no effect on young EFL learners in the initial stage of learning. The second is that speakers of a first-learned nonalphabetic language might possess a lesser degree of phonological sensitivity. These two points are discussed in more detail below.

5.2.2.i Nonwords as a Pure Measure of Phonological Memory

The lack of association between phonological memory and English instruction length, on one hand, shows that more English input did not promote phonological memory. On the other hand, it also supports the claim made in section 3.2.2 that the level of wordlikeness of nonwords would not affect the performance of beginning learners of English in an EFL context because they had received too little input of the target language and hence did not possess knowledge of language-specific phonological sequences (Speciale et al., 2004). Nonwords of high wordlikeness are said to trigger long-term lexical knowledge and hence contribute to a better performance of immediate verbal memory (Dollaghan et al., 1995; Gathercole, 1995). However, the lack of association between English instruction length and nonword repetition task scores shows that nonword repetition tasks, irrespective of their wordlikeness, are a pure measure of the beginners' phonological memory in an EFL context.

This is especially the case in the present study. Nonword repetition has been a pure measure of phonological memory in young Taiwanese learners because, owing to the nature of their native languages, measuring a learner's ability to store a multisyllabic phonological representation can be a way of discriminating between good and poor EFL learners. As the two native languages—Taiwanese and Mandarin Chinese—of the participants are both monosyllabic, a multisyllabic English word means a combination of at least two sound units to them. When the number of syllables in a word increases, the level of difficulty in repeating it naturally goes up. The finding in this study that nonword repetition was associated with vocabulary scores reflected the possibility that those learners who scored higher in the vocabulary assessments were likely to possess better phonological memory because they successfully retained and repeated the phonological representations of nonwords.

5.2.2.ii A Lesser Degree of Phonological Sensitivity

The finding that young learners' phonological sensitivity was not associated with the length of time they had been learning English clearly demonstrates that one or two more years' English input in the initial phase of learning the language had not helped them to acquire a good sense to discriminate phonologically similar words. This result has two implications. First, as previously stated, it implies that phonological sensitivity could be an underlying phonological processing ability that is not accounted for by extrinsic factors such as quantity of input. The other implication is that the monosyllabic and logographic nature of the Chinese language might have impeded the development of phonological sensitivity among the young learners, and this lack of phonological sensitivity becomes apparent when they begin to learn an alphabetic language such as English (Caravolas & Bruck, 1993; Cheung, 1999).

Aspects of phonological input and the nature of the orthography are found to have shaped the early development of phonological awareness in children, as well as to affect the rate and pattern of literacy skills, among child speakers of a transparent and a less transparent alphabetic language (Caravolas & Bruck, 1993; Winskel & Widjaja, 2007). If differences could be observed among alphabetic language users, the gap would be wider when it comes to those who use a nonalphabetic language. In fact, the lesser sensitivity to subsyllabic phonological units has been observed in first-learned logographic language users (Cheung, 1999; Holm & Dodd, 1996). The discussion in this section, however, is restricted to possible ways of accounting for a lack of association between English instruction length and phonological sensitivity measures, and will not extend to an attempt at determining whether these young learners with a first-learned nonalphabetic language background had less phonological sensitivity than those with an alphabetic language background.

The above discussions of the two extrinsic factors have demonstrated that an input of L1 translation qualitatively improved young learners' EFL vocabulary learning and that instruction time quantitatively increased their vocabulary size. The following sections will present discussions of the intrinsic factors in relation to their associations with young learners' EFL vocabulary acquisition.

5.3 Intrinsic Factor I—Phonological Memory

5.3.1 Determinant of Good Vocabulary Learners

Significant positive correlations were generally found between vocabulary scores and the two nonword repetition tasks, with the exception of two vocabulary tests (see Table 13). Possible reasons for the lack of association between these two vocabulary assessments—the pretest of Story 1 and Word Review I—and the second nonword repetition task may be a combination of the time of testing and a possible change in the structure of the participants' vocabulary knowledge. As further explanations will be provided in section 5.3.1.ii, the results of these two tests are hence excluded from the following section.

5.3.1.i Positive Association With Vocabulary Scores

The consistent positive correlations found in the present study between the participants' vocabulary assessment scores and their phonological memory performance confirm the link between nonword repetition and English vocabulary learning (Bowey, 1996, 2001; Gathercole & Baddeley, 1989; Gathercole et al., 1997; Gathercole & Thorn, 1998; Service & Kohonen, 1995). The results indicate that those who performed better in the nonword repetition tasks also scored higher in the vocabulary assessments, supporting the claim that performance in tests of immediate verbal memory is closely related to an individual's ability to learn new phonological forms of a language (Gathercole & Baddeley, 1993). The association between nonword repetition and vocabulary scores could be explained by the similarities between a nonword and a new lexical item. As nonword repetition evaluates a learner's ability to repeat back a string of unfamiliar phonological representations, it reflects whether the learner possesses the ability to temporarily store the new phonological form in his short-term memory in one encounter. Similarly, every time a new lexical item is presented, its phonological representation is stored in the short-term memory. If the phonological representation is successfully transferred to the long-term memory, learners would be able to recall the phonological form and make a guess as to its meaning when they encounter it the next time. The connection between the word's meaning and its phonological form is therefore strengthened every time they match (Nation, 2001). The connection between nonword repetition and phonological memory is hence reflected by the vocabulary scores in the present study. However, if learners have difficulty even in correctly repeating the phonological form they have just heard, it is very unlikely that this word will be transferred and stored in their long-term memory. There is no doubt that the ability to initially repeat and to store the new phonological form in the short-term memory is crucial to successful vocabulary acquisition.

5.3.1.ii Reflection of Phonological Memory by the Timing of Testing

At the beginning of the discussion of the correlations between nonword repetition and vocabulary scores, it was mentioned that two vocabulary assessments—the pretest for Story 1 and Word Review I—were exceptionally *un*correlated with the second nonword repetition task, while all the other vocabulary assessments were positively associated with both nonword repetition tasks (see Table 13). The lack of association between these two assessments and the second task could be explained by the time of testing and the possible change in the participants' vocabulary knowledge in the wake of an intensive input of the target language.

Nonword repetition appears to be very sensitive to the time of testing and reflects the learners' concurrent capability of phonological memory. The two vocabulary assessments were administered on the same day, three days prior to the beginning of the storytelling programme, and hence were best explained as representing the level of vocabulary knowledge of the young learners before the launch of the storytelling sessions. It was not surprising that the two assessments were still correlated with the first nonword repetition task, considering that the concurrent association is often found between vocabulary learning and phonological memory (Gathercole et al., 1997). However, as increased English instruction from the story sessions was likely to produce a change in the participants' phonological memory and the structure of their vocabulary knowledge, as was evidenced by the higher mean score in the second nonword repetition task and the increase in the mean number of their newly-learned words, the task was found eventually to be associated with only those vocabulary assessments administered after the story sessions, and not with the two administered before.

5.3.1.iii Increase of Phonological Memory Over Time

The moderate correlation between the two nonword repetition tasks suggests that those learners who achieved a high score in the first task were able to maintain their performance in the second one. The higher mean score in the second task shows that the participants on average performed better in the second task than they did in the first. The first task was conducted at the beginning of the study and the second one near the end of the study, with a lapse of five weeks between them. When the result is viewed from the developmental perspective of the phonological memory framework, it indicates that the young learners' phonological memory skill improved during the course of the study (Roodenrys et al., 1993).

5.3.2 The Downside of Nonword Repetition

5.3.2.i Reflection of Weak Association by Low Correlation Coefficients

The individual correlation coefficients between vocabulary scores and the two nonword repetition tasks obtained in the present study ranged from .262 to .501 (see

Table 13), several of which fell off the lower end of the range usually obtained, which is claimed to be between .34 and .61 (Baddeley et al., 1998). The low *r* values obtained in the present study might have reflected the low internal consistency of the nonword repetition items (Cronbach α values = .55 for the first task and .33 for the second task). However, a calculation of the mean correlation coefficient between vocabulary scores and nonword repetition gave .394 for the first task (based on 15 tests) and .368 for the second (based on 13 tests with an exclusion of the pretest for Story 1 and Word Review I), both figures being within the range claimed by Baddeley et al. Even after the addition of the two aforementioned vocabulary tests, whose correlation coefficients did not reach significance level, the mean correlation coefficient for the second task remained at .341, just within the lower end of the range usually obtained by Baddeley et al. However, the *r* values obtained in the present study should be regarded as a pure reflection of a weak association between nonword repetition and the young learners' vocabulary performance, rather than as the result of a methodological difference as claimed by Gathercole and Baddeley (1997).

A similar low correlation coefficient ($r = .35$) was obtained between nonword repetition and vocabulary age in Bowey's (1996) investigation of the hypothesis that phonological memory, rather than a latent phonological processing factor which was shared by phonological sensitivity as well, accounted for significant variation in young children's receptive vocabulary. The multiple regression analyses in Bowey's study revealed that phonological memory was no more strongly associated with receptive vocabulary than was phonological sensitivity, and Gathercole and Baddeley (1997) suggested that this result could be attributed to the low correlation coefficient between vocabulary and nonword repetition, which in turn was caused by the differences in methodological procedures from theirs in the aspect of the method of presenting the same set of nonwords. However, a closer look at the table of the correlations observed among the experimental variables in Bowey's (1996) study reveals that the correlation coefficients between vocabulary and two tasks of phonological sensitivity were .35 and .38, which were just as low or only slightly higher than that between vocabulary and nonword repetition. The low correlation coefficient and hence the methodological difference should not have been blamed as the cause when similarly low correlation coefficients were obtained between vocabulary and the two phonological sensitivity measures. The explanation for Bowey's (1996) results could be extended to the present study: The methodologically procedural differences should not be blamed for a lower correlation coefficient between nonword repetition and vocabulary scores. The low correlation coefficients, on the contrary, should be regarded as a true reflection of the weak association between the two variables.

Gathercole and Baddeley's (1997) criticism of Bowey's (1996) study centred on their different presentation methods of the same set of nonwords. Bowey (1996, 2001) chose to use live presentation to engage young children in the task, in contrast to a prerecorded presentation to control for presentation variance (Gathercole et al., 1991). Despite the criticism, Gathercole et al. (1997) adopted the same method of live presentation and found a moderate correlation ($r = .60, p < .05$) between nonword repetition and receptive vocabulary. Gathercole et al. (1999) used both presentation methods and obtained moderate correlation coefficients with vocabulary ($r = .54$, $p < .05$ for live presentation and $r = .61$, $p < .05$ for prerecorded presentation). A comparison of these results indicates that correlation coefficients are comparable from study to study, but that a low r value cannot be attributed to the use of different presentation methods. As previously stated, the low correlation coefficients obtained in the present study were hence simply a reflection of the association between nonword repetition and vocabulary learning among these EFL learners, rather than the result of a difference caused by the methodological design. The low correlation coefficients should not be held responsible when phonological memory does not contribute uniquely to the variance of vocabulary scores in the later hierarchical regression analyses.

5.3.2.ii Decrease of Association With Vocabulary Scores

The two mean correlation coefficients calculated for each nonword repetition task reveal a decline in the association between vocabulary scores and nonword repetition, indicating a possible decrease in the importance of the role played by nonword repetition in predicting young learners' vocabulary performance. A similar pattern was observed by Gathercole and Adams (1994), where correlations between nonword repetition and receptive vocabulary remained significant at Age 4 ($r = .26, p < .05$) but dropped to nonsignificance at Age 5 ($r = .22, p > .05$). In the present study, the decline in the association between nonword repetition and vocabulary learning might account for the finding that phonological memory was a less powerful predictor of the young learners' vocabulary performance in later multiple regression analyses.

5.4 Intrinsic Factor II—Phonological Sensitivity

The discussions of the link between phonological sensitivity and the results of the vocabulary assessments will focus mainly on the English measures. The Chinese measures, as stated in the chapter on methodology, were principally intended to act as precursors to their English counterparts, and hence the focus of discussions in relation to the Chinese measures will be on the manifestation of cross-language transfer of phonological processing skills.

5.4.1 Developmental Trend of the Phonological Sensitivity Skill

The results of the phonological sensitivity tasks conducted for the present study shed light on several aspects of the nature of the development of the phonological sensitivity skill which are in line with the developmental perspective of this skill as observed in native child speakers.

5.4.1.i Development of Phonological Sensitivity From Rhyme to Phoneme

The percentage rates of correct answers given by the subjects in the English phonological sensitivity tasks (see Table 9) were compared in order to evaluate their phonological processing skills on large and small units. The comparison reveals that, in these young EFL learners, large-unit phonological processing skills—rhyme awareness and head awareness—were better developed than the small-unit skill of phonemic awareness. Additional *t*-tests revealed that the young learners performed in-differentially between the two large-unit detection tasks but that there were significant differences between their performances in both large-unit tasks and the initial consonant isolation task (see section 4.2.4). This result is consistent with the findings of studies where a developmental perspective of phonological processing skills stipulates that large-unit rhyme sensitivity develops earlier than small-unit phonemic awareness (Chien, 2002; Goswami, 2002; Goswami & Bryant, 1990; Torgesen & Davis, 1996; Torgesen, Morgan, & Davis, 1992; Yang, 2006; Yopp, 1988).

This result from the present study has two implications in relation to the EFL context. The first is that the large-unit phonological processing skill developed regardless of reading instruction, and the second is that the skill was refined after the young learners started to have contact with the alphabetic language.

It has been argued that instruction in learning to read an alphabetic script is a prerequisite for the ability to explicitly represent spoken words as sequences of phonemes. In contrast, rhyme sensitivity develops without exposure to alphabetic reading instruction in children and illiterate adults (Goswami & Bryant, 1990; Morais et al., 1986). This, in turn, explains the developmental perspective that large-unit processing skills should emerge earlier than small-unit skills. In the present study, as all participants had been taught the alphabet and phonics rules for more than one year when these tasks were administered to them, they naturally possessed both some rhyme and some phonemic awareness. Although they had received more instruction in small-unit awareness, i.e., phonics, the result shows that they had more sophisticated large-unit phonological sensitivity. It was likely that large-unit phonological sensitivity was more of a universal skill reflected in the phonology of all languages and developed regardless of whether an opaque relationship existed between the orthography and the phonology of a language (Gottardo et al., 2001).

It is also likely that their rhyme awareness was refined after the young learners started having contact with the alphabetic language. It has been suggested that phonological awareness skills in young children develop gradually as a result of spoken vocabulary growth (Metsala, 1999; Metsala & Walley, 1998). According to Metsala and Walley's (1998) Lexical Restructuring Model, very young children start to recognise words in a more holistic manner because there is no need to represent words in a systematic manner when their vocabulary size is still small. However, owing to the increased demands made by a rapidly growing vocabulary, they gradually develop more segmental representations so that they can analyse and cross-reference with their existing phonological lexicon to facilitate efficient storage and online recognition. As the vocabulary grows, these phonological representations are re-structured, and the segmentation of sound develops from a large-unit rhyme to a small-unit phoneme. In the present study, the young learners were given a large quantity of English vocabulary, which included phonologically similar words such as *rug* and *bug*, semantically similar words like *king* and *queen*, or both phonologically and semantically similar words such as *triangle* and *rectangle*. They needed to pay attention to the phonological representations in both the large and the small units in order to discriminate between such words. In order to counteract the effects of phonological similarity, the children's large-unit phonological processing skill had to be refined. This facilitated a re-structuring of their existing vocabulary knowledge.

At the same time, their phonemic awareness also continued to develop. When the children developed the ability to segment word sounds at the phonemic level, their sensitivity to large units was expected to become keener as a result.

It should be pointed out here that the comparison between large and small units of phonological representations did not include the two tasks of rhyme and head detection and production. These two tasks were excluded at present because they involved two different processing skills in one single task—one being rhyme and the other head. It is vital that the phoneme task be compared with the large-unit tasks which are at the same processing level when examining the developmental relationship between large- and small-unit phonological processing skills in young EFL learners. It is hence inappropriate to use the scores from the rhyme and head detection and production tasks for the comparison here. However, a discussion of the association between these two-layered tasks and the one-layered tasks will be presented in the next section, when the difficulty levels of the phonological sensitivity measures are compared.

5.4.1.ii Levels of Difficulty in Phonological Sensitivity Measures

Among the five English phonological sensitivity measures, the young learners scored highest in rhyme detection and lowest in rhyme and head production (see Table 9). At first glance, the results may be wrongly construed to indicate that the small-unit phonological skill of the young EFL learners was more highly developed, because they scored higher in the initial consonant isolation task than in the rhyme and head production task between the two production tasks. As previously stated, it would be inappropriate to compare the rhyme detection, head detection, and initial consonant isolation tasks with the rhyme and head detection and production tasks in the discussion of the developmental trend of phonological processing skills, because the former three and the latter two tasks placed two different levels of cognitive demands on the young learners (Adams, 1990). Each of the former three tasks required the children to exercise their phonological skills in one aspect in a single task, only needing to heed the sound unit(s) in the specified position, for example, the rhyme (-VC) in the rhyme detection task or the onset (C-) in the initial consonant isolation task. By contrast, each of the two latter tasks combined the processing of rhyme and head in one single task, which meant that the subjects had to simultaneously exercise the two different processing skills to differentiate the sound patterns in both the CV- and -VC positions. These five tasks were in fact at two different levels of difficulty in terms of cognitive demands, the three former tasks being less cognitively demanding than the two latter tasks.

However, the task type—detection or production—obviously played a role in affecting task performance and complicated the difficulty levels of the tasks, as the subjects were found to have scored lower on the more cognitively demanding rhyme and head detection task than they did on the initial consonant isolation task. Performance in the two types of task will be discussed in relation to vocabulary scores in the next section.

5.4.2 Intimate Relationships Between Phonological Sensitivity and Vocabulary Scores

5.4.2.i Higher Correlations in Relation to Time Lapses

As with the nonword repetition, the mean score of the five English phonological sensitivity measures was positively correlated with all 15 vocabulary assessments (see Table 14), suggesting that phonological sensitivity is closely associated with young learners' vocabulary knowledge (Bowey, 2001; Cheng, 2005; de Jong et al., 2000; Gathercole et al., 1991). This result demonstrates on one hand that the significant relationships found between phonological processing skills and vocabulary development among alphabetic language users are also manifest in speakers of a first-learned logographic language. On the other hand, it also indicates that phonological sensitivity is a prerequisite to all learners for success in learning the alphabetic language of English. In addition, a comparison of the correlation coefficients between vocabulary scores and the two phonological processing skills suggests that a more intimate relationship existed between phonological sensitivity and vocabulary scores than between phonological memory and vocabulary scores in the present study. The correlation *r* values reaching significance levels obtained for the former two variables ranged from .303 to .730, while those obtained for the latter ranged from .264 to .501.

An interesting pattern of association appears in relation to time lapses in the present study. It was found that the correlation coefficients between phonological sensitivity and written vocabulary assessments became higher when a longer period of time had elapsed between explicit vocabulary teaching and the time of vocabulary testing. This pattern appears to suggest that phonological sensitivity might have exercised a long-term influence over young children's EFL vocabulary learning, so that children with better phonological sensitivity would tend to have retained a greater vocabulary knowledge.

5.4.2.ii Lack of Association Between Pretests and Individual Tasks

When the English phonological processing tasks were examined individually, the positive correlations were largely replicated between vocabulary scores and each individual measure, with minor exceptions (see Table 14), indicating that those who performed well in learning vocabulary were generally good at analysing and manipulating rhymes, heads, and phonemes. This result confirms the findings of other studies that positive correlations existed between vocabulary learning and phonological sensitivity at different size units (Anthony et al., 2002; Bowey, 1996; de Jong et al., 2000).

An examination of the correlation table reveals that the exception to a positive relationship occurred in the pretests for Stories 1 and 3, the pretests being for the most part dissociated from the three detection tasks. This finding seems to confirm the previous result showing a lack of association between vocabulary scores and the second nonword repetition task (see Table 13), which was attributed to the time of testing and a possible change in the participants' vocabulary knowledge. However, taking into account the finding mentioned previously that a stronger association was observed between phonological sensitivity and vocabulary scores as more time elapsed, a different explanation is provided below.

The lack of association between vocabulary pretests and phonological sensitivity might indicate that those learners who performed well in phonological sensitivity measures were not necessarily early starters in learning English and hence did not score high in the vocabulary pretests, in which high scores were largely explained by longer English instruction. However, given the same amount of story instruction, those who had originally had lesser vocabulary knowledge gradually caught up and were able to retain the target vocabulary, as was evidenced by the stronger correlation between their performances in subsequent vocabulary assessments and the phonological sensitivity tasks. This interpretation explains the finding that English instruction length was associated with the participants' vocabulary performance in the pretests, but that instruction time was not associated with their phonological processing skills. It demonstrates that English instruction length contributed to the participants' vocabulary scores but that vocabulary learning still depended on their phonological processing skills.

5.4.2.iii Stronger Association Between Vocabulary Scores and Production Tasks

The other interesting association pattern observed in the individual tasks (see Table 14) was that consistent correlations appeared to exist between vocabulary scores and production tasks. In the two *production* tasks, participants were asked to name the shared rhyme or the shared head in the rhyme and head production task and to name the first consonant of a word in the initial consonant isolation task. In contrast, *detection* tasks required the participants to click the mouse to indicate whether the sound units in a word pair were the same in the designated positions in the three tasks of rhyme detection, head detection, and rhyme and head detection. In a comparison of the two categories, it appears that each individual production task was associated with all the vocabulary assessment results, while each of the three detection tasks was *dis*sociated with at least one vocabulary assessment. This result, showing a consistent association pattern, indicates that the production tasks were more strongly associated with vocabulary scores than the detection tasks. In other words, the production tasks had a greater capacity to distinguish between good and poor learners than the detection tasks in that those young learners who performed well on the two production tasks consistently performed well on *all* vocabulary assessments as well.

The stronger association between vocabulary scores and production tasks could be attributed to the fact that the participants had to give oral responses in the production tasks in contrast to mouse-clicking in the detection tasks, where the participants could simply have guessed the right answers. The production tasks placed more demands on the young learners as they had to demonstrate their skills of analysis and manipulation of speech sounds as well as of speech production. It is likely that oral output might reflect the level of young learners' phonological sensitivity more accurately.

5.4.3 Cross-Language Transfer of Phonological Processing Skills

Intercorrelations were generally observed when the same measures in the two different languages were computed against each other (see Table 15), suggesting that cross-language transfer of phonological processing skills occurred to the young learners. The result that the subjects had scored higher on four Chinese phonological sensitivity tasks (see Table 9 for the percentage rates) demonstrates that their Chinese phonological sensitivity skills were more developed than their English skills and hence were able to transfer the skill from the native language to the foreign language, rather than vice versa. This finding of cross-language transfer of phonological processing skill is consistent with the findings of previous research conducted with groups of ESL learners who had an alphabetic language background (Cisero & Royer, 1995; Comeau et al., 1999; Durgunoglu, Nagy, & Hancin-Bhatt, 1993) and with learners who had a nonalphabetic language background (Gottardo, Yan, Siegel, & Wade-Woolley, 2001). Since the native language decoding skill is related to second language learning among ESL learners (Meschyan & Hernandez, 2002), the cross-language transfer of phonological processing skill might depend highly on the transparency of grapheme-phoneme correspondence rules in their L1 orthography (Koda, 1997; Rickard Liow & Poon, 1998). Although it has often been argued that speakers with a first-learned logographic language background possess phonological processing skills only at the word level and not at the phoneme level, owing to the nature of their native language (Cheung, 1999; Hanley et al., 1999), the findings of the present study indicate the opposite. The strategies of teaching Chinese in Taiwan might account for this unexpected finding.

While strategies of teaching Chinese to young children vary from one region to another, as pointed out earlier, Taiwanese children learn to read Chinese with the assistance of Zhu-Yin-Fu-Hao, a phonetic script which spells the sounds of Chinese characters. As Zhu-Yin-Fu-Hao is put alongside each Chinese character in the texts of Chinese textbooks from the first grade to the fourth grade, and alongside new words from the fifth grade onwards, Taiwanese children are used to combining phonetic symbols in order to pronounce new words. In other words, Zhu-Yin-Fu-Hao has inadvertently taught Taiwanese children to manipulate phonological units at the phoneme level, which enables them to apply their L1 phonological processing skills when tackling English phonological sensitivity tasks.

However, Zhu-Yin-Fu-Hao has its limitations in the cross-language transfer of phonological processing skill, owing to the nature of the Chinese language. While all the items in the Chinese phonological sensitivity tasks were presented in the CVN structure where C can be replaced by a glide, consonants or glides in the onset position seldom exist alone in the spoken language and are often combined with a vowel to form a word structure, with CV and CVN being legitimate phonological representations. Among the 37 phonetic symbols, only seven can have an initial consonant which can be legitimately isolated from the CVN structure. It is possible that the violation of legitimate Chinese phonological structures in the spoken language caused by an isolation of the initial sound might have been the cause of the dissociation between Chinese initial sound isolation and other phonological sensitivity task scores. Despite this dissociation, the results of the English initial consonant isolation task performed by the participants were still correlated with the results of other English phonological sensitivity measures (see Appendix L) as well as with all vocabulary scores (see Table 14), and this result seems to indicate that the cross-language transfer of phonological processing skill is an underlying process that is related to a child's ability to reflect on phonology, even if not specifically his L1 phonology (Gottardo et al., 2001).

In sections 5.3 and 5.4, the two phonological processing skills have been discussed individually in terms of their association with young learners' EFL vocabulary learning. However, the present study also aimed to compare the predictive power of these two skills in relation to EFL vocabulary learning, and the relevant discussion will be presented below.

5.5 A Latent Vocabulary Learning Factor

5.5.1 A Common Component Construct

When eight variables were entered in the principal-component analysis, the function of which is to extract the common underlying dimension shared by the variables, two factors were extracted for the present study (see Table 16). The result that all eight variables loaded highly on Factor 1 suggests that Word Review II, the two nonword repetition tasks, and the five English phonological sensitivity tasks all shared a common construct which linked the two phonological processing skills and vocabulary learning. Among the variables, nonword repetition has been empirically proved to be associated with young children's vocabulary knowledge (Bowey, 2001; Gathercole & Baddeley, 1989; Service & Kohonen, 1995). A variety of tasks designed to tap children's phonological sensitivity, for example, rhyme oddity and phoneme identity, have also been shown to correlate with children's vocabulary learning (de Jong et al., 2000), prereading ability (Burgess & Lonigan, 1998) and reading ability (Gottardo et al., 1996). In addition to the nonword repetition tasks and the English phonological sensitivity tasks, which tested the relationship between phonological processing capabilities and vocabulary knowledge, the Word Review II assessment also represented the children's vocabulary learning outcome. Taken together, the result of this extracted factor could be regarded as a latent vocabulary learning factor and this has also been found in other studies (Gathercole et al., 1991; Metsala & Walley, 1998; Stahl & Murray, 1998).

5.5.2 A Unique Nonword Repetition Factor

An intriguing result was, however, revealed by Factor 2, on which both nonword repetition tasks loaded moderately (see Table 16), suggesting that phonological memory, measured by nonword repetition, constituted a unique nonword repetition factor which was linked weakly with vocabulary learning but not with phonological sensitivity. A similar result was obtained in a study conducted by Gathercole et al.

(1991), in which two factors were extracted, one representing a common phonological processing factor and the other a unique phonological memory factor. The result that both nonword repetition tasks loaded moderately on the two factors in the present study might indicate that nonword repetition is associated with phonological sensitivity and shares the same component construct in terms of vocabulary learning, but that otherwise it is a unique factor unassociated with phonological sensitivity.

5.5.3 Rhyme and Phoneme: A Continuum of Phonological Sensitivity Skills

Also revealed by the factor loadings on Factor 2 was the fact that the four phonological sensitivity tasks measuring large phonological units loaded nonsignificantly, while the initial consonant isolation task which measured phonemic awareness loaded moderately. This result seems to indicate that rhyme sensitivity and phonemic awareness are two independent skills (Muter & Snowling, 1998; Yopp, 1988) and hence argues against the one-factor model, which posits that large-size and small-size phonological sensitivity is an overlapping sensitivity to syllables, rhymes, and phonemes (Anthony et al., 2002; Schatschneider et al., 1999).

However, this result could be interpreted in line with a broader developmental conceptualisation of phonological sensitivity, which suggests that continuity exists between large-size and small-size phonological processing skills (Adams, 1990; Anthony et al., 2002; Goswami, 2002; Goswami & Bryant, 1990). Although findings discussed earlier indicated that the young children performed differentially in the large-unit and small-unit processing tasks and appeared to have better rhyme sensitivity than phonemic processing skills, their performances in the five tasks were intercorrelated with each other (see Appendix L). This indicates that their comparable performance at different levels of phonological sensitivity was a natural occurrence constrained by the developmental trend of phonological processing progression.

5.6 Phonological Sensitivity as Best Predictor of EFL Vocabulary Learning

English phonological sensitivity and Word Review I were chosen as the best predictors of young learners' EFL vocabulary learning in the 14 vocabulary assessments over the other four variables—age, phonological memory, English instruction length, and Chinese phonological sensitivity, which were also entered into the regression model (see section 4.5.1). The interesting factor distinguishing these two predictors was the time of administering these vocabulary assessments. Word Review I was the best predictor of a series of vocabulary assessments conducted at the beginning of each storytelling session, pretests and in-class tests, for example. However, as time progressed and the gap between story sessions and assessment time increased, English phonological sensitivity emerged as the best predictor in the later administered vocabulary assessments (see Appendix M). This result echoes the finding discussed earlier of an intimate association between English phonological sensitivity and vocabulary scores (see section 5.4.2.i), and indicates that phonological sensitivity played a more important role in young children's EFL vocabulary learning as time elapsed, and that retention of target vocabulary items became difficult when they were not recycled.

Since Word Review I is defined as an achievement task representing children's achievement from previous learning, the results of this task were picked as the determinant of children's vocabulary performance in the vocabulary assessments held earlier, because what children had previously learned would naturally constitute their vocabulary knowledge base. The replacement of these results by English phonological sensitivity in those vocabulary assessments administered later proved that the importance of their role in accounting for children's vocabulary learning had diminished. The emergence of phonological sensitivity as the best predictor further demonstrates that those with better phonological sensitivity performance soon caught up with their peers under the new learning situation.

5.6.1 Unique Contribution by Phonological Sensitivity to EFL Vocabulary Learning

As EFL vocabulary learning was accounted for by various variables, a hierarchical regression analysis was used at this stage to help determine the variance of vocabulary scores explained by different variables, i.e., the accountability of each variable in vocabulary assessments, and hence to compare whether phonological memory or English phonological sensitivity contributed uniquely to vocabulary scores after the variance explained by other variables was removed. The first four variables entered into the regression analyses—age, English instruction length, Word Review I, and Chinese phonological sensitivity—jointly accounted for different percentage rates of variance in various vocabulary assessments, some of which explained more than 50% of the variance, indicating that the four variables were importantly associated with young children's EFL vocabulary learning (see section 4.5.2).

When the first nonword repetition task score and the English phonological sensitivity mean score were entered in the next two steps after the four variables, the results indicated that they explained the additional variance in vocabulary scores, but that the variance accounted for by each of them was comparable (see Table 17 & Appendix N). Throughout the vocabulary assessments in the present study, when English phonological sensitivity was entered in the fifth step, it consistently explained more variance (range from 2.5% to 19.2%) in the vocabulary scores than did nonword repetition, which was entered in the sixth step and explained very little variance (range from none to 2.4%). This result shows that English phonological sensitivity explained more variance in vocabulary scores than did phonological memory in the variance that remained after the variance accounted for by the four variables entered first had been removed.

However, such a pattern was not found when nonword repetition was entered prior to phonological sensitivity. When nonword repetition was entered in the fifth step, it explained relatively little variance, ranging from 0.3% to 7.2%, in comparison to the results when phonological sensitivity was entered in the fifth step. Meanwhile, when phonological sensitivity was entered in the sixth step after nonword repetition in order to determine the variance it uniquely explained after the variance explained by all the other five variables had been removed, the results indicated that it still accounted for a relatively large variance, ranging from 1.6% to 14.4%. This result clearly indicates that English phonological sensitivity explained more variance than did phonological memory and hence contributed uniquely to young learners' EFL vocabulary scores. A less distinctive pattern was shown when the mean of the two nonword repetition tasks was entered in the hierarchical regression analyses. However, phonological sensitivity still generally accounted for more variance than did phonological memory and contributed uniquely to vocabulary scores after phonological memory was statistically controlled (see the right panel of Appendix N). This finding confirms the previous studies which showed that phonological sensitivity remained a stronger predictor of young EFL learners' vocabulary scores than phonological memory (de Jong & van der Leij, 1999; Gottardo et al., 1996; Metsala, 1999).

5.6.2 Overlapping Variance by Two Phonological Processing Skills

Judging by the results obtained from the hierarchical multiple regression analyses, phonological memory and phonological sensitivity seem to have overlapped with each other in most tests when explaining the variance in vocabulary scores (Bowey, 1996; de Jong & van der Leij, 1999). When the variance in vocabulary scores explained by the first four variables was removed from the equation, phonological memory and phonological sensitivity were found to have jointly explained the same amount of variance in each vocabulary assessment regardless of which variable was entered first (see Appendix N). In spite of the fact that the explained variance fluctuated between the two variables, they still jointly accounted for the same amount of variance in vocabulary scores in each test and hence should have overlapped in explaining the variance. Taking for example the posttest for Story 3 (see the left panel of Table 17), the two skills jointly explained 21.6% of the variance. When nonword repetition was entered in the fifth step, it accounted for 7.2% of the variance and phonological sensitivity explained 14.4% in the sixth step. But when phonological sensitivity was entered in the fifth step and found to account for 19.2% of the variance, nonword repetition only accounted for 2.4% of the variance in the sixth step. This result indicates that nonword repetition only accounted uniquely for 2.4% of the variance in this posttest, while it could have shared 4.8% (= 7.2% - 2.4%) of the variance with phonological sensitivity, which explained a higher 19.2% (= 14.4% + 4.8%) of the variance accounted for by phonological sensitivity when it was entered prior to nonword repetition. The phenomenon of fluctuating percentage rates of variance accounted for by phonological memory on the one hand and phonological sensitivity on the other, across all vocabulary assessments, is a clear indication that the two skills overlapped with each other in explaining the variance in the young learners' vocabulary scores.

5.6.3 Inflation of Predictive Power of Nonword Repetition

Both stepwise and hierarchical regression analyses have shown that English phonological sensitivity was the best predictor and contributed uniquely to young learners' EFL vocabulary learning after the other variables were statistically controlled. The possibility that the predictive power of nonword repetition with regard to vocabulary learning had been inflated might explain why it appeared to play a far less important role. Despite the claim by Gathercole and her colleagues that nonword repetition is a pure measure of phonological memory, other studies have shown that nonword repetition might reflect a range of phonological processing skills, such as speech perception and production (Bowey, 1997; Dollaghan et al., 1995). Nonword repetition requires a correct phonological representation as well as articulation, and is also influenced by an individual's vocabulary knowledge. With respect to the range of phonological processing skills it involves, nonword repetition is perhaps unfit to be construed as a pure phonological memory test (Muter & Snowling, 1998).

Furthermore, since the factor analysis shows that nonword repetition shared the same component construct with phonological sensitivity, the former might not be as good a predictor of young children's vocabulary performance as other phonological processing measures that do not require a strong memory component. Bowey (1997) commented that the unique contribution of nonword repetition to vocabulary development could only be sustained when "vocabulary is more strongly correlated with nonword repetition than with measures of phonological processing that do not include a strong memory component" (1997, p. 298).

5.6.4 Rhyme and Head Production as the Best Task Predictor

In order to determine the best predictor of vocabulary scores among the phonological processing tasks, the scores for the 10 Chinese and English phonological sensitivity tasks, as well as the scores for the first nonword repetition and the mean of the two nonword repetition tasks, were entered as independent variables in the stepwise regression analyses, while vocabulary scores were held as dependent variables. The results show that the English rhyme and head production task was the best task predictor of the young learners' EFL vocabulary scores and that English initial consonant isolation came in as the second best task predictor (see Appendix O), confirming the previous finding of closer correlations between production tasks and vocabulary scores. As English rhyme and head production has been identified as the most cognitively demanding task among the five English phonological sensitivity tasks, this result seems to imply that the task requiring sophisticated phonological analysis and manipulation as well as an element of speech output was more effective in predicting the children's vocabulary scores.

However, the result also points to a future link with the orthography of the target language. Stahl and Murray (1994) have suggested that onset-rime analysis, rather than vowel-coda analysis, within syllables is a minimal requirement for early reading success, even though phonemic awareness might be a better predictor of individual differences (Hulme et al., 2002; Muter, Hulme et al., 1998). The result that the rhyme and head production task as well as the initial consonant isolation task were chosen as the best predictors of EFL vocabulary learning in the present study shows that the subjects were equipped with the minimal requirement for later reading success.

5.7 Summary of the Chapter

The discussions contained in this chapter have related the experimental statistics derived from the present study to the findings of other relevant research and attempted to explain their causes or effects. Despite the pedagogical method of creating a picture-to-concept link, an investigation into the L1 effects shows that the subjects did not access English meaning conceptually and had to employ L1 translation in the early stage of foreign language learning. This result contradicts the hypothesis of L1 effects (1.5.1.i), but is consistent with the findings of psycholinguistic studies of bilingual memory representations that a bilingual's two languages are connected at the lexical level in the initial stage of learning.

The examination of English instruction length proves that time was a crucial factor in determining the amount of the subjects' vocabulary, since more time allowed them to learn more words, consequently confirming the benefits of an early start to the English curriculum. However, instruction time was dissociated with phonological memory and phonological sensitivity, the two factors positively associated with EFL vocabulary learning. The dissociation between instruction length and vocabulary gains shows that a greater exposure to the target language in the initial stage did not enhance the speed of vocabulary learning.

The positive association between vocabulary knowledge and the two phonological processing skills found in the present study has proved that the factors crucial to the vocabulary development of native child speakers were similarly important to children with a first-learned logographic language background learning English as a foreign language. While the two skills were equally important to EFL vocabulary learning and shared a common construct of vocabulary learning factor, the association between phonological memory and vocabulary scores was found to weaken while the association between phonological sensitivity and vocabulary learning became stronger as more time elapsed between explicit vocabulary teaching and the time of tests. The stronger predictive power of phonological sensitivity among the two skills in predicting EFL children's vocabulary learning was manifested in the finding that it still explained more of the variance in vocabulary scores than phonological memory, and consequently made a unique contribution to the children's EFL vocabulary learning.

Several other findings with respect to the phonological sensitivity skill were also revealed by the study. Firstly, a cross-language transfer of phonological processing skills observed in the present study suggests that children's phonological processing ability might be related to the ways they are taught to read and may not be related specifically to their L1 phonology. Secondly, the development of phonological sensitivity among young EFL learners was in line with the developmental perspective and was found to progress from large units to small units, in the same way as that of native child speakers.

Although the findings discussed in this chapter indicate the inevitability of the route of L1 translation and the importance of phonological processing skills to EFL vocabulary learning, and thus have implications for the teaching of English as a foreign language, there are also certain limitations to their application in the EFL classroom. Both implications and limitations will be discussed in the final chapter.

CHAPTER 6: CONCLUSION

6.1 Introduction

The present study set out to investigate the factors which are favourable to the vocabulary acquisition of young learners of English as a foreign language. Four factors believed to influence vocabulary learning qualitatively and quantitatively were examined in the study. These are L1 translation, length of English instruction, phonological memory, and phonological sensitivity. The first two factors are termed extrinsic factors since use of the L1 assists EFL learners' memorisation and retention of vocabulary and is in-differentially available to all EFL learners. At the same time, greater exposure to English in an EFL context has been proved to modulate learners' performance quantitatively, especially in the early stage of L2 learning. These factors are contrasted with the other two factors, which are referred to as intrinsic factors, since phonological memory and phonological sensitivity are related to an individual's vocabulary learning ability and vary from one learner to another.

In the present study, phonological memory was measured by nonword repetition tasks and phonological sensitivity by five different tasks designed to tap the children's large- and small-size phonological processing skills. In addition to these tasks, computerised and paper-and-pencil vocabulary assessments were conducted to investigate individual learners' differences in vocabulary learning in relation to their phonological processing skills. As a large vocabulary size in an EFL learner with a nonalphabetic language background is often attributed to an early start and greater exposure to the target language, length of L2 instruction was examined with respect to the two phonological processing skills in an attempt to determine which factor best predicted the young learners' vocabulary learning. Various methods of statistical analysis, including *t*-tests, correlation analyses, a factor analysis, and hierarchical and stepwise regression analyses, were used to investigate the relationships between these key factors and young learners' EFL vocabulary acquisition.

6.2 Answers to the Research Questions

6.2.1 Answers to the Main Research Question

What is the best predictor of young learners' successful acquisition of receptive vocabulary in an EFL context?

As the results from the present study show, English phonological sensitivity was the best predictor of young EFL learners' concurrent vocabulary performance. The greater power of phonological sensitivity to predict young learners' EFL vocabulary acquisition may be seen in the hierarchical regression analysis, where it was demonstrated to account for additional variance in vocabulary scores after a large share of variance explained by the other five variables entered beforehand had been removed. Phonological memory, which was measured by nonword repetition, by contrast, did not explain as much of the variance in vocabulary scores as phonological sensitivity after the other variables were statistically controlled.

6.2.2 Answers to the Subquestions

Subquestion 1: Is the use of L1 a boost or a hindrance to young learners' acquisition of English vocabulary in the foreign language classroom?

Generally, the use of L1 was more of a boost than a hindrance to these young EFL learners, as was evidenced by both the amount of vocabulary gains in written vocabulary assessments and the reaction times in the online vocabulary assessments. The control group, who were given an extra input of L1 translation equivalents to English lexical items in explicit vocabulary teaching, consistently responded faster than the experimental group, who were given instruction only in the target language, despite the fact that the between-group differences in the reaction times were not consistently significant. In terms of vocabulary learning, the control group had higher vocabulary gains made between two vocabulary assessments and hence showed a better retention of newly learned vocabulary items. Taken together, these results suggest that L1 use is beneficial to learners in the initial stage of L2 vocabulary learning.

Subquestion 2: Does a longer length of English instruction result in better acquisition of English vocabulary? Is it helpful to young learners' phonological processing *skills?*

The pupils who reported having received extracurricular English instruction consistently performed better in the vocabulary assessments than those who had received less or no such instruction, as was evidenced by the positive association between English instruction length and vocabulary scores. This result clearly indicates that English vocabulary knowledge was associated with the amount of instruction the young learners had received, suggesting that time is an important factor in EFL vocabulary learning.

However, longer English instruction did not appear to enhance phonological processing skills, as a complete lack of association was observed between vocabulary scores and the subjects' performance in the tasks relating to the two skills. In other words, the advantage of having a larger vocabulary in the initial stage of EFL learning as a result of longer instruction length is likely to disappear in the wake of a later intensive exposure to the target language.

Subquestion 3: Is nonword repetition a good measure to predict young learners' vocabulary acquisition? What is the association between nonword repetition and receptive vocabulary?

A large body of research supports the power of nonword repetition in predicting native child speakers' receptive vocabulary, and it was also found to be a good measure for predicting the vocabulary acquisition of young EFL learners in the present study. Used to measure the young learners' phonological memory, nonword repetition was found to be positively associated with their vocabulary scores, indicating that those children with a better phonological memory also scored higher in the vocabulary assessments. However, the decline in this association between nonword repetition and vocabulary scores seems to suggest that its predictive power waned over the course of the study.

Subquestion 4: Is phonological sensitivity a good predictor of young learners' receptive vocabulary acquisition? What is the association between phonological sensitivity and receptive vocabulary?

Being a good predictor of native child speakers' literacy skills, phonological sensitivity was equally effective in predicting young EFL learners' vocabulary performance. Positive correlations were found between the subjects' vocabulary scores and their performance in phonological sensitivity tasks. The finding that it was the best predictor of young EFL learners' vocabulary learning outcome in the present study was confirmed by the stronger association between vocabulary and phonological sensitivity task scores as more time elapsed between explicit vocabulary teaching and testing times, in contrast to the decline in the association between the subjects' vocabulary and phonological memory performances. Among the phonological sensitivity tasks employed in the present study, the rhyme and head production task, which required an output of speech and was the most cognitively demanding, was the best task predictor of children's vocabulary scores.

Development of phonological processing skills from large- to small-size units was observed among the young EFL learners, as they performed better in the large-size tasks than in the small-size tasks of the same difficulty level, which is consistent with the developmental perspective of phonological processing skills observed among native child speakers.

Owing to the specific strategies of teaching Chinese in Taiwan, a trend of cross-language transfer of phonological processing skills, as was evidenced by the intercorrelations between tasks in the two different languages, was also observed in the present study. The result appears to suggest an underlying process for the young learners with a first-learned logographic language to reflect on all phonology and not specifically on their L1 phonology.

Subquestion 5: Do these two aspects of phonological processing skills explain independent or overlapping variation in young children's acquisition of receptive vocabulary?

The principal-component analysis showed that phonological memory and phonological sensitivity shared the same component construct with vocabulary learning in the present study. Further hierarchical regression analyses showed that the two aspects of phonological processing skills jointly explained the same amount of variance in vocabulary scores, while the overlapping variance was demonstrated by the fluctuating percentage rates of variance between the two factors. However, phonological sensitivity still explained more variance than did nonword repetition in young learners' vocabulary scores and hence can be seen to have made a unique contribution after the other variables were statistically controlled.

6.3 Findings Related to the Hypotheses

6.3.1 Hypothesis of First Language Effects

It was hypothesised that young learners who were given L1 translation equivalents to the target lexical items would need longer reaction times in response to auditory and visual cues in the online vocabulary assessments than those who were offered no L1 translation.

The finding from the present study shows that the group who were given L1 translation equivalents in explicit vocabulary teaching, to the contrary of the hypothesis, had shorter reaction times than the other group who were pedagogically instructed to map FL phonological representations directly onto pictures. The interview which took place immediately after the online vocabulary assessments reveals that the two groups of learners, irrespective of the pedagogical difference, employed L1 translation to help access the FL word meaning. A further analysis of the online processing behaviour shows that the difference in reaction times between the two groups occurred probably because the group who were not provided L1 glosses might have spent milliseconds more in locating the L1 translation of the auditory cue and hence showed slower reaction times. This finding is consistent with the findings of bilingual memory research which revealed that the two languages of L2 learners are connected at the lexical level in the initial stage of learning.

6.3.2 Hypothesis With Respect to Relationships Between EFL Vocabulary Learning and Intrinsic Factors

It was hypothesised that the two phonological processing capabilities—phonological memory and phonological sensitivity—crucial to vocabulary development among native child speakers, were similarly important to young EFL learners' vocabulary learning.

This hypothesis is proved correct by the finding that the two phonological processing skills were positively correlated with the young EFL learners' vocabulary scores, suggesting that young learners who performed well in vocabulary assessments had good phonological memory and phonological sensitivity. However, in terms of the power to predict EFL vocabulary learning, English phonological sensitivity was a better predictor because it explained more of the variance in vocabulary scores than did phonological memory and was thus found to make a unique contribution to the children's EFL vocabulary learning. This result reflects the association patterns between the two skills and the vocabulary scores: The association between phonological memory and vocabulary scores became weaker, while the association between phonological sensitivity and vocabulary learning became stronger as more time elapsed between explicit vocabulary teaching and the times of testing.

6.4 Implications of the Present Study

6.4.1 Provision of L1 Glosses to Facilitate EFL Vocabulary Learning

The computerised aspect of the research design of the present study has shown that taking the L1 route is likely to be inevitable in EFL beginners, whatever the pedagogical differences. This finding is consistent with the body of research which suggests that learners' L2 is connected with their L1 at the lexical level in the early stages of learning a foreign language. As use of the L1 was proved to have facilitated the children's FL vocabulary learning, there is no reason to block its use in the classroom completely. In fact, the approach in which English is taught often encourages children to think of the foreign language as a set of words (Cameron, 2001). Consequently, Vygotsky's theory for language learning which stresses the importance of the word as unit might help in constructing a theoretical framework for teaching English to the children in an EFL context.

In Taiwan, L1 has long been used as an aid for comprehending L2 in classroom instruction. Its wide use has attracted criticisms that the amount of English input is consequently reduced to a minimum in the very limited instruction time available. White and Genesee (1996) once commented that they did not consider ESL instruction in Quebec, Canada, to constitute significant exposure. The reasons they listed included the use of the children's L1 as the medium of instruction, limited instruction hours, minimal exposure outside the ESL classroom and the low English proficiency of nonnative teachers. These happen to be the same shortcomings facing Taiwan's EFL teaching. Use of the L2 in the EFL classroom is crucial for the provision of a language model and opportunities for learners to practise the foreign language. Despite the finding that taking the L1 route is inevitable in beginning language learners, the literature also points out that an ultimate switch to conceptual mediation is absolutely essential if one is to achieve fluency in L2, as thinking in L2 cuts off L1 translation and results in no delay in L2 processing. Given the fact that humans have a limited capacity for information processing, it is necessary for language users to pay attention to higher levels of linguistic information in communicative situations. Conceptual mediation will hence eventually be necessary for this purpose.

It is perhaps, in this case, more realistic to reflect on how to take advantage of the L1's inevitable presence in EFL learners' minds in order to strengthen the connection between a familiar L1 word and its FL equivalent in the initial stages of EFL learning. L1 glosses could be readily provided to the learners in the textbooks so that the L1-L2 connection is strengthened. However, as previously stated, the ultimate replacement of the L1 by the target language is necessary so that learners will learn to switch their mode of language processing to L2 for their own benefit as they progress. As research on speakers of an alphabetic language background shows that it takes two to three years for them to start mediating an alphabetic L2 conceptually, more research is needed to find out whether the same amount of time applies to speakers of a first-learned logographic language in an EFL context.

6.4.2 Inclusion of English Sensitivity Training in Elementary English Education

While Carroll (1969, p. 57) sees it as "a matter of conscience", there is no easy answer to the question as to at what age English teaching should begin. That English is one of the most important foreign languages taught in Taiwan is manifested by the phenomenon that many young children are sent to extracurricular lessons. The introduction of English as a subject required in national examinations also speaks volumes. However, although many parents believe that the age of learning a foreign language should be advanced to as young as possible, so that their children might learn more English to ensure their successful English acquisition before the critical period for language learning terminates, the present study shows that an early start for a child did not necessarily guarantee a lasting advantage. On one hand, a longer period of English instruction was not found to promote the phonological processing skills related to new word learning. On the other hand, the advantage of a slightly larger vocabulary as a result of an early start to acquire L2 vocabulary is likely to be only short-lived, because the late-starters often catch up with the early-starters within a few years (Oller & Nagato, 1974).

Nevertheless, the present study has revealed that the component abilities which ensure successful literacy development among children who are native speakers of English are similarly important to children learning English as a foreign language. This implies that if English teaching is advanced to an early age, efforts should be made to promote these component abilities in the children in their English education at elementary school level so that these abilities will assist them in further language learning.

6.4.3 L2 Learners as Unique Individuals

The findings on how these young learners stored and processed EFL vocabulary and how individual differences in their phonological processing skills affected their vocabulary acquisition demonstrate that they did not learn a new language as did its native speakers and are consequently compatible with Cook's (2001) portrait of L2 users in the multi-competence model. The finding from the experiment on the L1 effects suggests that the first language should be recognized in language teaching if the languages of learners are always linked in the mind. Meanwhile, the finding from the development of their phonological processing skills and a possible cross-language transfer of such skills further suggests that the overall knowledge they possessed—their two first languages and the L2 interlanguage in the same mind—might have had an effect on the outcome of their EFL vocabulary learning. Taking into account the fact that L2 learners are not native speakers and never can be, the model recognises the distinctive nature of the L2 learners and considers them in their own right.

6.5 Original Contributions to SLA Knowledge

This present study has made contributions to knowledge on young learners' EFL vocabulary learning, specifically on the group of learners with a first-learned nonalphabetic language background. While similar research has been done in EFL contexts, those young learners enjoy the advantage of having a phonemically-based first language. The findings from the present study reveal that the variables which facilitate native child speakers' vocabulary acquisition perform similar functions on Taiwanese children. While lowering the age to start English instruction allows young learners more time to learn the language, teaching sound awareness, which is likely to be a prerequisite for learning an alphabetic language, might serve as an alternative to allow more time for native language education.

The other contribution is to do justice to the status of the L1 in young learners' initial stages of EFL vocabulary learning. The result shows that a provision of L1 glosses might be as effective as a provision of picture aids when they process and store English lexical items. This finding is compatible with the research in bilingual lexicon representations that the L2 beginners have their two languages interconnected at the lexical level.

6.6 Limitations of the Present Study

Although the results of this study indicate that phonological processing abilities are equally important to young EFL learners as they are to native child speakers of English in the development of literacy skills, the application of this result is confined to the learning of the phonological forms of new words, i.e. the prereading stage of EFL learning. The other component skills of literacy, such as letter knowledge and word reading, which are crucial to advanced literacy skills, were not included in the assessments for fear that such assessments might create more complications, since the children's familiarity with the English alphabet is generally enhanced by the length of their extracurricular instruction. The other shortcoming of using Chinese phonological sensitivity tasks as precursors to the English tasks is that no assessment could be conducted to test consonant clusters, the phonological feature which is distinctive in English phonology but lacking in Chinese.

In addition, due to the small sample of subjects, the findings of the present study are best suited to predict the concurrent learning result of these EFL students and could be generalised only to their counterparts of the same school year in a similar context where limited English is taught. The results might be subject to changes caused by more variables if they were to predict the learning result of the next stage of word recognition (Torgesen & Burgess, 1998).

6.7 Suggestions for Further Research

As the present study has revealed the importance of phonological sensitivity in young EFL learners' learning of new words, further research is suggested, first to replicate the results of this study to ensure its validity. As various phonological sensitivity tasks might tap different levels of linguistic complexity, a variety of tasks using different tapping methods should be attempted in further research in order to identify the best task predictor suitable for children of different ages and with different experiences in reading instruction. An inclusion of tasks on segmenting consonant clusters is recommended.

A younger group of subjects who have received minimal instruction in Zhu-Yin-Fu-Hao but have not yet received English instruction could be recruited to further examine the claim on the possibility of a cross-language transfer of phonological processing skills from L1 to L2 and not vice versa. It is likely that this group of younger learners will be able to manipulate their L1 sounds but not their L2 if the transfer occurs from L1 to L2. Learners who have never been instructed in Zhu-Yin-Fu-Hao might pose problems in experimental instructions.

6.8 Conclusion

This study represents an initial attempt to compare in EFL children the predictive power of the two phonological processing abilities which are empirically proved to facilitate the vocabulary development of native child speakers. By investigating the individual differences in the young EFL learners' vocabulary acquisition, the study has provided insights into the factors favourable to EFL learners. In addition, the study has also attempted to examine the effect of L1 use in EFL word learning with respect to the bilingual lexico-semantic memory. The result offers an alternative angle from which to re-consider both the status of the L1 in the EFL classroom and the design of teaching materials. It is hoped that this study has contributed to and offered insights into the area of research concerned with EFL vocabulary learning among young learners with a first-learned logographic language background.

REFERENCES

Adams, A.-M., & Gathercole, S. E. (2000). Limitations in working memory: Implications for language development. *International Journal of Language & Communication Disorders. 35*, 95-116.

Adams, M. J. (1990). *Beginning to read: Thinking and learning about print.* Cambridge, MA: The MIT Press.

Aitchison, J. (2003). *Words in the mind: An introduction to the mental lexicon.* (3rd ed.). Oxford: Blackwell.

Alderson, J. C. (1984). Reading in a foreign language: A reading problem or a language problem? In J. C. Alderson & A. H. Urquhart (Eds.), *Reading in a foreign language* (pp. 1-24). London: Longman.

Anthony, J. L., Lonigan, C. J., Burgess, S. R., Driscoll, K., Phillips, B. M., & Cantor, B. G. (2002). Structure of preschool phonological sensitivity: Overlapping sensitivity to rhyme, words, syllables, and phonemes. *Journal of Experimental Child Psychology, 82*, 65-92.

Asher, J. J., & Garcia, R. (1969). The optimal age to learn a foreign language. *The Modern Language Journal, 53*, 334-341.

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

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., Papagno, C., & Vallar, G. (1988). When long-term learning depends on short-term storage. *Journal of Memory and Language, 27*, 586-596.

Beck, I. L., McKeown, M. G., & Omanson, R. C. (1987). The effects and uses of diverse vocabulary instructional techniques. In M. G. McKeown & M. E. Curtis (Eds.), *The Nature of vocabulary acquisition* (pp. 147-162). Hillsdale, NJ: Lawrence Erlbaum.

Beck, I. L., Perfetti, C. A., & McKeown, M. G. (1982). Effects of long-term vocabulary instruction on lexical access and reading comprehension. *Journal of Educational Psychology, 74*, 506-521.

Bensoussan, M. (1992). Learners' spontaneous translations in an L2 reading comprehension task: Vocabulary knowledge and use of schemata. In P. J. L. Arnaud & H. Béjoint (Eds.), *Vocabulary and applied linguistics* (pp. 102-112). London: Macmillan Academic and Professional Ltd.

Bialystok, E. (1991). Introduction. In E. Bialystok (Ed.), *Language processing in bilingual children* (pp. 1-9). Cambridge: Cambridge University Press.

Bialystok, E. (2001). *Bilingualism in development: Language, literacy, & cognition.* Cambridge: Cambridge University Press.

Bialystok, E., Majumder, S., & Martin, M. M. (2003). Developing phonological awareness: Is there a bilingual advantage? *Applied Psycholinguistics, 24,* 27-44.

Biemiller, A. (2003). Vocabulary: needed if more children are to read well. *Reading Psychology, 24,* 323-335.

Block, C. C. (1997). *Teaching the language arts* (2nd ed.). Boston: Allyn and Bacon.

Bowey, J. A. (1994). Phonological sensitivity in novice readers and nonreaders. *Journal of Experimental Child Psychology, 58,* 134-159.

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. (1997). What does nonword repetition measure? A reply to Gathercole and Baddeley. *Journal of Experimental Child Psychology, 67,* 295-301.

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

Bowey, J. A. (2002). Reflections on onset-rime and phoneme sensitivity as predictors of beginning word reading. *Journal of Experimental Child Psychology, 82,* 29-40.

Bowey, J. A., Cain, M. T., & Ryan, S. M. (1992). A reading-level design of phonological skills underlying fourth-grade children's word reading difficulties. *Child Development, 63,* 999-1011.

Bradley, L., & Bryant, P. E. (1983). Categorizing sounds and learning to read—a causal connection. *Nature, 301*, 419-421.

Brady, S. A., & Shankweiler, D. P. (1991). (Eds.). *Phonological processes in literacy: A tribute to Isabelle Y. Liberman*. Hillsdale, NJ: Lawrence Erlbaum.

Braze, D., Tabor, W., Shankweiler, D. P., & Mencl, W. E. (2007). Speaking up for vocabulary: Reading skill differences in young adults. *Journal of Learning Disabilities, 40*, 226-243.

Brett, A., Rothlein, L., & Hurley, M. (1996). Vocabulary acquisition from listening to stories and explanations of target words. *The Elementary School Journal, 96*, 415-422.

British Association for Applied Linguistics. (2006). *Recommendations on Good Practice in Applied Linguistics* (2nd ed.). Retrieved August 20, 2007, from http://www.baal.org.uk/about_goodpractice_full.pdf.

Brown, G. D. A., & Hulme, C. (1996). Nonword repetition, STM, and word age-of-acquisition: A computational model. In S. E. Gathercole (Ed.), *Models of short-term memory* (pp. 129-148). Hove: Psychology Press.

Bryant, P. E., & Goswami, U. (1987). Phonological awareness and learning to read. In J. R. Beech & A. M. Colley (Eds.), *Cognitive approaches to reading* (pp. 213-243). New York: John Wiley & Sons.

Burgess, S. R., & Lonigan, C. J. (1998). Bidirectional relations of phonological sensitivity and prereading abilities: Evidence from a preschool sample. *Journal of Experimental Child Psychology*, *70*, 117-141.

Burstall, C., Jamieson, M., Cohen, S., & Hargreaves, M. (1974). *Primary French in the balance*. Windsor: NFER Publishing Co.

Cameron, L. (2001). *Teaching languages to young learners*. Cambridge: Cambridge University Press.

Cameron, L. (2002). Measuring vocabulary size in English as an additional language. *Language Teaching Research, 6*, 145-173.

Caravolas, M., & Bruck, M. (1993). The effect of oral and written language input on children's phonological awareness: A cross-linguistic study. *Journal of Experimental Child Psychology*, *55*, 1-30.

Carey, S. (1978). The child as word learner. In M. Halle, J. Bresnan, & G. Miller (Eds.), *Linguistic theory and psychological reality* (pp. 264-293). Cambridge, MA: The MIT Press.

Carroll, J. B. (1969). Psychological and educational research into second language teaching to young children. In H. H. Stern (Ed.), *Languages and the young school child* (pp. 56-68). London: Oxford University Press.

Čeponienė, R., Service, E., Kurjenluoma, S., Cheour, M., & Näätänen, R. (1999).

Children's performance on pseudoword repetition depends on auditory trace quality: Evidence from event-related potentials. *Developmental Psychology, 35*, 709-720.

Chang, S.-M. (2005). *The effects of using glosses in facilitating English vocabulary learning and reading comprehension.* Unpublished master's dissertation, Southern Taiwan University, Taiwan.

Chao, C.-K. (2004). *Effects of vocabulary level and syntactic competence on English reading comprehension of EFL English-major students in Taiwan.* Unpublished master's dissertation, Leader University, Taiwan.

Chen, C.-Y. (2005). *A decade of changes in the English ability of university freshmen in Taiwan.* Unpublished master's dissertation, Tunghai University, Taiwan.

Chen, H.-C. (2005). *The teaching methods reflected in junior high school English textbooks*. Unpublished master's dissertation, National Kaohsiung Normal University, Taiwan.

Chen, H.-C., & Leung, Y.-S. (1989). Patterns of lexical processing in a nonnative language. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 15*, 316-325.

Chen, J.-H. (2004). *The effectiveness of the memory of teaching and learning* *strategies in English vocabulary of senior high EFL learners in Taiwan.* Unpublished master's dissertation, National Kaohsiung Normal University, Taiwan.

Chen, P.-C. (2005). *Perceptual learning style preferences among Taiwanese EFL junior high school students.* Unpublished master's dissertation, National Taiwan Normal University, Taiwan.

Chen, Y.-J. (2003). *The effects of phonological decoding training on English word recognition skills in Taiwanese EFL children.* Unpublished master's dissertation, National Taipei University of Education, Taiwan.

Chen, Y.-M. (2004). *The effect of marginal glosses on reading comprehension and vocabulary learning.* Unpublished master's dissertation, National Taiwan Normal University, Taiwan.

Chen, Y.-P. (2004). *Investigation of 7th graders' EFL proficiency discrepancy in relation to their family background.* Unpublished master's dissertation, National Kaohsiung Normal University, Taiwan.

Chen, Y.-Y. (2004). *The effect of early start in English learning on senior high school students' receptive skills.* Unpublished master's dissertation, National Kaohsiung Normal University, Taiwan.

Cheng, M.-C. (2005). *An investigation into English vocabulary learning strategies* *used by junior high school students in Taiwan.* Unpublished master's dissertation, National Chengchi University, Taiwan.

Chern, C. L. (1993). Chinese students' word-solving strategies in reading in English. In T. Huckin, M. Haynes, & J. Coady (Eds.), *Second language reading and vocabulary learning* (pp. 67-85). Norwood, NJ: Ablex.

Cheung, H. (1999). Improving phonological awareness and word reading in a later learned alphabetic script. *Cognition, 70,* 1-26.

Chien, L.-C. (2002). *A developmental study on phonological awareness, spelling and reading in Taiwanese EFL children.* Unpublished master's dissertation, National Taipei University of Education, Taiwan.

Chuang, S.-Y. (2004). *The effect of marginal glosses upon junior high school students' incidental vocabulary learning through a simplified novel.* Unpublished master's dissertation, National Kaohsiung Normal University, Taiwan.

Chou, C.-T. (1989). *A study on the effect of the early start in learning English as reflected in children's later English competence* (Research Rep. No. NSC78-0301-H003-011). Taipei, Taiwan: National Science Council.

Cisero, C. A., & Royer, J. M. (1995). The development and cross-language transfer of phonological awareness. *Contemporary Educational Psychology, 20,* 275-303.

Coady, J., Magoto, J., Hubbard, P., Graney, J., & Mokhtari, K. (1993). High frequency vocabulary and reading proficiency in ESL readers. In T. Huckin, M. Haynes, & J. Coady (Eds.), *Second language reading and vocabulary learning* (pp. 217-228). Norwood, NJ: Ablex.

Collins, A. M., & Loftus, E. F. (1975). A spreading-activation theory of semantic processing. *Psychological Review, 82*, 407-428.

Comeau, L., Cormier, P., Grandmaison, E., & Lacroix, D. (1999). A longitudinal study of phonological processing skills in children learning to read in a second language. *Journal of Educational Psychology, 91*, 29-43.

Cook, V. (1990). Timed comprehension of binding in advanced L2 learners of English. *Language Learning, 40*, 557-599.

Corder, S. P. (1994). A role for the mother tongue. In S. M. Gass & L. Selinker (Eds.), *Language transfer in language learning* (pp. 18-31). Philadelphia: John Benjamins.

Cummins, J. (1991). Interdependence of first- and second-language proficiency in bilingual children. In E. Bialystok (Ed.), *Language processing in bilingual children* (pp. 70-89). Cambridge: Cambridge University Press.

de Gelder, B., Vroomen, J., & Bertelson, P. (1993). The effects of alphabetic reading competence on language representation in bilingual Chinese subjects. *Psychological Research, 55*, 315-321.

de Groot, A. M. B., & van Hell, J. G. (2005). The learning of foreign language vocabulary. In J. F. Kroll & A. M. B. de Groot (Eds.), *Handbook of bilingualism: Psycholinguistic approaches* (pp. 9-29). Oxford: Oxford University Press.

de Jong, P. F., Seveke, M., & van Veen, M. (2000). Phonological sensitivity and the acquisition of new words in children. *Journal of Experimental Child Psychology, 76*, 275-301.

de Jong, P. F., & van der Leij, A. (1999). Specific contributions of phonological abilities to early reading acquisition: Results from a Dutch latent variable longitudinal study. *Journal of Educational Psychology, 91*, 450-476.

DeKeyser, R. M. (2000). The robustness of critical period effects in second language acquisition. *Studies in Second Language Acquisition, 22*, 499-533.

Dollaghan, C. (1985). Child meets word: "Fast mapping" in preschool children. *Journal of Speech and Hearing Research, 28*, 449-454.

Dollaghan, C. (1987). Fast mapping in normal and language-impaired children. *Journal of Speech and Hearing Disorders, 52*, 218-222.

Dollaghan, C. (1998). Spoken word recognition in children with and without specific language impairment. *Applied Psycholinguistics, 19*, 193-207.

Dollaghan, C., Biber, M., & Campbell, T. (1993). Constituent syllable effects in a nonsense-word repetition task. *Journal of Speech and Hearing Research, 36*, 1051-1054.

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

Dufva, M., & Voeten, M. J. M. (1999). Native language literacy and phonological memory as prerequisites for learning English as a foreign language. *Applied Psycholinguistics, 20*, 329-348.

Durgunoglu, A. Y., Nagy, W., & Hancin-Bhatt, B. (1993). Cross-language transfer of phonological awareness. *Journal of Educational Psychology, 85*, 453-465.

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.

Elley, W. B. (1989). Vocabulary acquisition from listening to stories. *Reading Research Quarterly, 24*, 174-187.

Ellis, G., & Brewster, J. (1991). *The storytelling handbook: A guide for primary teachers of English.* London: Penguin Group.

Ellis, N. C., & Beaton, A. (1993). Factors affecting the learning of foreign language vocabulary: Imagery keyword mediators and phonological short-term memory. *The Quarterly Journal of Experimental Psychology, 46A*, 533-558.

Ellis, N. C., & Hennelly, R. A. (1980). A bilingual word-length effect: Implications for intelligence testing and the relative ease of mental calculation in Welsh and

English. *British Journal of Psychology, 71*, 43-51.

Ellis, N. C., & Sinclair, S. G. (1996). Working memory in the acquisition of vocabulary and syntax: Putting language in good order. *Quarterly Journal of Experimental Psychology, 49A*, 234-250.

Ervin, S. M., & Osgood, C. E. (1954). Second language learning and bilingualism. *Journal of Abnormal Social Psychology, Supplement*, 139-146.

Favreau, M., & Segalowitz, N. S. (1983). Automatic and controlled processes in the first- and second-language reading of fluent bilinguals. *Memory and Cognition, 11*, 565-574.

Fernald, A., Swingley, D., & Pinto, J. P. (2001). When half a word is enough: Infants can recognize spoken words using partial phonetic information. *Child Development, 72*, 1003-1015.

Field, A. (2000). *Discovering statistics using SPSS for Windows*. London: Sage Publications.

Flege, J. E., Munro, M. J., & MacKay, I. R. A. (1995). Factors affecting strength of perceived foreign accent in a second language. *Journal of the Acoustical Society of America, 97*, 3125-3134.

Gathercole, S. E. (1995). Is nonword repetition a test of phonological memory or long-term knowledge? It all depends on the nonwords. *Memory and Cognition,* *23*, 83-94.

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., & 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. (1990). 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. (1993). *Working memory and language.* Hove: Lawrence Erlbaum.

Gathercole, S. E., & Baddeley, A. D. (1997). Sense and sensitivity in phonological memory and vocabulary development: A reply to Bowey (1996). *Journal of Experimental Child Psychology, 67*, 290-294.

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., & Pickering, S. J. (1999). Estimating the capacity of phonological short-term memory. *International Journal of Psychology, 34*, 378-382.

Gathercole, S. E., Service, E., Hitch, G. J., Adams, A.-M., & Martin, A. J. (1999). Phonological short-term memory and vocabulary development: Further evidence on the nature of the relationship. *Applied Cognitive Psychology, 13*, 65-77.

Gathercole, S. E., & Thorn, A. S. C. (1998). Phonological short-term memory and foreign language learning. In A. F. Healy & L. E. Bourne, Jr. (Eds.), *Foreign language learning: Psycholinguistic studies on training and retention* (pp. 141-158). Mahwah, NJ: Lawrence Erlbaum.

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.

Gelman, S. A., & Ebeling, K. S. (1998). Shape and representational status in children's early naming. *Cognition, 66*, 35-47.

Goswami, U. (2002). In the beginning was the rhyme? A reflection on Hulme, Hatcher, Nation, Brown, Adams and Stuart (2002). *Journal of Experimental Child Psychology, 82*, 47-57.

Goswami, U., & Bryant, P. (1990). *Phonological skills and learning to read.* Hove: Lawrence Erlbaum.

Gottardo, A., Stanovich, K. E., & Siegel, L. S. (1996). The relationships between phonological sensitivity, syntactic processing, and verbal working memory in the reading performance of third-grade children. *Journal of Experimental Child Psychology, 63*, 563-582.

Gottardo, A., Yan, B., Siegel, L. S., & Wade-Woolley, L. (2001). Factors related to English reading performance in children with Chinese as a first language: More evidence of cross-language transfer of phonological processing. *Journal of Educational Psychology, 93*, 530-542.

Goulden, R., Nation, I. S. P., & Read, J. (1990). How large can a receptive vocabulary be? *Applied Linguistics, 11*, 341-363.

Guion, S. G., Flege, J. E., Liu, S. H., & Yeni-Komshian, G. H. (2000). Age of learning effects on the duration of sentences produced in a second language. *Applied Psycholinguistics, 21*, 205-228.

Gupta, P., & MacWhinney, B. (1997). Vocabulary acquisition and verbal short-term memory: Computational and neural bases. *Brain and Language, 59*, 267-333.

Hakuta, K., Butler, Y. G., & Witt, D. (2000). *How long does it take English learners to attain proficiency?* University of California Linguistic Minority Research Institute Policy Report 2000-1. Retrieved August 20, 2007, from http://caselinks.education.ucsb.edu/casetrainer/CLADContent/CladLanguage/no de7/theory/HowLong.pdf.

Hanley, J. R., Tzeng, O., & Huang, H.-S. (1999). Learning to read Chinese. In M. Harris & G. Hatano (Eds.), *Learning to read and write: A cross-linguistic perspective* (pp. 173-195). Cambridge: Cambridge University Press.

Hansen, J., & Bowey, J. A. (1994). Phonological analysis skills, verbal working memory, and reading ability in second-grade children. *Child Development, 65*, 938-950.

Haynes, M., & Baker, I. (1993). American and Chinese readers learning from lexical familiarization in English text. In T. Huckin, M. Haynes, & J. Coady (Eds.), *Second language reading and vocabulary learning* (pp. 130-152). Norwood, NJ: Ablex.

Hazenberg, S., & Hulstijn, J. H. (1996). Defining a minimal receptive secnd-language vocabulary for non-native university students: An empirical investigation. *Applied Linguistics, 17*, 145-163.

Hernandez, A., Li, P., & MacWhinney, B. (2005). The emergence of competing modules in bilingualism. *Trends in Cognitive Sciences, 9*, 220-225.

Holm, A., & Dodd, B. (1996). The effect of first written language on the acquisition of English literacy. *Cognition, 59*, 119-147.

Hong, Y-F. (2005). *The effect of L1 vocabulary glosses on EFL elementary school* *students' reading comprehension and reading process.* Unpublished master's dissertation, National Chung Cheng University, Taiwan.

Hsieh, D. (2005). *The impact of vocabulary translation on elementary school EFL learners' comprehension and vocabulary learning.* Unpublished master's dissertation, National Taipei University of Education, Taiwan.

Huang, S.-H. (2006). *A study of the correlation between early start and subsequent performance in English among senior high school students in Taiwan.* Unpublished master's dissertation, National Kaohsiung Normal University, Taiwan.

Huang, Y.-C. (2002). *The effects of vocabulary glosses and example sentences on junior high school EFL students' reading comprehension and vocabulary learning.* Unpublished master's dissertation, National Cheng Kung University, Taiwan.

Huckin, T., & Coady, J. (1999). Incidental vocabulary acquisition in a second language. *Studies in Second Language Acquisition, 21*, 181-193.

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., Hatcher, P. J., Nation, K., Brown, A., Adams, J., & Stuart, G. (2002).

Phoneme awareness is a better predictor of early reading skill than onset-rime awareness. *Journal of Experimental Child Psychology, 82*, 2-28.

Hulstijn, J. H., Hollander, M., & Greidanus, T. (1996). Incidental vocabulary by advanced foreign language students: The influence of marginal glosses, dictionary use, and reoccurrence of unknown words. *The Modern Language Journal, 80*, 327-339.

Jakobovits, L. A., & Lambert, W. E. (1961). Semantic satiation among bilinguals. *Journal of Experimental Psychology, 62*, 576-582.

Jenkins, J. R. (1968). Effects of incidental cues and encoding strategies on paired-associate learning. *Journal of Educational Psychology, 59*, 410-413.

Jiang, P.-S. (2004). *The relationship between EFL learners' vocabulary size and reading comprehension.* Unpublished master's dissertation, National Kaohsiung Normal University, Taiwan.

Johnson, J. S., & Newport, E. L. (1989). Critical period effects in second language learning: The influence of maturational state on the acquisition of English as a second language. *Cognitive Psychology, 21*, 60-99.

Kessler, B., & Treiman, R. (2001). Relationships between sounds and letters in English monosyllables. *Journal of Memory and Language, 44*, 592-617.

Koda, K. (1997). Orthographic knowledge in L2 lexical processing. In J. Coady &

T. Huckin (Eds.), *Second language vocabulary acquisition* (pp. 35-52). Cambridge: Cambridge University Press.

Kolers, P. A. (1963). Interlingual word associations. *Journal of Verbal Learning and Verbal Behaviour*, *2*, 291-300.

Kroll, J. F. (1993). Accessing conceptual representations for words in a second language. In R. Schreuder & B. Weltens (Eds.), *The bilingual lexicon* (pp. 53-81). Amsterdam: John Benjamins.

Kroll, J. F., & Curley, J. (1988). Lexical memory in novice bilinguals: The role of concepts in retrieving second language words. In M. Gruneberg, P. Morris, & R. Sykes (Eds.), *Practical aspects of memory* (Vol. 2, pp. 389-395). London: John Wiley and Sons.

Kroll, J. F., Michael, E., Tokowicz, N., & Dufour, R. (2002). The development of lexical fluency in a second language. *Second Language Research*, *18*, 137-171.

Kroll, J. F., & Potter, M. C. (1984). Recognizing words, pictures, and concepts: A comparison of lexical, object, and reality decisions. *Journal of Verbal Learning and Verbal Behaviour*, *23*, 39-66.

Kroll, J. F., & Sholl, A. (1992). Lexical and conceptual memory in fluent and nonfluent bilinguals. In R. J. Harris (Ed.), *Cognitive processing in bilinguals* (pp. 191-204). Amsterdam: Elsevier Science.

Kroll, J. F., & Stewart, E. (1994). Category interferences in translation and picture naming: Evidence for asymmetric connections between bilingual memory representations. *Journal of Memory and Language, 33*, 149-174.

Kroll, J. F., & Sunderman, G. (2003). Cognitive processes in second language learners and bilinguals: The development of lexical and conceptual representations. In C. J. Doughty & M. H. Long (Eds.), *The handbook of second language acquisition* (pp. 104-129). Oxford: Blackwell.

Laufer, B. (1992). How much lexis is necessary for reading comprehension? In P. J. L. Arnaud & H. Béjoint (Eds.), *Vocabulary and applied linguistics* (pp. 126-132). London: Macmillan Academic and Professional Ltd.

Laufer, B. (1998). The development of passive and active vocabulary in a second language: Same or different? *Applied Linguistics, 19*, 255-271.

Laufer, B., & Paribakht, T. S. (1998). The relationship between passive and active vocabularies: Effects of language learning context. *Language Learning, 48*, 365-391.

Laufer, B., & Shmueli, K. (1997). Memorizing new words: Does teaching have anything to do with it? *RELC Journal, 28*, 89-108.

Lengyel, Z. (1995). Some critical remarks on the phonological component. In D. Singleton & Z. Lengyel (Eds.), *The age factor in second language acquisition. A* *critical look at the critical period hypothesis* (pp. 124-134). Clevedon: Multilingual Matters.

Lenneberg, E. H. (1967). *Biological foundations of language.* New York: John Wiley & Sons.

Liang, Y.-P. (2005). *The relationship between vocabulary skills and vocabulary test: An assessment perspective.* Unpublished master's dissertation, National Yunlin University of Science and Technology, Taiwan.

Liao, C.-P. (2005). *The influence of translation, picture clues and written contexts on Taiwanese sixth graders' word recognition in English learning.* Unpublished master's dissertation, National Taipei University of Education, Taiwan.

Lin, P.-Y. (2005). *The English speech perception abilities of elementary school students with learning disabilities.* Unpublished master's dissertation, National Changhua University of Education, Taiwan.

Liu, W.-L. (2006). *A case study of five factors of early English learning for English speaking performance of senior high school students.* Unpublished master's dissertation, National Kaohsiung Normal University, Taiwan.

Lonigan, C. J., Burgess, S. R., Anthony, J. L., & Barker, T. A. (1998). Development of phonological sensitivity in 2- to 5-year-old children. *Journal of Educational Psychology, 90,* 294-311.

Lotto, L., & de Groot, A. M. B. (1998). Effects of learning method and word type on acquiring vocabulary in an unfamiliar language. *Language Learning, 48*, 31-69.

Lundberg, I. (1991). Phonemic awareness can be developed without reading instruction. In S. A. Brady & D. P. Shankweiler (Eds.), *Phonological processes in literacy: A tribute to Isabelle Y. Liberman* (pp. 47-53). Hillsdale, NJ: Lawrence Erlbaum.

Macnamara, J. (1966). *Bilingualism and primary education.* Edinburgh: Edinburgh University Press.

MacWhinney, B. (2005). New directions in the competition model. In M. Tomasello & D. I. Slobin (Eds.), *Beyond nature-nurture: Essays in honor of Elizabeth Bates* (pp. 81-110). Mahwah, NJ: Lawrence Erlbaum.

Mann, V. A. (1986). Phonological awareness: The role of reading experience. *Cognition, 24*, 65-92.

Marsh, L. G., & Maki, R. H. (1976). Efficiency of arithmetic operations in bilinguals as a function of language. *Memory and Cognition, 4*, 459-464.

Masur, E. F. (1995). Infants' early verbal imitation and their later lexical development. *Merrill-Palmer Quarterly, 41*, 286-306.

McBride-Chang, C. (1995). What is phonological awareness? *Journal of Educational Psychology, 87*, 179-192.

McDonald, J. L. (2000). Grammaticality judgments in a second language: Influences of age of acquisition and native language. *Applied Psycholinguistics, 21*, 395-423.

McDougall, S., Hulme, C., Ellis, A., & Monk, A. (1994). Learning to read: The role of short-term memory and phonological skills. *Journal of Experimental Child Psychology, 58*, 112-133.

McLaughlin, B., Rossman, T., & McLeod, B. (1983). Second language learning: An information-processing perspective. *Language Learning, 33*, 135-158.

McLeod, B., & McLaughlin, B. (1986). Restructuring or automaticity? Reading in a second language. *Language Learning, 36*, 109-123.

Meara, P. (1984). The study of lexis in interlanguage. In A. Davies, C. Criper, & A. P. R. Howatt (Eds.), *Interlanguage* (pp. 225-235). Edinburgh: Edinburgh University Press.

Meara, P. (1992). Network structures and vocabulary acquisition in a foreign language. In P. J. L. Arnaud & H. Béjoint (Eds.), *Vocabulary and applied linguistics* (pp. 62-70). London: Macmillan Academic and Professional Ltd.

Meara, P. (1993). The bilingual lexicon and the teaching of vocabulary. In R. Schreuder & B. Weltens (Eds.), *The bilingual lexicon* (pp. 279-297). Amsterdam: John Benjamins.

Meara, P. (1997). Towards a new approach to modelling vocabulary acquisition. In N. Schmitt & M. McCarthy (Eds.), *Vocabulary: Description, acquisition and pedagogy* (pp. 109-121). Cambridge: Cambridge University Press.

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.

Meschyan, G., & Hernandez, A. (2002). Is native-language decoding skill related to second-language learning? *Journal of Educational Psychology, 94*, 14-22.

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

Metsala, J. L., & Ehri, L. C. (1998). *Word recognition in beginning literacy.* Mahwah, NJ: Lawrence Erlbaum.

Metsala, J. L., & 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). London: Lawrence Erlbaum.

Miller, G. A. (1978). Semantic relations among words. In M. Halle, J. Bresnan, & G. A. Miller (Eds.), *Linguistic theory and psychological reality* (pp. 60-118).

Cambridge, MA: The MIT Press.

Miyake, A., & Friedman, N. P. (1998). Individual differences in second language proficiency: Working memory as language aptitude. In A. F. Healy & L. E. Bourne, Jr. (Eds.), *Foreign language learning: Psycholinguistic studies on training and retention* (pp. 339-364). Mahwah, NJ: Lawrence Erlbaum.

Morais, J., Bertelson, P., Cary, L., & Alegria, J. (1986). Literacy training and speech segmentation. *Cognition, 24*, 45-64.

Moyer, A. (1999). Ultimate attainment in L2 phonology. The critical factors of age, motivation, and instruction. *Studies in Second Language Acquisition, 21*, 81-108.

Muijs, D. (2004). *Doing Quantitative Research in Education with SPSS*. London: Sage.

Muter, V., & Snowling, M. (1998). Concurrent and longitudinal predictors of reading: The role of metalinguistic and short-term memory skills. *Reading Research Quarterly, 33*, 320-337.

Muter, V., Hulme, C., Snowling, M., & Taylor, S. (1998). Segmentation, not rhyming, predicts early progress in learning to read. *Journal of Experimental Child Psychology, 71*, 3-27.

Nagy, W. E., & Herman, P. A. (1987). Breadth and depth of vocabulary knowledge: Implications for acquisition and instruction. In M. G McKeown & M. E. Curtis

(Eds.), *The nature of vocabulary acquisition* (pp. 19-35). Hillsdale, NJ: Lawrence Erlbaum.

Nation, I. S. P. (1982). Beginning to learn foreign vocabulary: A review of the research. *RELC Journal, 13*, 14-36.

Nation, I. S. P. (1990). *Teaching and learning vocabulary.* Boston, MA: Heinle & Heinle.

Nation, I. S. P. (2001). *Learning vocabulary in another language.* Cambridge: Cambridge University Press.

Nieh, P.-L. (2004). *Elementary graduates' EFL proficiency in Taiwan: A factorial analysis.* Unpublished PhD thesis, National Kaohsiung Normal University, Taiwan.

Oller, J. W., Jr., & Nagato, N. (1974). The long-term effect of FLES: An experiment. *The Modern Language Journal, 58*, 15-19.

Oyama, S. (1982a). A sensitive period for the acquisition of a nonnative phonological system. In S. D. Krashen, R. C. Scarcella, & M. H. Long (Eds.), *Child-adult differences in second language acquisition* (pp. 20-38). Rowley, MA: Newbury House. (Reprinted from *Journal of Psycholinguistic Research*, 5, 261-285, 1976.)

Oyama, S. (1982b). The sensitive period and comprehension of speech. In S. D.

Krashen, R. C. Scarcella, & M. H. Long (Eds.), *Child-adult differences in second language acquisition* (pp. 39-51). Rowley, MA: Newbury House. (Reprinted from *Working Papers on Bilingualism*, 16, 1-17, 1978).

Papagno, C., & Vallar, G. (1995). Verbal short-term memory and vocabulary learning in polyglots. *Quarterly Journal of Experimental Psychology, 48A*, 98-107.

Paribakht, T. S., & Wesche, M. (1997). Vocabulary enhancement activities and reading for meaning in second language vocabulary acquisition. In J. Coady & T. Huckin (Eds.), *Second language vocabulary acquisition: A rationale for pedagogy* (pp. 174-200). Cambridge: Cambridge University Press.

Penno, J., Wilkinson, I. A. G., & Moore, D. (2002). Vocabulary acquisition from teacher explanation and repeated listening to stories: do they overcome the Matthew Effect? *Journal of Educational Psychology, 94*, 23-33.

Perfetti, C. A., & Lesgold, A. M. (1977). Discourse comprehension and sources of individual differences. In M. A. Just & P. A. Carpenter (Eds.), *Cognitive processes in comprehension* (pp. 141-183). Hillsdale, NJ: Lawrence Erlbaum.

Polka, L., & Werker, J. F. (1994). Developmental changes in perception of nonnative vowel contrasts. *Journal of Experimental Psychology: Human Perception and Performance, 20*, 421-435.

Potter, M. C., & Faulconer, B. A. (1975). Time to understand pictures and words. *Nature, 253*, 437-438.

Potter, M. C., So, K.-F., von Eckardt, B., & Feldman, L. B. (1984). Lexical and conceptual representation in beginning and proficient bilinguals. *Journal of Verbal Learning and Verbal Behaviour, 23*, 23-38.

Read, C., Zhang, Y., Nie, H., & Ding, B. (1986). The ability to manipulate speech sounds depends on knowing alphabetic writing. *Cognition, 24*, 31-44.

Rickard Liow, S. J., Green, D., & Tam, M. M. L.-J. (1999). The development of visual search strategies in biscriptal readers. *International Journal of Bilingualism, 3*, 333-349.

Rickard Liow, S. J., & Poon, K. K. L. (1998). Phonological awareness in multilingual Chinese children. *Applied Psycholinguistics, 19*, 339-362.

Rixon, S. (1999). Where do the words in EYL textbooks come from? In S. Rixon (Ed.), *Young learners of English: Some research perspectives* (pp. 55-71). Essex: Longman.

Robbins, C., & Ehri, L. C. (1994). Reading storybooks to kindergartners helps them learn new vocabulary words. *Journal of Educational Psychology, 86*, 54-64.

Roodenrys, S., Hulme, C., & Brown, G. (1993). The development of short-term memory span: Separable effects of speech rate and long-term memory. *Journal of Experimental Child Psychology, 56*, 431-442.

Schatschneider, C., Francis, D. J., Foorman, B. R., & Fletcher, J. M. (1999). The dimensionality of phonological awareness: An application of item response theory. *Journal of Educational Psychology, 91*, 439-449.

Schreuder, R., & Weltens, B. (1993). The bilingual lexicon: An overview. In R. Schreuder & B. Weltens (Eds.), *The bilingual lexicon* (pp. 1-10). Amsterdam: John Benjamins.

Segalowitz, N., & Hulstijn, J. (2005). Automaticity in bilingualism and second language learning. In J. F. Kroll & A. M. B. de Groot (Eds.), *Handbook of bilingualism: Psycholinguistic approaches* (pp. 371-388). Oxford: Oxford University Press.

Segalowitz, N., Watson, V., & Segalowitz, S. (1995). Vocabulary skill: single-case assessment of automaticity of word recognition in a timed lexical decision task. *Second Language Research, 11*, 121-136.

Seliger, H. W., Krashen, S. D., & Ladefoged, P. (1982). Maturational constraints in the acquisition of second languages. In S. D. Krashen, R. C. Scarcella, & M. H. Long (Eds.), *Child-adult differences in second language acquisition* (pp. 13-19). Rowley, MA: Newbury House. (Reprinted from *Language Sciences, 38*, 20-22, 1975).

Sénéchal, M., Thomas, E., & Monker, J. (1995). Individual differences in 4-year-old children's acquisition of vocabulary during storybook reading. *Journal of Educational Psychology, 87*, 218-229.

Service, E. (1992). Phonology, working memory, and foreign-language learning. *The Quarterly Journal of Experimental Psychology, 45A*, 21-50.

Service, E., & Kohonen, V. (1995). Is the relation between phonological memory and foreign language learning accounted for by vocabulary acquisition? *Applied Psycholinguistics, 16*, 155-172.

Shatil, E., & Share, D. L. (2003). Cognitive antecedents of early reading ability: A test of the modularity hypothesis. *Journal of Experimental Child Psychology, 86*, 1-31.

Shih, H.-H. (2004). *The effects and the strategies of vocabulary learning on EFL junior and senior high school students in Taiwan.* Unpublished master's dissertation, National Kaohsiung Normal University, Taiwan.

Siegel, L. S. (1998). Phonological processing deficits and reading disabilities. In J. L. Metsala & L. C. Ehri (Eds.), *Word recognition in beginning literacy* (pp. 141-160). London: Lawrence Erlbaum.

Smith, M. (1978). The acquisition of word meaning: An introduction. *Child Development, 49*, 950-952.

Smith, M. C., & Magee, L. E. (1980). Tracing the time course of picture-word processing. *Journal of Experimental Psychology: General, 109*, 373-392.

Speciale, G, Ellis, N. C., & Bywater, T. (2004). Phonological sequence learning and short-term store capacity determine second language vocabulary acquisition. *Applied Psycholinguistics, 25*, 293-321.

Stahl, S. A., & Murray, B. A. (1994). Defining phonological awareness and its relationship to early reading. *Journal of Educational Psychology, 86*, 221-234.

Stahl, S. A., & Murray, B. A. (1998). Issues involved in defining phonological awareness and its relationship to early reading. In J. L. Metsala & L. C. Ehri (Eds.), *Word recognition in beginning literacy* (pp. 65-87). London: Lawrence Erlbaum.

Stanovich, K. E. (1992). Speculations on the causes and consequences of individual differences in early reading acquisition. In P. B. Gough, L. C. Ehri, & R. Treiman (Eds.), *Reading acquisition* (pp. 307-342). Hove: Lawrence Erlbaum.

Sun, L.-Y. (2002). *An investigation of the correlation between phonological awareness and word recognition ability in EFL elementary students.* Unpublished master's dissertation, National Taichung University, Taiwan.

Swan, M. (1997). The influence of the mother tongue on second language vocabulary acquisition and use. In N. Schmitt & M. McCarthy (Eds.), *Vocabulary: Description, acquisition and pedagogy* (pp. 156-180). Cambridge: Cambridge

University Press.

Taiwan, Ministry of Education. (2006). The Area of Language Arts—English. In *The General Guidelines of Grade 1-9 Curriculum of Elementary and Junior High School Education.* (Rev. ed.) Taipei, Taiwan: MOE. Retrieved August 20, 2007, from http://www.edu.tw/EDU_WEB/EDU_MGT/EJE/EDU5147002/9CC /English.doc.

Torgesen, J. K., & Burgess, S. R. (1998). Consistency of reading-related phonological processes throughout early childhood: Evidence from longitudinal-correlational and instructional studies. In J. L. Metsala & L. C. Ehri (Eds.), *Word recognition in beginning literacy* (pp. 161-188). Mahwah, NJ: Lawrence Erlbaum.

Torgesen, J. K., & Davis, C. (1996). Individual differences variables that predict response to training in phonological awareness. *Journal of Experimental Child Psychology, 63*, 1-21.

Torgesen, J. K., Morgan, S. T., & Davis, C. (1992). Effects of two types of phonological awareness training on word learning in kindergarten children. *Journal of Educational Psychology, 84*, 364-370.

Torgesen, J. K., Wagner, R. K., Rashotte, C. A., Rose, E., Lindamood, P., Conway, T., et al. (1999). Preventing reading failure in young children with phonological processing disabilities: Group and individual responses to instruction. *Journal of*

Educational Psychology, 91, 579-593.

Treiman, R., & Zukowski, A. (1991). Levels of phonological awareness. In S. A. Brady & D. P. Shankweiler (Eds.), *Phonological processes in literacy: A tribute to Isabelle Y. Liberman* (pp. 67-83). Hillsdale, NJ: Lawrence Erlbaum.

Tsai, B.-Y. (2005). *The relationship between receptive English vocabulary sizes and listening comprehension competence of college EFL students*. Unpublished master's dissertation, National Kaohsiung Normal University, Taiwan.

Tsai, C.-H. (2001). *Parents' attitudes towards English learning for their kindergarten children*. Unpublished master's dissertation, Nanhua University, Taiwan.

Vallar, G., & Baddeley, A. D. (1984). Fractionation of working memory: Neuropsychological evidence for a phonological short-term store. *Journal of Verbal Learning and Verbal Behaviour, 23*, 151-161.

van Hell, J. G., & Candia Mahn, A. (1997). Keyword mnemonics versus rote rehearsal: Learning concrete and abstract foreign words by experienced and inexperienced learners. *Language Learning, 47*, 507-546.

Vygotsky, L. (1986). *Thought and language*. Cambridge, MA: The MIT Press.

Wagner, R. K., & Torgesen, J. K. (1987). The nature of phonological processing and its causal role in the acquisition of reading skills. *Psychological Bulletin, 101*, 192-212.

Weinreich, U. (1953). *Languages in contact: Findings and problems.* New York: Linguistic Circle of New York.

Wellhousen, K. (1993). Eliciting and examining young children's storytelling. *Journal of Research in Childhood Education, 7*, 62-65.

Werker, J. F., & Tees, R. C. (2002). Cross-language speech perception: Evidence for perceptual reorganization during the first year of life. *Infant Behaviour and Development, 25*, 121-133.

White, L., & Genesee, F. (1996). How native is near-native? The issue of ultimate attainment in adult second language acquisition. *Second Language Research, 12*, 233-265.

Williams, E. (1996). Reading in two languages at year five in African primary schools. *Applied Linguistics, 17*, 182-209.

Wimer, C. C., & Lambert, W. E. (1959). The differential effects of word and object stimuli on the learning of paired associates. *Journal of Experimental Psychology. 57*, 31-36.

Winskel, H., & Widjaja, V. (2007). Phonological awareness, letter knowledge, and literacy development in Indonesian beginner readers and spellers. *Applied Psycholinguistics, 28*, 23-45.

Wu, M.-H. (2005). Relationship between sixth-graders' use of memory strategies and background variables in English vocabulary learning. Unpublished master's dissertation, National Taipei University of Education, Taiwan.

Yang, H.-F. (2006). The correlation study of real-pseudo words in primary students' phonological awareness. Unpublished master's dissertation, National Taichung University, Taiwan.

Yin, C.-W. (2006). *The effect of early start and parental socioeconomic status on sixth-graders' English proficiency*. Unpublished master's dissertation, National Kaohsiung Normal University, Taiwan.

Yopp, H. K. (1988). The validity and reliability of phonemic awareness tests. *Reading Research Quarterly*, *23*, 159-177.

Yovanoff, P., Duesbery, L., Alonzo, J., & Tindal, G. (2005). Grade-level invariance of a theoretical causal structure predicting reading comprehension with vocabulary and oral reading fluency. *Educational Measurement: Issues and Practice*, *24*, 4-12.

Zimmerman, C. B. (1997). Historical trends in second language vocabulary instruction. In J. Coady & T. Huckin (Eds.), *Second language vocabulary acquisition: A rationale for pedagogy* (pp. 5-19). Cambridge: Cambridge University Press.

Appendix A
Questionnaire

Class ____________ Name ____________________

Gender __________ Age __________

A. English Experience

1. Have you ever lived in an English-speaking country?

□ Yes, I have lived in ____________(country) for _____ year(s).

□ No, I have not.

2. English Education

□ I have attended an English-medium school for ____________ year(s).

□ I have never studied at an English-medium school.

3. Extra-curricular English (ECE) lessons

Nursery and Reception Year

a) □ I have attended a kindergarten where English was taught for __________ year(s).

b) □ I have studied at an English cram school for__________ year(s).

4. Primary School

a) □ In Grade 1, I attended ECE lessons for __ months and __ hours per week.

b) □ In Grade 2, I attended ECE lessons for __ months and __ hours per week.

c) □ In Grade 3, I attended ECE lessons for __ months and __ hours per week.

B. Language Use (MC=Mandarin Chinese, TWN=Taiwanese)

5. At home

a) I speak to the elderly most often in □MC □TWN □other (_____).

b) I speak to siblings & cousins most often in □MC □TWN □other (_____).

c) I speak to my neighbours most often in □MC □TWN □other (_____).

6. At school

a) I speak to teachers most often in □MC □TWN □other (_____).

b) I speak to schoolmates most often in □MC □TWN □other (_____).

C. Attitudes towards English Skills

7. English has four skills & they are listening(L), speaking(S), reading(R) and writing (W).

a) Among them, the most important skill to me is □L □S □R □W.

b) Among them, the one I want most to master is □L □S □R □W.

c) Among them, the easiest skill to pick up is □L □S □R □W.

d) Among them, the most difficult one to me is □L □S □R □W.

D. Computer Use

8. I have been using a computer for ______________(length of time).

Appendix B

Nonword Repetition Task

Nonwords	Phoneme	Syllable	Consonant cluster (in syllable)				
			1st	2nd	3rd	4th	5th
a. rubid	5	2					
b. pennel	5	2					
c. prindle	7	2	*				
d. thickery	6	3					
e. barrazon	7	3					
f. hampent	7	2		*			
g. sladding	6	2	*				
h. bannifer	6	3					
i. glistering	9	3	*	*			
j. doppelate	7	3					
k. fennerizer	8	4					
l. commerine	7	3					
m. voltularity	11	5					
n. sepretennial	11	5		*			
o. trumpetine	9	3	*				
p. stopograttic	11	4	*		*		
q. commeecitate	9	4					
r. confrantually	12	5		*			
s. defermication	11	5					
t. versatrationist	13	5			*		*

Note. Adapted from S. E. Gathercole's (1995) Children's Test of Nonword Repetition.
The shaded areas represent the number of syllables a nonword possesses.
* marks the syllable in which a consonant cluster is located.

Appendix C.i

Chinese Phonological Sensitivity Assessment: Rhyme Detection

a. Demonstration and Practice Items

No.	Demonstration item		No.	Practice item	
1	biao	piao	1	pie	pian
2	sua*	dua*	2	liang*	xia
3	duo	dio	3	lun*	dun*
4	hio*	tiou*	4	kue*	xue
5	luai*	guai	5	duei	huei
6	min*	qian	6	quan	rueng*
7	ruan*	buan*	7	nong*	zhong
8	kien*	miou*	8	niao*	gie*
9	nuo*	tuo	9	pui*	muan*
10	hie*	qiou	10	duan*	kuan*

b. Test Items

No.	Pair		No.	Pair	
1	jiu	xiang	11	guen*	cui
2	biu*	hiu*	12	lia*	pia*
3	giao*	liao	13	ming*	riao*
4	nia*	jia	14	huai*	duai*
5	shuei*	lua*	15	liong*	qun
6	lue*	xuan	16	puan*	tuan
7	hua	shua	17	qing	liu
8	luei*	fua*	18	giang*	hian*
9	qua*	bua*	19	qiong	jiong
10	ruo*	huo	20	shuang	chui

Note. * indicates a nonword.

Appendix C.ii

Chinese Phonological Sensitivity Assessment: Head Detection

a. Demonstration and Practice Items

No.	Demonstration item		No.	Practice item	
1	bie	biao	1	guai	guan
2	sia*	siang*	2	ruang*	rua*
3	nia*	niao*	3	lue*	luan*
4	ling	luo	4	niai*	nuai*
5	shuei*	huai	5	luei*	liong*
6	lueng*	luan*	6	hiang*	hiao*
7	tia*	tua*	7	ming*	nong*
8	duang*	dun*	8	fia*	fua*
9	jia	jiou	9	ziou*	zian*
10	kua	qiong	10	zhuang	xiei

b. Test Items

No.	Pair		No.	Pair	
1	cong	zhong	11	pian	pien
2	jio*	jiai*	12	xuan	xuen
3	kuo*	kuai	13	qiai*	qio*
4	duei	duen	14	pong*	piong*
5	rui*	mian*	15	jiang	kuang
6	guen*	ruen*	16	lia*	lun*
7	liang*	lien*	17	nin*	ning*
8	niang*	nuen*	18	kiao*	kiang*
9	giong*	giao*	19	mui*	miang*
10	miao	tong	20	ruan*	rueng*

Note. * indicates a nonword.

Appendix C.iii

Chinese Phonological Sensitivity Assessment:

Rhyme and Head Detection/Production

a. Demonstration and Practice Items

No.	Demonstration item		No.	Practice item	
1	biao	piao	1	sua*	dua*
2	mie	miao	2	lun*	dun*
3	luan*	liou	3	liei	liao
4	luai*	guai	4	queng	jun
5	kue*	xue	5	duan*	kuan*
6	kian*	kiang*	6	hiou*	hian*
7	tuang*	tua*	7	ruan*	buan*
8	riong*	run*	8	nong*	zhong
9	nuo*	tuo	9	hie*	ruai*
10	gua	liou	10	nua*	nuei*

b. Test Items

No.	Pair		No.	Pair	
1	biai*	bia*	11	pui*	puang*
2	tuang*	qiu	12	lia*	pia*
3	giao*	liao	13	chuen	shua
4	nue*	xue	14	huai*	duai*
5	kun*	qun	15	guen*	gueng
6	tiao	ting	16	ding	hia*
7	shuai	shuan	17	luan*	kuan*
8	ruei*	kuo*	18	qiang	piang*
9	kuo*	kuei	19	lue*	luan*
10	ruo*	huo	20	xin	xian

Note. * indicates a nonword.

Appendix C.iv

Chinese Phonological Sensitivity Assessment: Initial Sound Isolation

a. Demonstration and Practice Items

No.	Demonstration item	No.	Practice item
1	xia	1	nue*
2	jiong	2	chuang
3	sua*	3	mia*
4	nuen*	4	luang*
5	tiou*	5	bui*
6	liong*	6	qiai*
7	piai*	7	duan*
8	miou*	8	zhuo
9	bun*	9	kuo*
10	shui*	10	huai*

b. Test Items

No.	Item	No.	Item
1	kuai	11	duai*
2	hia*	12	kian*
3	ruan*	13	ging*
4	duei	14	qua*
5	fia*	15	ming
6	bua*	16	piang*
7	lin*	17	zuai*
8	que	18	guen*
9	tuo	19	nia*
10	nian	20	jiai*

Note. * indicates a nonword.

Appendix D.i

English Phonological Sensitivity Assessment: Rhyme Detection

a. Demonstration and Practice Items

No.	Demonstration item		No.	Practice item	
1	cab	lab	1	chief	choose
2	beef	leaf	2	wade	weed
3	moon	noon	3	beg	leg
4	surf	turf	4	life	like
5	date	deep	5	name	game
6	gum	jam			
7	cope	cup			
8	tone	those			
9	verb	verge			
10	fire	five			

b. Test Items

No.	Pair		No.	Pair	
1	choice	voice	14	fish	wish
2	fight	fought	15	cook	hook
3	hop	mop	16	seem	seen
4	sheet	shirt	17	zoom	booze
5	herd	jade	18	loss	mousse
6	size	wise	19	such	touch
7	hall	mall	20	shell	Share
8	fun	food	21	mine	turn
9	curve	church	22	perm	purse
10	vase	thick	23	birth	worth
11	get	gap	24	judge	budge
12	shout	doubt	25	vet	web
13	cash	young			

Appendix D.ii

English Phonological Sensitivity Assessment: Head Detection

a. Demonstration and Practice Items

No.	Demonstration item		No.	Practice item	
1	bag	bat	1	chair	check
2	come	cut	2	rice	nice
3	dim	dish	3	firm	fan
4	chat	church	4	choose	was
5	sheep	ship	5	nurse	nerd
6	wife	tide			
7	gate	hate			
8	guide	fade			
9	wait	wake			
10	light	nose			

b. Test Items

No.	Pair		No.	Pair	
1	gaze	gap	14	base	pig
2	jack	jab	15	let	lot
3	vine	wine	16	night	hurt
4	rug	pub	17	feet	heat
5	boom	doom	18	toil	toys
6	birth	wrath	19	down	doubt
7	soup	soon	20	thin	thick
8	merge	mirth	21	nap	nape
9	tom	top	22	shirk	surge
10	type	town	23	pack	race
11	pile	pal	24	should	shook
12	poke	pope	25	read	reach
13	nut	numb			

Appendix D.iii

English Phonological Sensitivity Assessment:

Rhyme and Head Detection/Production

a. Demonstration and Practice Items

No.	Demonstration item		No.	Practice item	
1	keep	keen	1	rub	cub
2	jerk	job	2	hack	shack
3	fetch	letch	3	pun	pub
4	wide	hide	4	life	lid
5	half	ham	5	rice	right
6	jill	jim			
7	done	ban			
8	hope	soap			
9	lake	late			
10	house	mouse			

b. Test Items

No.	Pair		No.	Pair	
1	buck	bug	14	move	mood
2	hike	like	15	shop	top
3	couch	vouch	16	mare	wear
4	soil	boil	17	zoom	zoos
5	deer	dig	18	pass	mass
6	side	sad	19	book	bush
7	pall	call	20	rash	tall
8	phone	foam	21	fate	faith
9	girl	gird	22	leave	weave
10	wage	rage	23	cheap	cheat
11	shape	tell	24	met	debt
12	mouth	south	25	hitch	pitch
13	curb	cob			

Appendix D.iv

English Phonological Sensitivity Assessment: Initial Consonant Isolation

a. Demonstration and Practice Items

No.	Demonstration item	No.	Practice item
1	goose	1	wap
2	mar	2	couch
3	judge	3	ripe
4	neat	4	fin
5	chap	5	vouch
6	hiss		
7	poach		
8	fake		
9	lice		
10	rouse		

b. Test Items

No.	Item	No.	Item
1	bail	14	rife
2	kin	15	foil
3	shook	16	cat
4	robe	17	tease
5	zoom	18	their
6	geese	19	yacht
7	third	20	surge
8	lose	21	hawk
9	fate	22	chose
10	noun	23	dive
11	jazz	24	pouch
12	vote	25	mud
13	wok		

Appendix E.i

Text of *Silly Willy*

"Get out of bed, Willy. It is time to get dressed."

"I can do it! My pants go on my head." "No, silly Willy! You look like a bunny!"

"My woolen hat goes on my nose." "No, silly Willy! You look like a duck!"

"My sneakers go on my hands." "No, silly Willy! You look like a seal!"

"My gloves go on my feet." "No, silly Willy! You look like a frog!"

"I am a bunny, a duck, a seal, and a frog." "You are very silly, Willy!"

"Your pants do not go on your head. Your pants go on your feet."

"Your woolen hat does not go on your nose. Your woolen hat goes on your head. Your sneakers do not go on your hands. Your sneakers go on your feet."

"Your gloves do not go on your feet. Your gloves go on your hands."

"And here is a kiss for your nose. Now you are not silly, Willy!"

"But, I AM!"

By Maryann Cocca-Leffler, East & West Co., Ltd.

Note. Underlined words are target lexical items taught to the subjects. Some words have been altered and their synonyms are adopted for the convenience of the experiment. Other words that are not mentioned in the text but whose pictures appear in the illustrations have been used.

Appendix E.ii

Text of *King Big Wig*

Big Wig was a king. He liked big things. He liked his big wig and his big crown.

He liked his big bed and his big castle. Every night, the queen gave him a big bowl of ice cream. The king liked that too.

One day, the king was sailing his toy boat. He got a big idea. If big things are good, bigger things are better!

So the king got a bigger wig and a bigger crown. He got a bigger bed and a bigger castle.

But was the king happy?

No. His wig was too big! It fell in his eyes and it tickled his nose.

Achoo!

His crown was too big! It hurt his head. "Ow! I am going to bed," the king said.

So off went the king. But he did not find his bed. His castle was too big! "Bigger is not better," the king said.

So what did the king do? He left his bigger castle. He got rid of his bigger wig and his bigger crown.

The king got back his old wig and his old crown and his old castle. Then the king got into his old bed. But was the king happy?

Yes. He was. Until the queen came in with a bigger bowl of ice cream.

Then the king was very happy! "Sometimes bigger is better," the king said.

Story by Portia Aborio and Illustration by Sonja Lamut.

Note. Underlined words are target lexical items taught to the subjects. Other words that are not mentioned in the text but whose pictures appear in the illustrations have been used.

Appendix E.iii

Text of *Lots of Hearts*

I am making cards. I have paper, tape, and paint.

My dog wags her tail. She wants to help. But I say no.

I make lots of hearts. I make a card for my dad. It has a pink heart and red roses on it. I hide it in his book.

I make a card for my mom. It has a red heart and pink bows on it. I hide it in her boot.

I make a card for my goldfish. It has a blue heart and a yellow shell on it. I tape it to the fishbowl.

And I make a card for my dog. It has a green heart and a big bone on it. I hide it in her doghouse.

Now I say to Dad, "Look in your book. You will find a card." Dad looks. But there is no card.

I say to Mom, "Look in your boot. You will find a card." Mom looks. But there is no card!

I go to the fishbowl. There is no card for the goldfish!

I go to the doghouse. Here are the cards! My dog has all the cards! And my dog also has a big surprise for me! My dog had puppies! What a happy Valentine's Day!

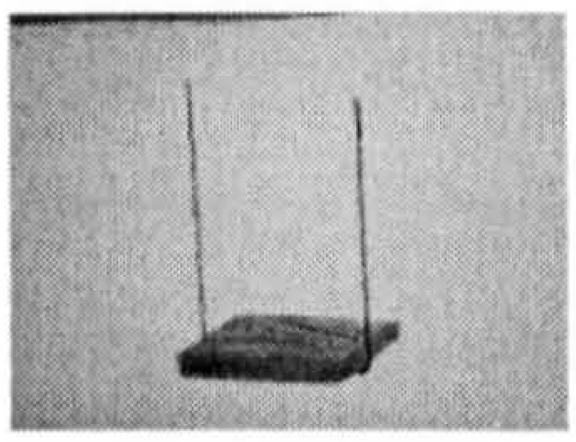
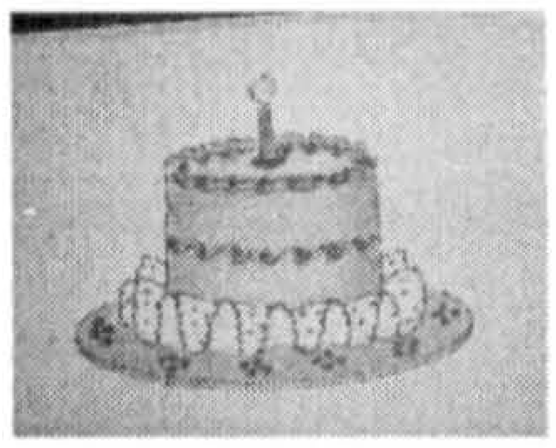
By Maryann Cocca-Leffler, East & West Book Co.,Ltd.

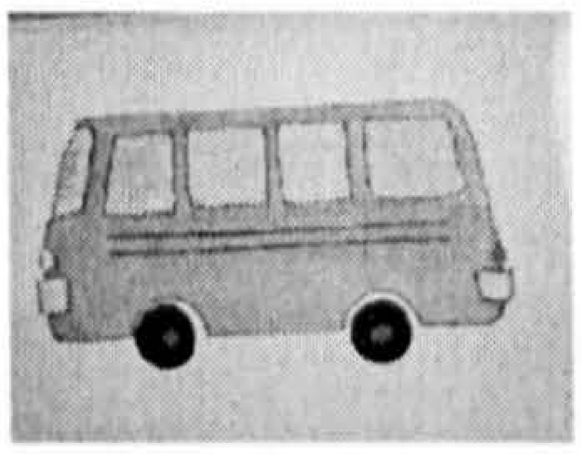
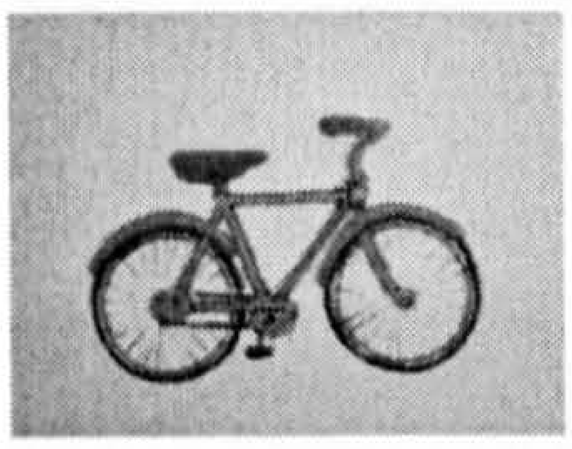
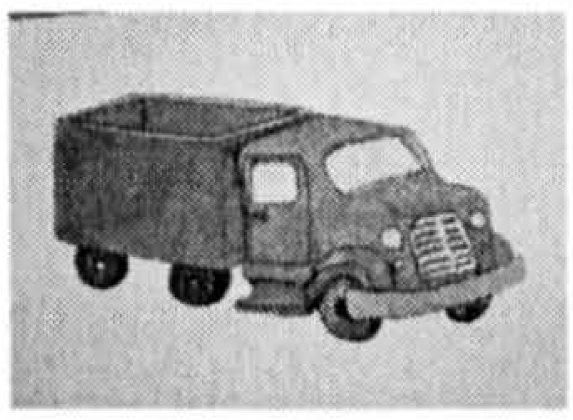
Note. Underlined words are target lexical items taught to the subjects. Other words that are not mentioned in the text but whose pictures appear in the illustrations have been used.

Appendix F.i

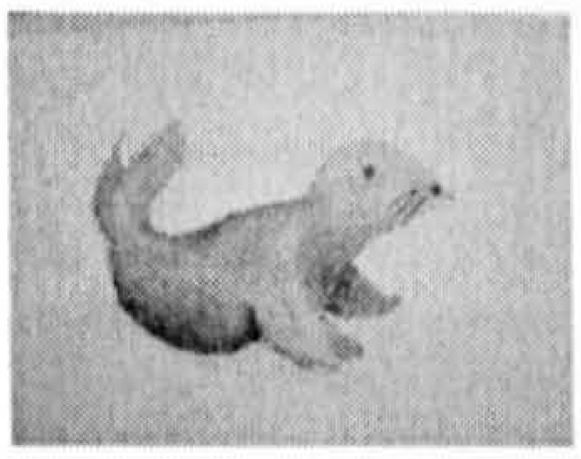
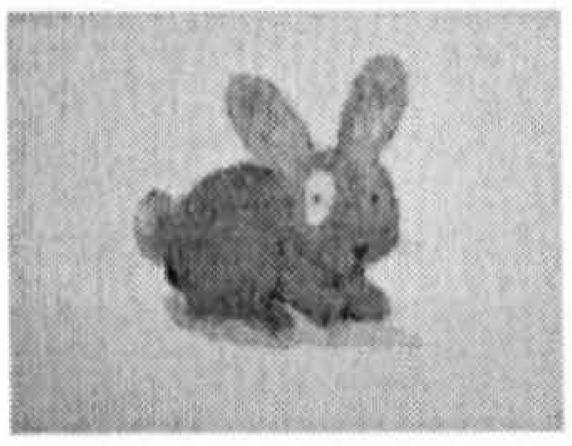
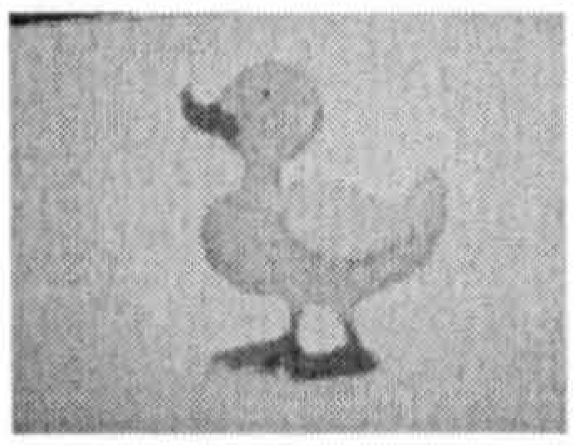
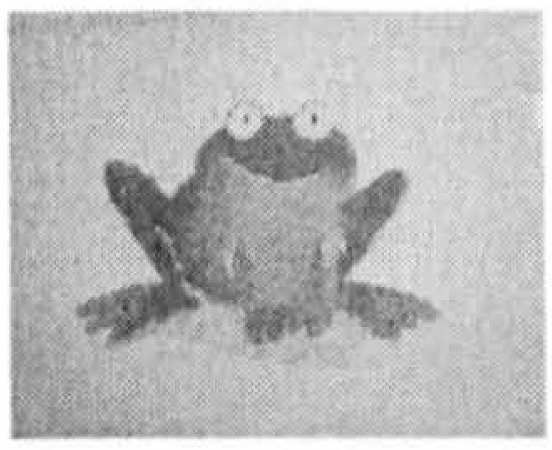
Written Vocabulary Assessment for Story 1

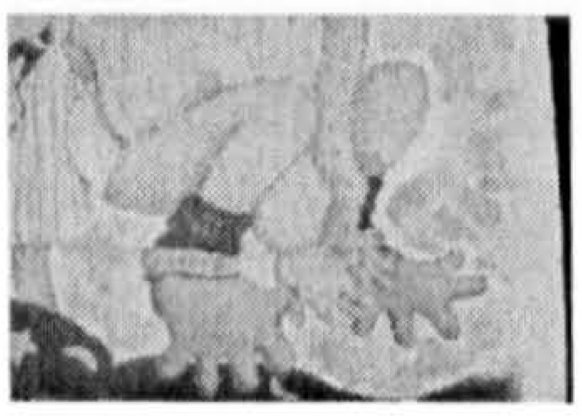
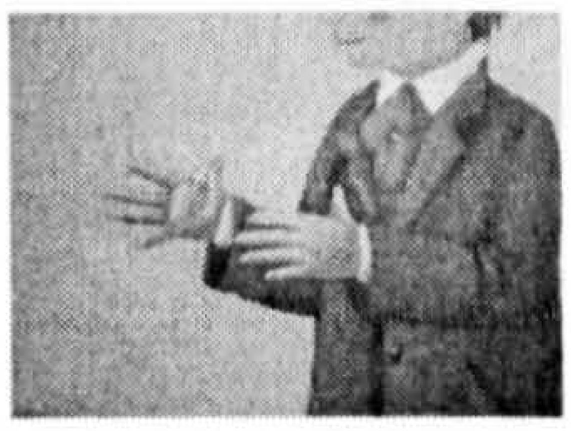
A.

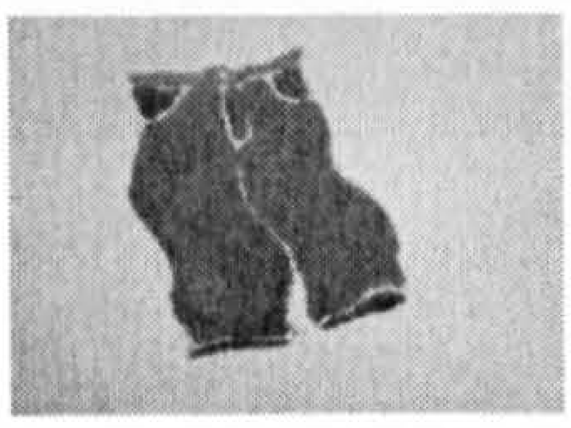
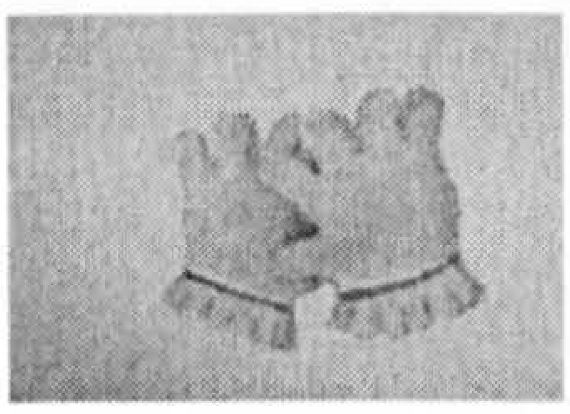
  

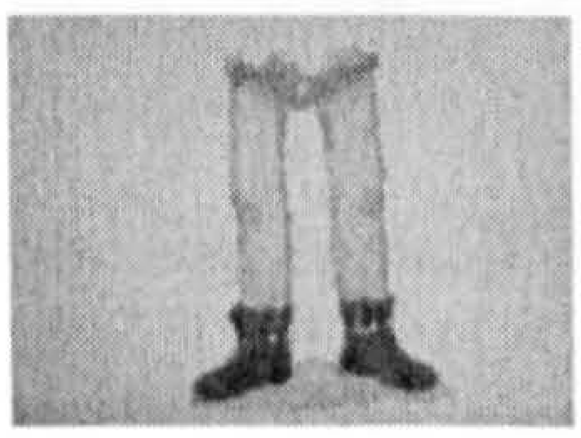
B.

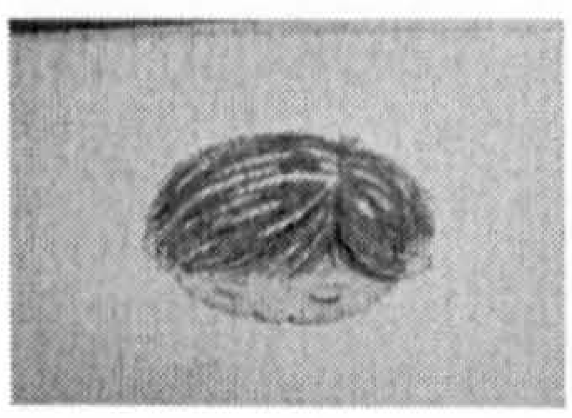
   

C.

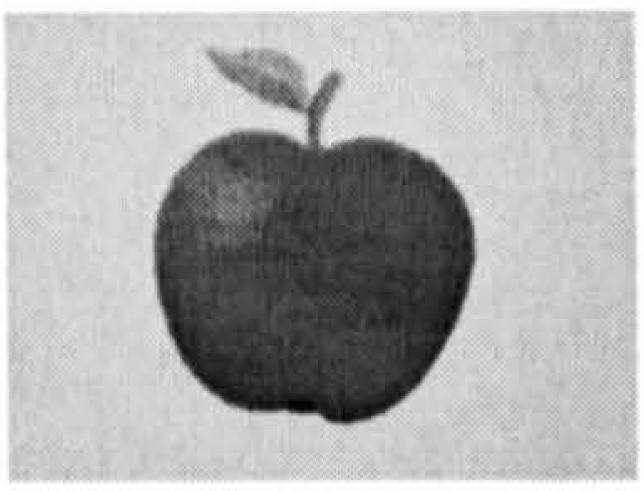
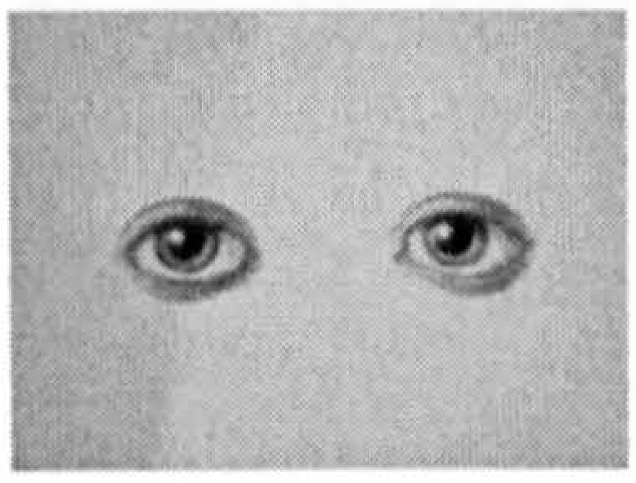
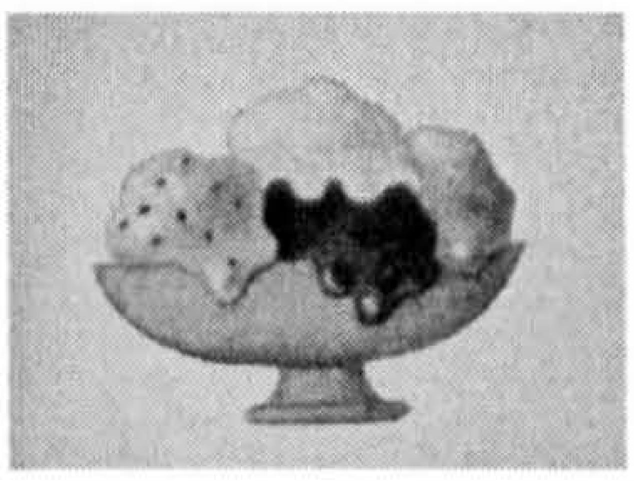
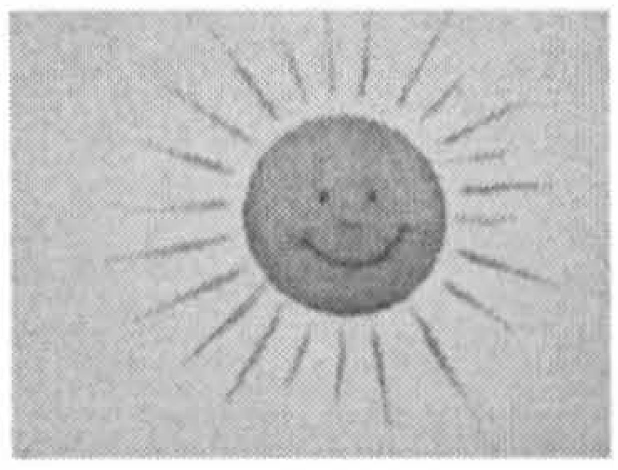
   

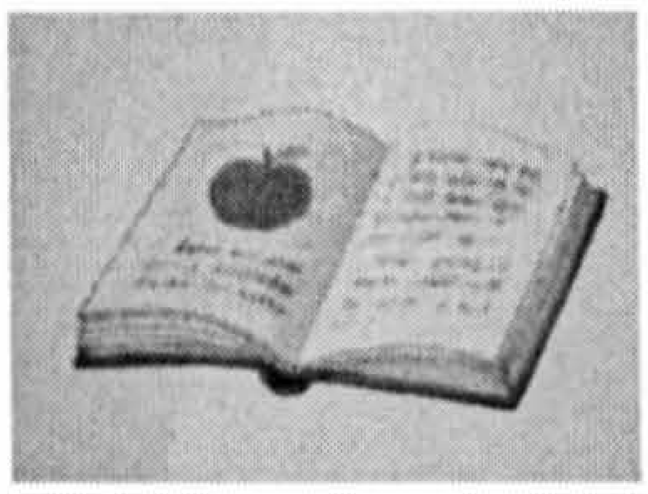
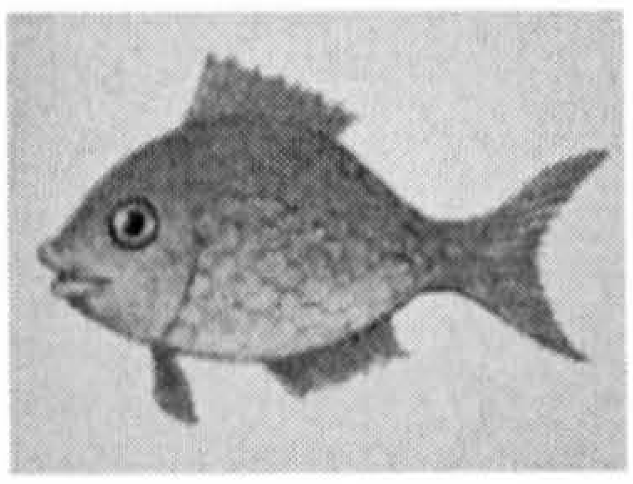
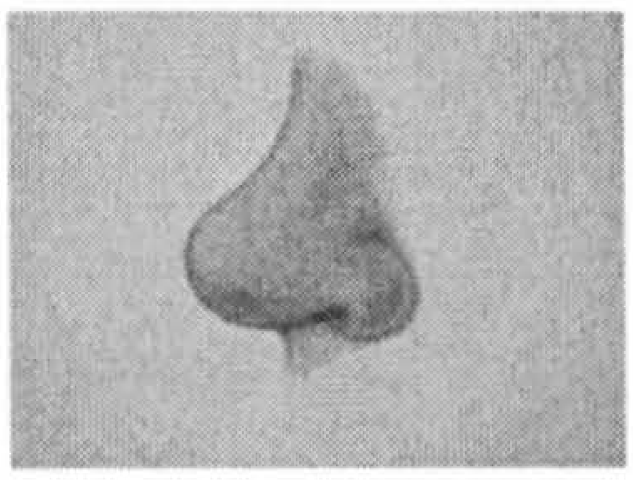
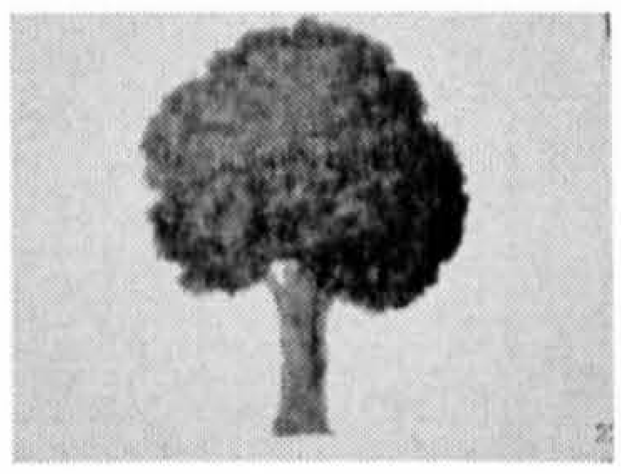
  

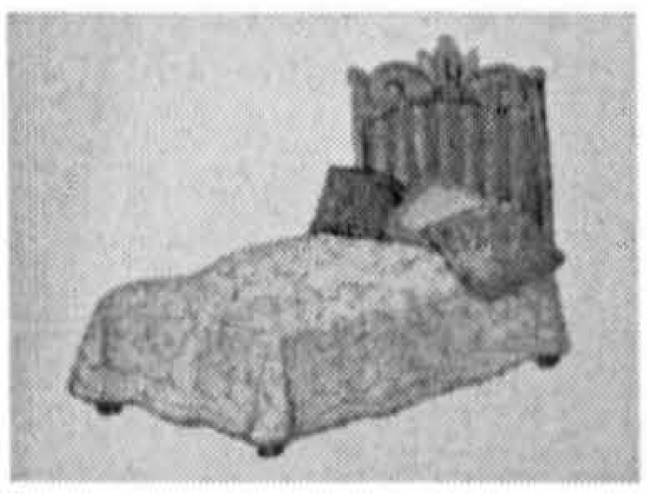
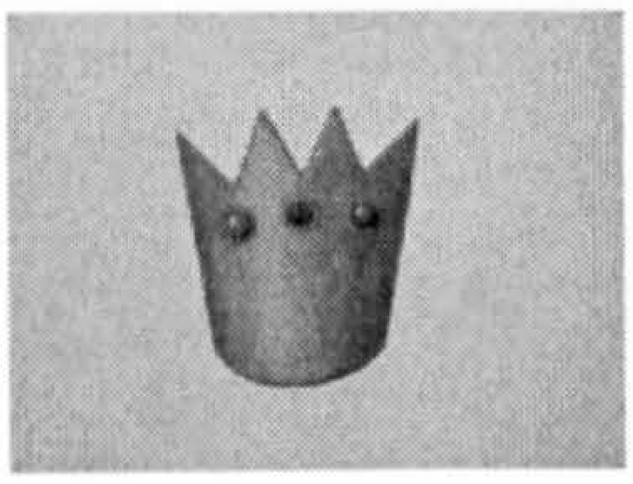
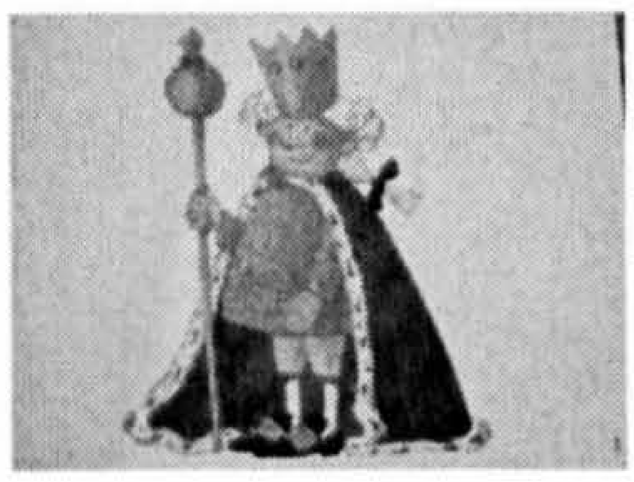
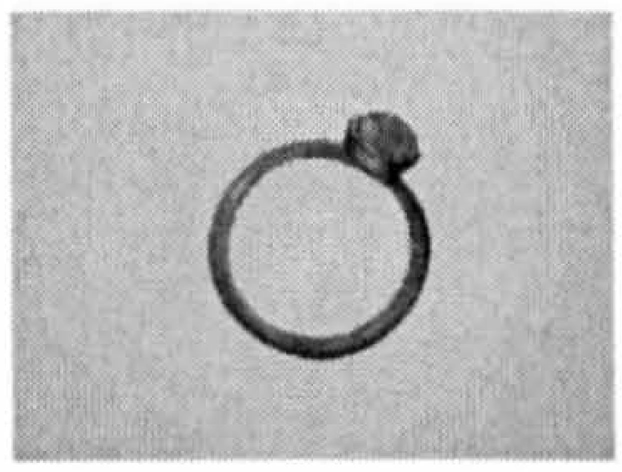
Appendix F.ii
Written Vocabulary Assessment for Story 2

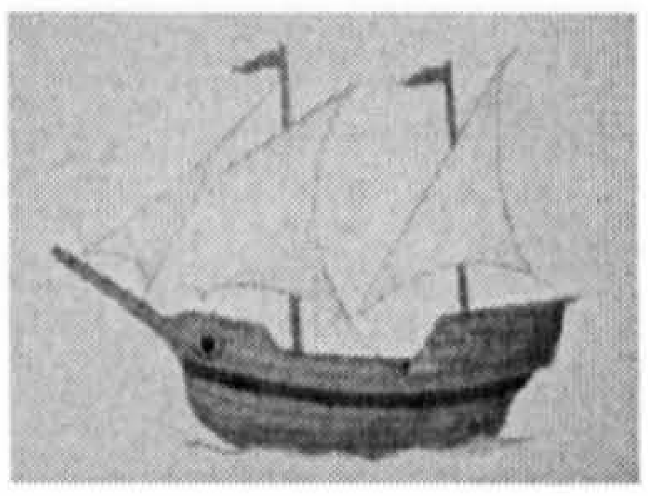
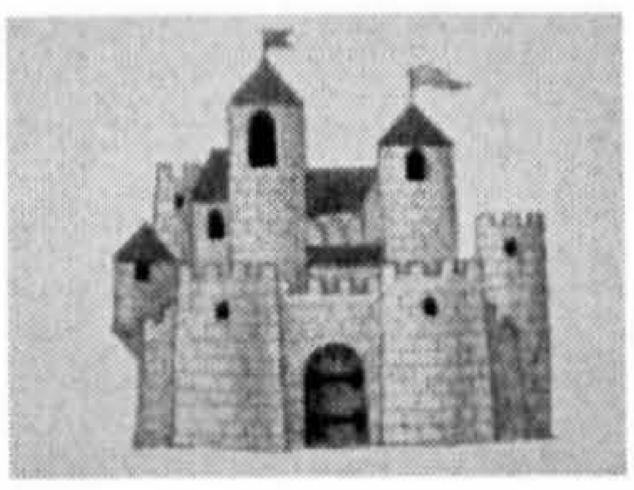
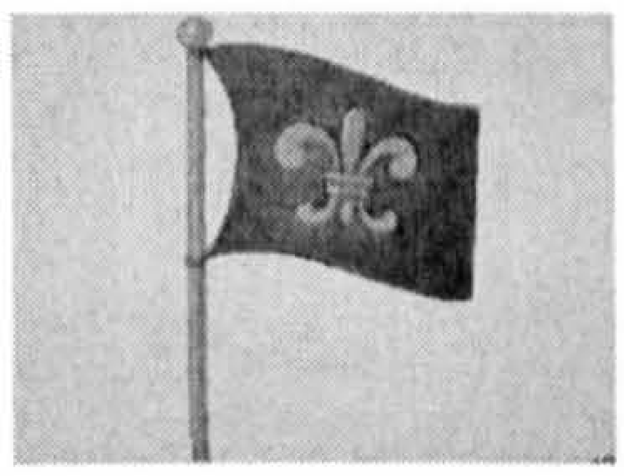
A.

 ______  ______  ______  ______

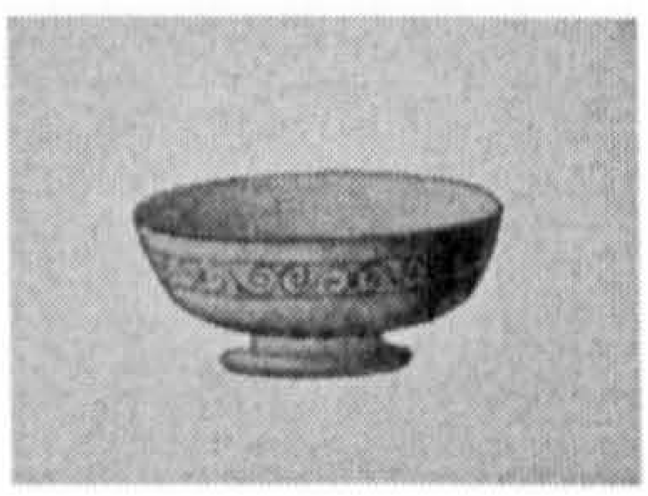
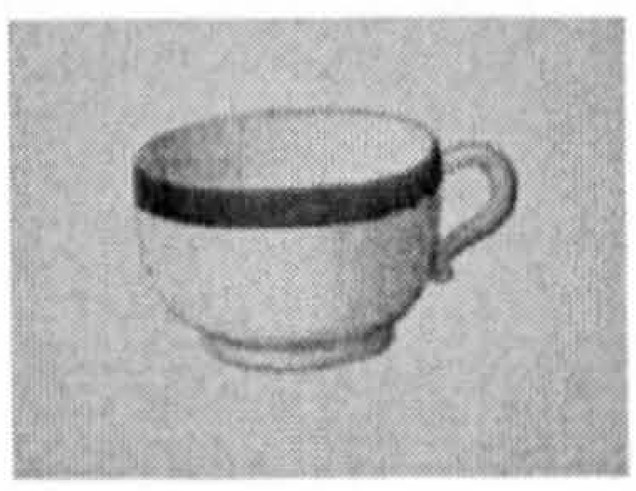
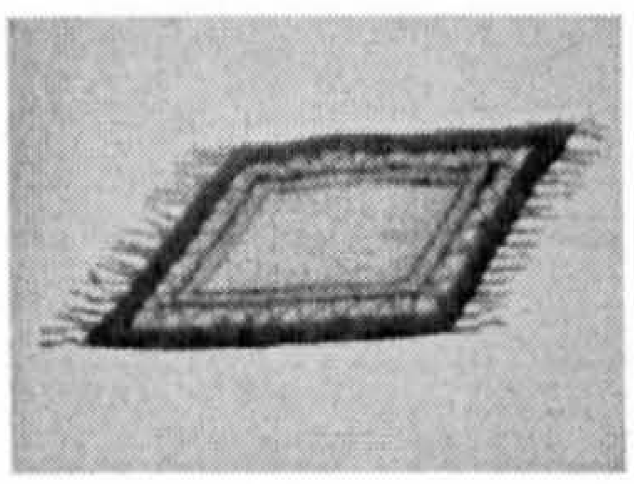
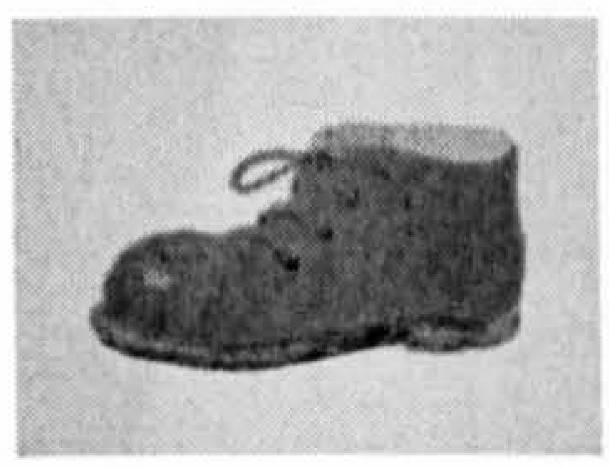
 ______  ______  ______  ______

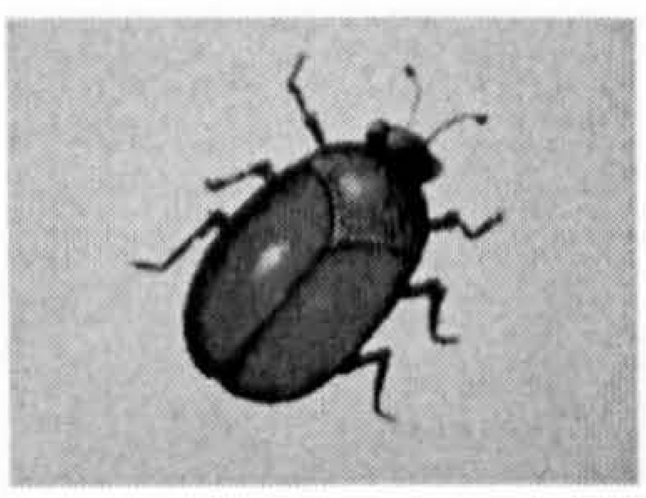
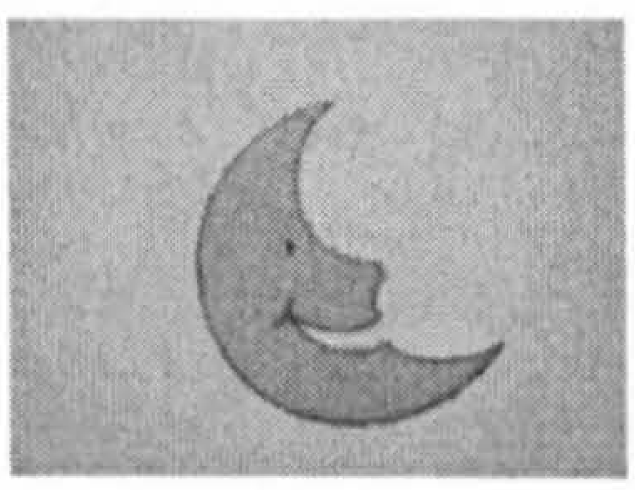
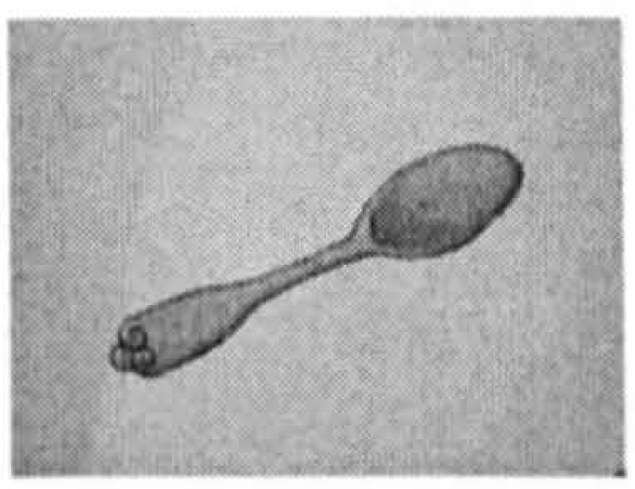
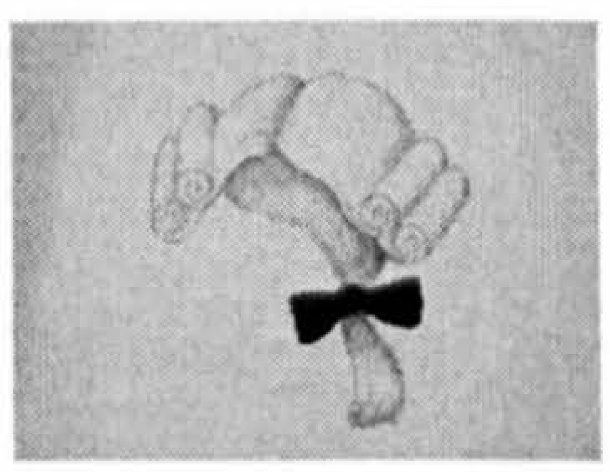
B.

 ______  ______  ______  ______

 ______  ______  ______  ______

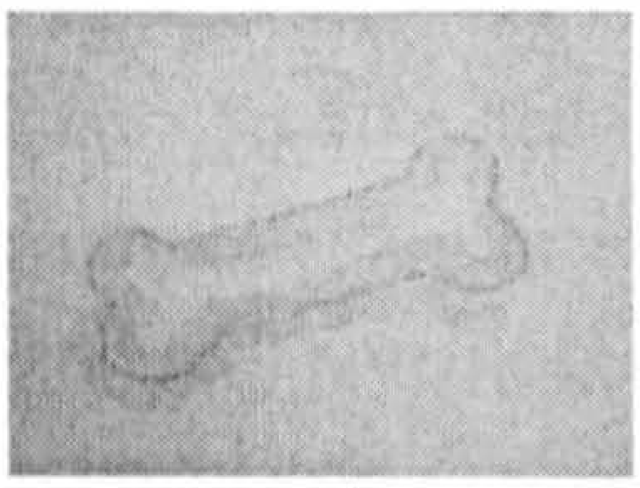
C.

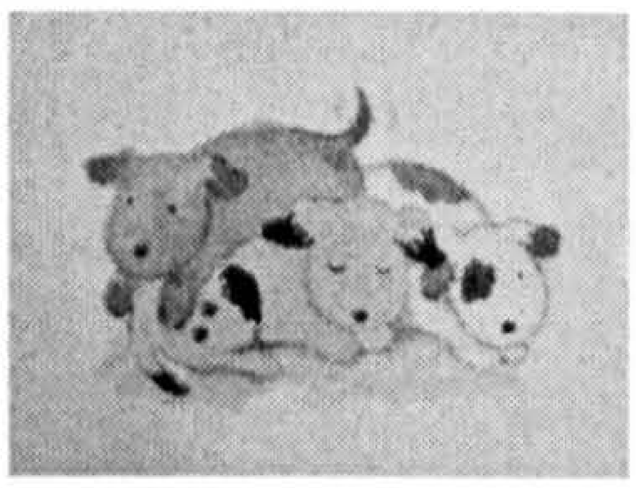
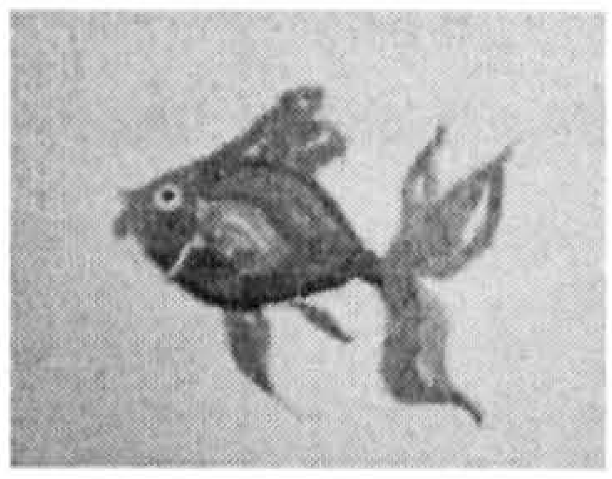
 ______  ______  ______  ______

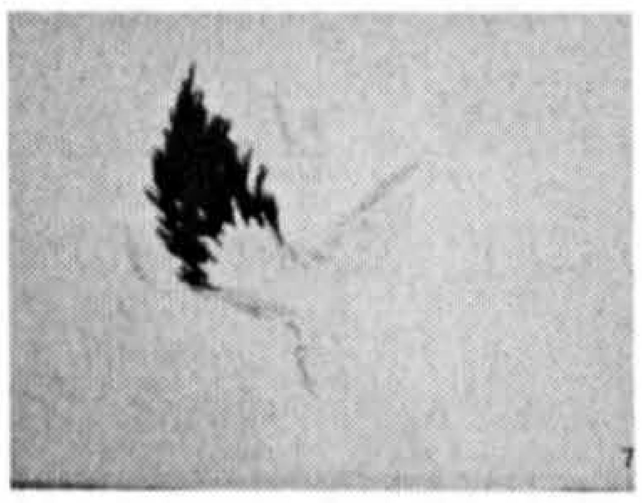
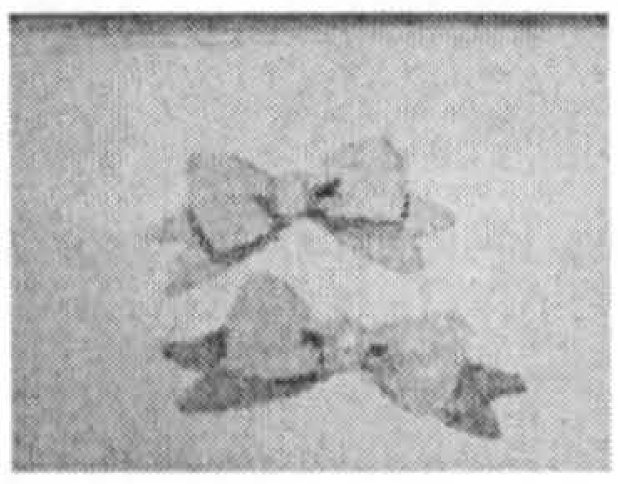
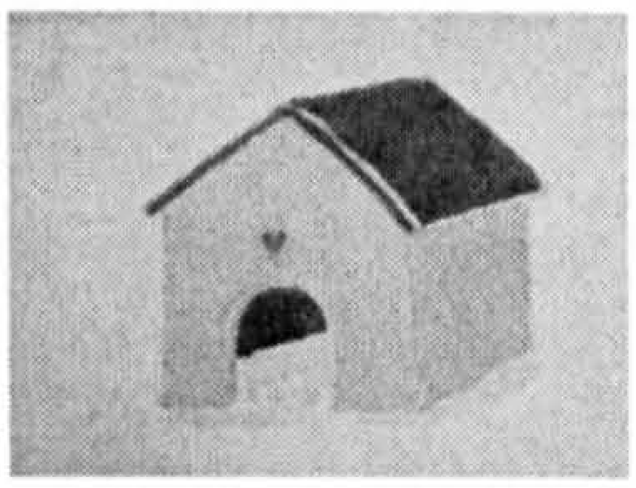
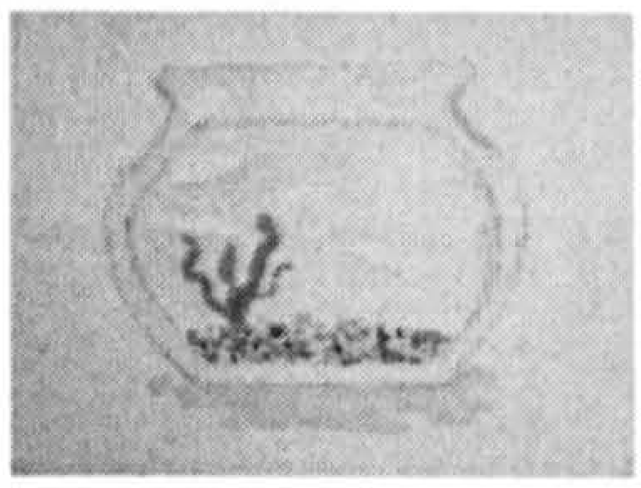
 ______  ______  ______  ______

Appendix F.iii
Written Vocabulary Assessment for Story 3

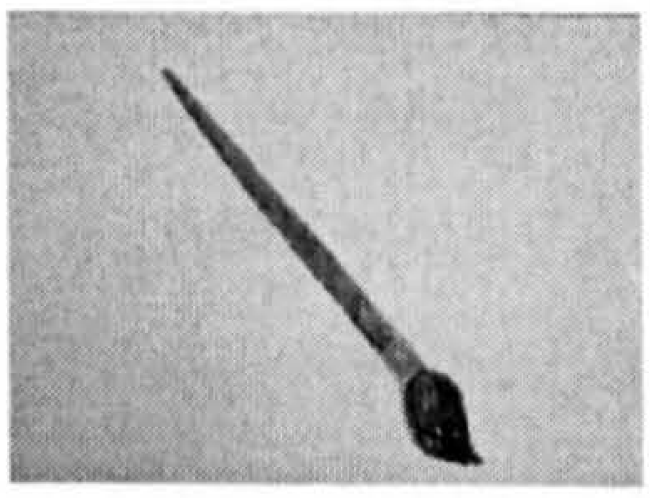
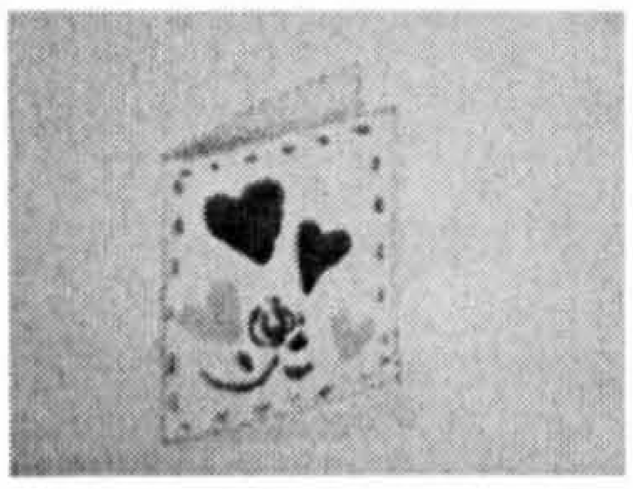
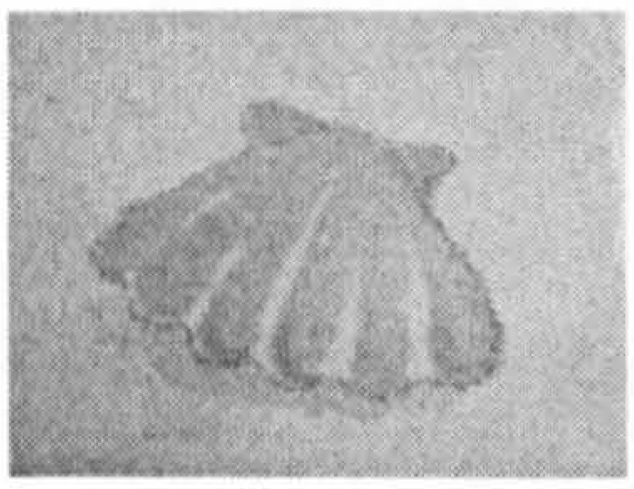
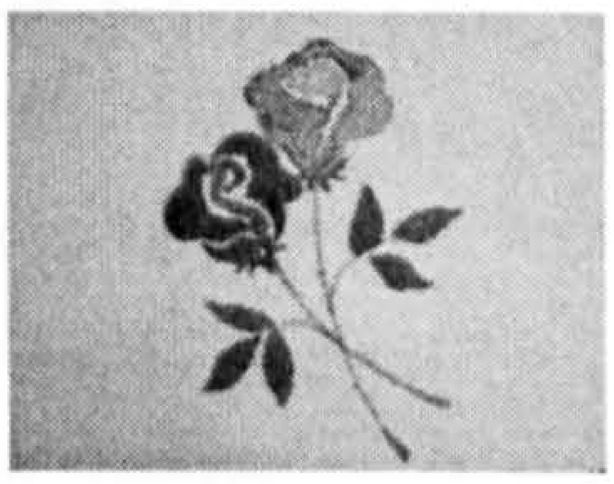
A.

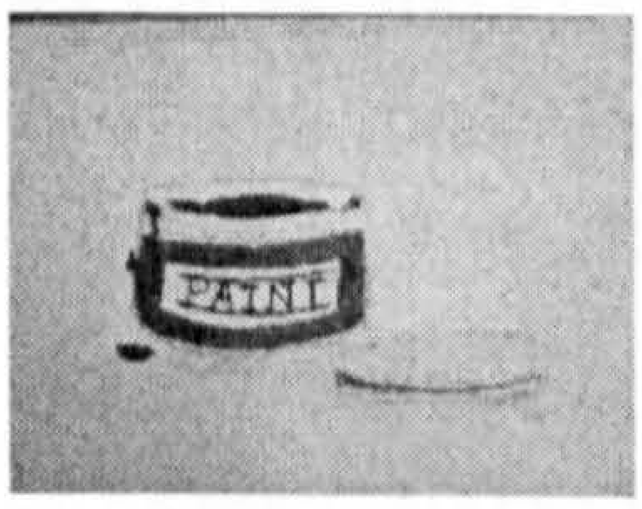


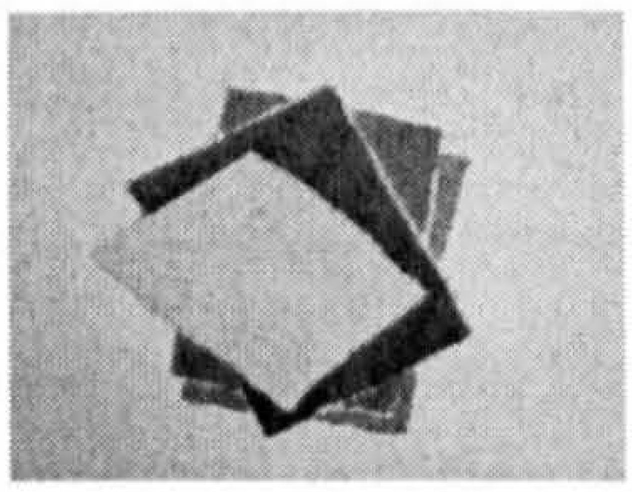
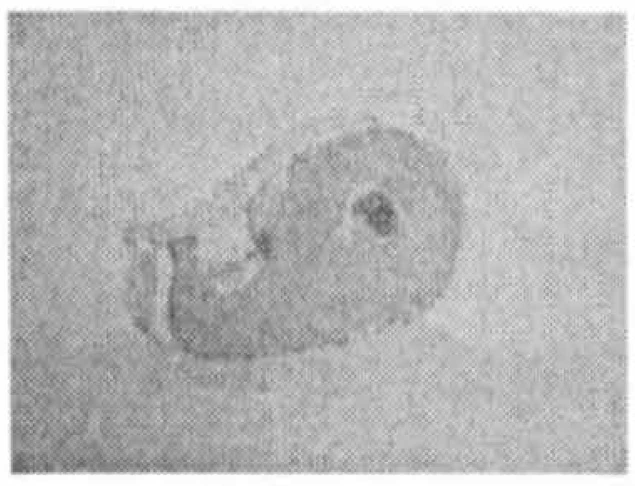



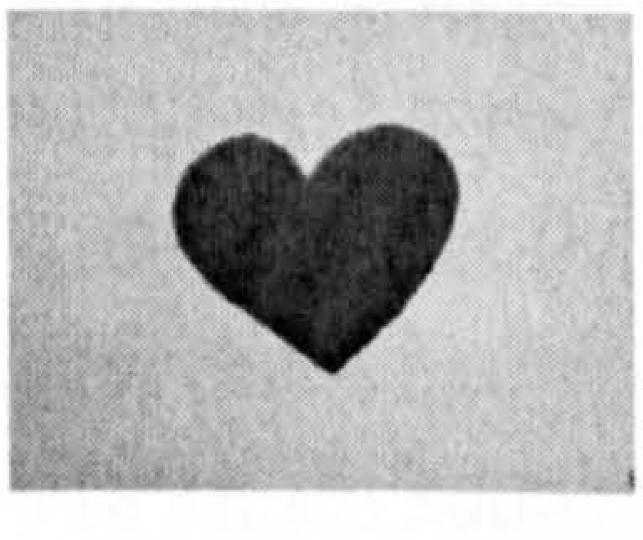
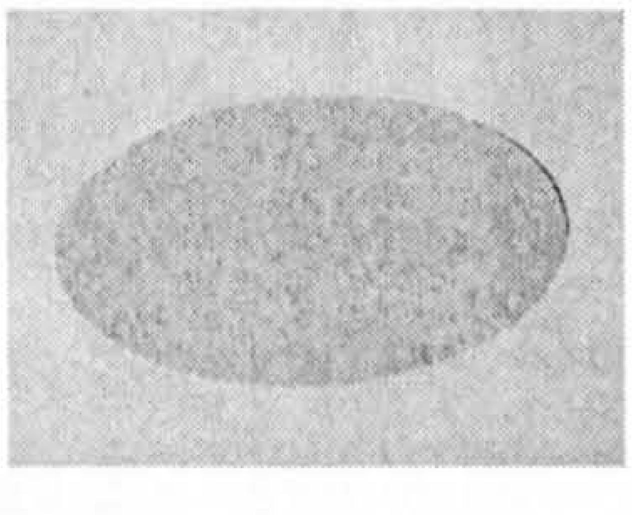
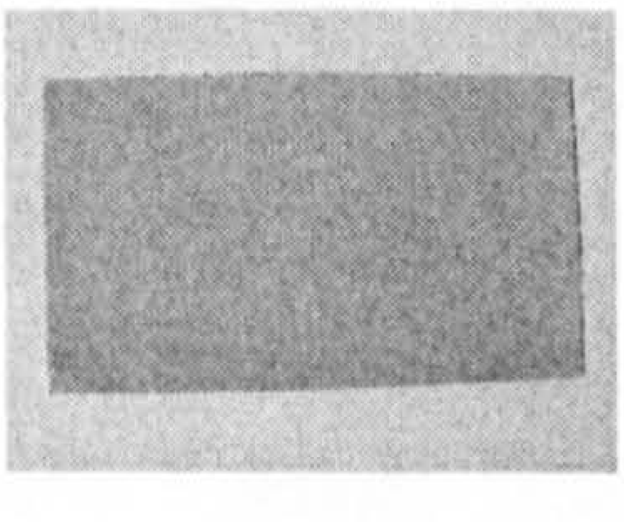
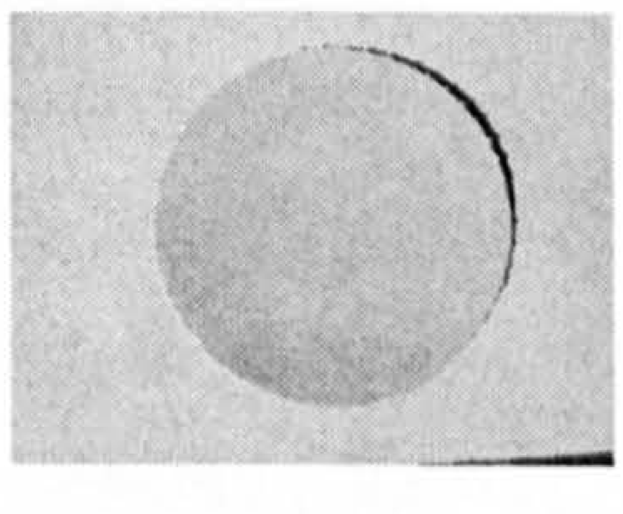
B.

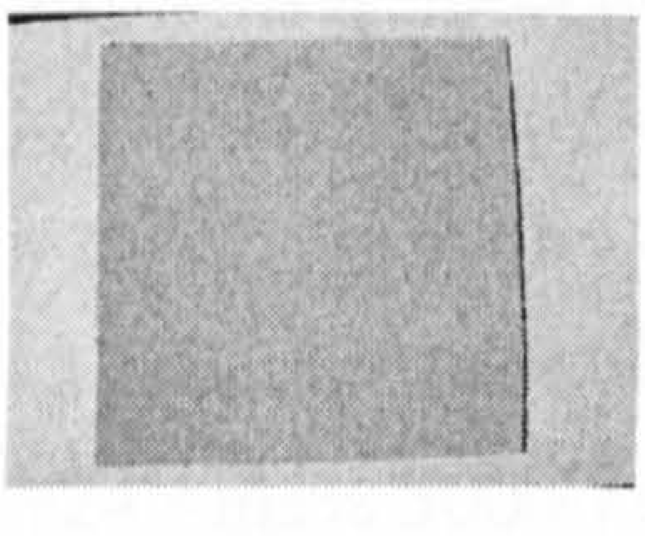
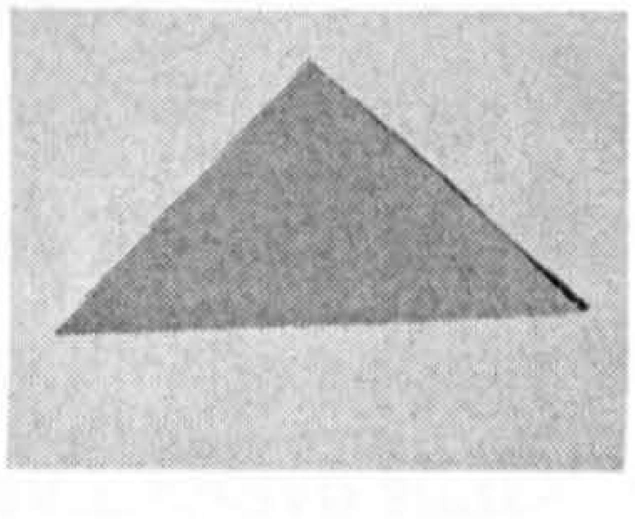







C.

Appendix G.i

Script of Online Vocabulary Assessment for Story 2

```
<ep> <n 30> <cr> <azk> <nfb> <fd 120> <t 6000> <id "keyboard"> <mr +Space> <id
#mouse> <umnr> <umpr> <mnr +#3> <mpr +#4> <dbc 255255255> <dwc 0> <vm
800,600,600,16,0> <eop>
0 "Press SPACEBAR to start.";
+101 <ms% 500> / * <wav 2> <svp start> "sun" / <bmp> "sun" / <ms% 1000> / ;
-102 <ms% 500> / * <wav 2> <svp start> "lip"/ <bmp> "book" / <ms% 1000> / ;
-103 <ms% 500> / * <wav 2> <svp start> "banana"/ <bmp> "apple" / <ms% 1000> / ;
+104 <ms% 500> / * <wav 2> <svp start> "ice cream"/ <bmp> "ice cream" / <ms% 1000>
/;
0 "Press SPACEBAR to start.";
-1 <ms% 500> / <wav 2> <svp start> "star" / * <bmp> "moon" / <ms% 1000> / ;
+2 <ms% 500> / <wav 2> <svp start> "king" / * <bmp> "king" / <ms% 1000> / ;
+3 <ms% 500> / <wav 2> <svp start> "boat" / * <bmp> "boat" / <ms% 1000> /;
-4 <ms% 500> / <wav 2> <svp start> "frog" / * <bmp> "flag" / <ms% 1000> /;
-5 <ms% 500> / <wav 2> <svp start> "throw" / * <bmp> "castle" / <ms% 1000> /;
-6 <ms% 500> / <wav 2> <svp start> "rug" / * <bmp> "bug" / <ms% 1000> /;
+7 <ms% 500> / <wav 2> <svp start> "cup" / * <bmp> "cup" / <ms% 1000> /;
+8 <ms% 500> / <wav 2> <svp start> "ring" / * <bmp> "ring" / <ms% 1000> /;
-9 <ms% 500> / <wav 2> <svp start> "fork" / * <bmp> "spoon" / <ms% 1000> /;
+10 <ms% 500> / <wav 2> <svp start> "bowl" / * <bmp> "bowl" / <ms% 1000> /;
-11 <ms% 500> / <wav 2> <svp start> "ears" / * <bmp> "nose" / <ms% 1000> /;
-12 <ms% 500> / <wav 2> <svp start> "king" / * <bmp> "queen" / <ms% 1000> / ;
-13 <ms% 500> / <wav 2> <svp start> "wag" / * <bmp> "wig" / <ms% 1000> / ;
+14 <ms% 500> / <wav 2> <svp start> "bug" / * <bmp> "bug" / <ms% 1000> /;
+15 <ms% 500> / <wav 2> <svp start> "rug" / * <bmp> "rug" / <ms% 1000> /;
+16 <ms% 500> / <wav 2> <svp start> "crown" / * <bmp> "crown" / <ms% 1000> /;
+17 <ms% 500> / <wav 2> <svp start> "shoe" / * <bmp> "shoe" / <ms% 1000> / ;
+18 <ms% 500> / <wav 2> <svp start> "castle" / * <bmp> "castle" / <ms% 1000> /;
+19 <ms% 500> / <wav 2> <svp start> "flag" / * <bmp> "flag" / <ms% 1000> / ;
+20 <ms% 500> / <wav 2> <svp start> "moon" / * <bmp> "moon" / <ms% 1000> /;
-21 <ms% 500> / <wav 2> <svp start> "folder" / * <bmp> "ring" / <ms% 1000> /;
+22 <ms% 500> / <wav 2> <svp start> "spoon" / * <bmp> "spoon" / <ms% 1000> /;
-23 <ms% 500> / <wav 2> <svp start> "rock" / * <bmp> "bowl" / <ms% 1000> /;
+24 <ms% 500> / <wav 2> <svp start> "wig" / * <bmp> "wig" / <ms% 1000> /;
-25 <ms% 500> / <wav 2> <svp start> "sprout" / * <bmp> "crown" / <ms% 1000> /;
```

+26 <ms% 500> / <wav 2> <svp start> "queen" / * <bmp> "queen" / <ms% 1000> /;
0"The end. Thank you." <end>;

Appendix G.ii

Script of Online Vocabulary Assessment for Story 3

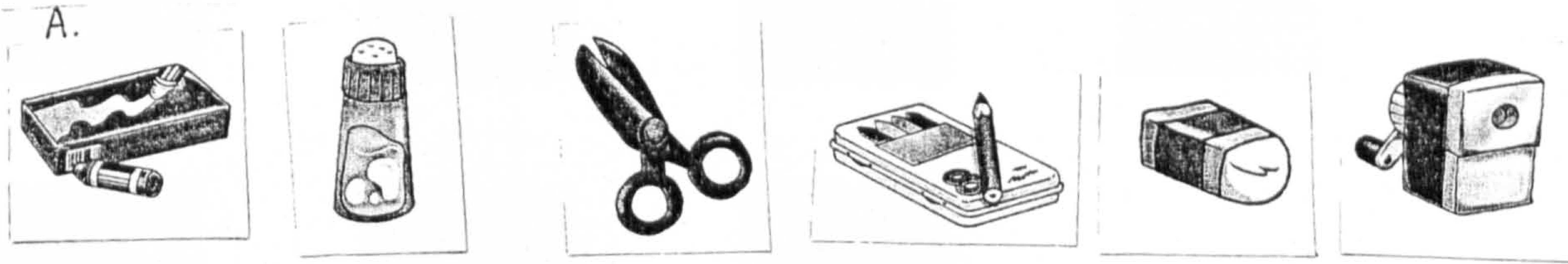
```
<ep> <n 30> <cr> <azk> <nfb> <fd 120> <t 6000> <id "keyboard"> <mr +Space> <id
#mouse> <umnr> <umpr> <mnr +#3> <mpr +#4> <dbc 255255255> <dwc 0> <vm
800,600,600,16,0> <eop>
0 "Press SPACEBAR to start.";
+101 <ms% 500> / * <wav 2> <svp start> "heart" / <bmp> "heart" / <ms% 1000> / ;
-102 <ms% 500> / * <wav 2> <svp start> "gold"/ <bmp> "goldfish" / <ms% 1000> / ;
-103 <ms% 500> / * <wav 2> <svp start> "watch"/ <bmp> "card" / <ms% 1000> / ;
+104 <ms% 500> / * <wav 2> <svp start> "brush"/ <bmp> "brush" / <ms% 1000> /;
0 "Press SPACEBAR to start.";
-1 <ms% 500> / * <wav 2> <svp start> "sail" / <bmp> "tail" / <ms% 1000> / ;
+2 <ms% 500> / * <wav 2> <svp start> "round"/ <bmp> "round" / <ms% 1000> /;
-3 <ms% 500> / * <wav 2> <svp start> "lilies"/ <bmp> "roses" / <ms% 1000> / ;
-4 <ms% 500> / * <wav 2> <svp start> "oval"/ <bmp> "square" / <ms% 1000> /;
+5 <ms% 500> / * <wav 2> <svp start> "shell"/ <bmp> "shell" / <ms% 1000> / ;
+6 <ms% 500> / * <wav 2> <svp start> "puppies" / <bmp> "puppies" / <ms% 1000> / ;
+7 <ms% 500> / * <wav 2> <svp start> "tape"/ <bmp> "tape" / <ms% 1000> /;
-8 <ms% 500> / * <wav 2> <svp start> "booty"/ <bmp> "boot" / <ms% 1000> / ;
+9 <ms% 500> / * <wav 2> <svp start> "fishbowl"/ <bmp> "fishbowl" / <ms% 1000> /;
+10 <ms% 500> / * <wav 2> <svp start> "paint"/ <bmp> "paint" / <ms% 1000> / ;
+11 <ms% 500> / * <wav 2> <svp start> "rectangle"/ <bmp> "rectangle" / <ms% 1000> / ;
-12 <ms% 500> / * <wav 2> <svp start> "tape" / <bmp> "paper" / <ms% 1000> / ;
+13 <ms% 500> / * <wav 2> <svp start> "square" / <bmp> "square" / <ms% 1000> / ;
+14 <ms% 500> / * <wav 2> <svp start> "bone"/ <bmp> "bone" / <ms% 1000> / ;
+15 <ms% 500> / * <wav 2> <svp start> "bows"/ <bmp> "bows" / <ms% 1000> / ;
-16 <ms% 500> / * <wav 2> <svp start> "round"/ <bmp> "oval" / <ms% 1000> /;
+17 <ms% 500> / * <wav 2> <svp start> "tail" / <bmp> "tail" / <ms% 1000> / ;
-18 <ms% 500> / * <wav 2> <svp start> "coop"/ <bmp> "doghouse" / <ms% 1000> / ;
+19 <ms% 500> / * <wav 2> <svp start> "triangle" / <bmp> "triangle" / <ms% 1000> / ;
+20 <ms% 500> / * <wav 2> <svp start> "paper"/ <bmp> "paper" / <ms% 1000> / ;
-21 <ms% 500> / * <wav 2> <svp start> "share"/ <bmp> "shell" / <ms% 1000> /;
-22 <ms% 500> / * <wav 2> <svp start> "pencil"/ <bmp> "brush" / <ms% 1000> / ;
-23 <ms% 500> / * <wav 2> <svp start> "doghouse"/ <bmp> "fishbowl" / <ms% 1000> / ;
-24 <ms% 500> / * <wav 2> <svp start> "kittens"/ <bmp> "puppies" / <ms% 1000> / ;
+25 <ms% 500> / * <wav 2> <svp start> "oval"/ <bmp> "oval" / <ms% 1000> / ;
-26 <ms% 500> / * <wav 2> <svp start> "triangle" / <bmp> "rectangle" / <ms% 1000> / ;
```

0"The end. Thank you." <end>;

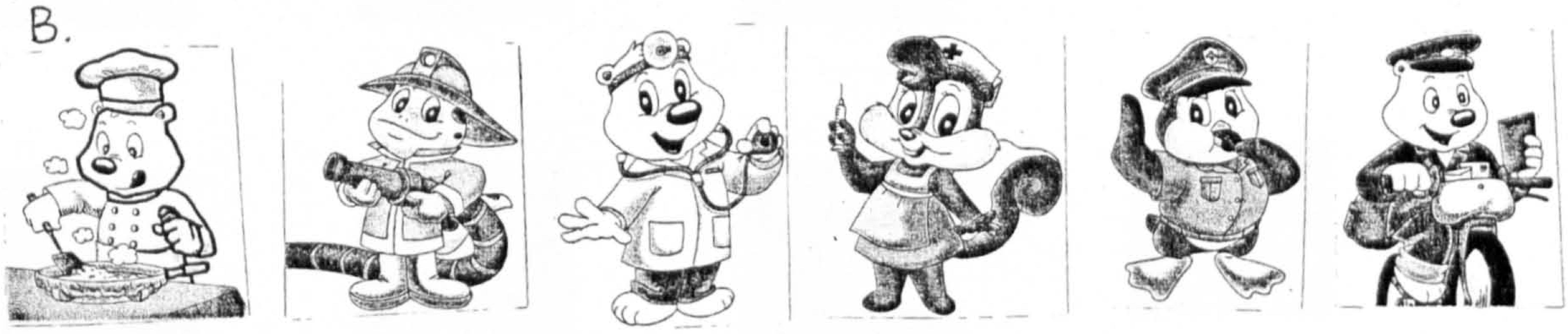
Appendix H.i

Word Review I

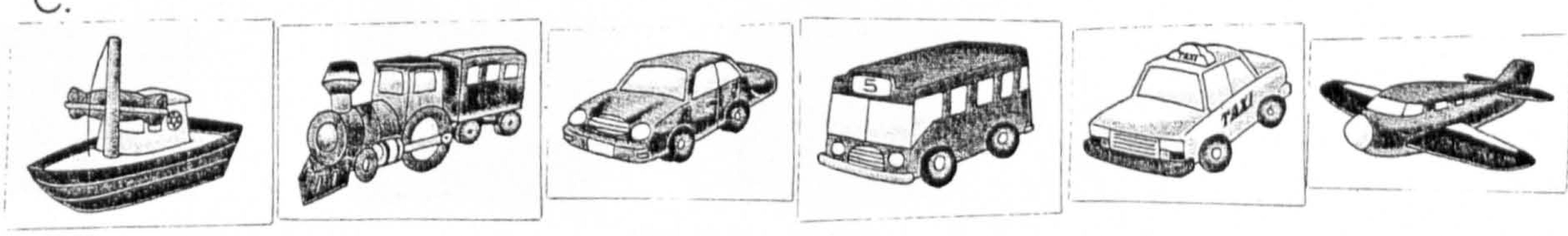
A.



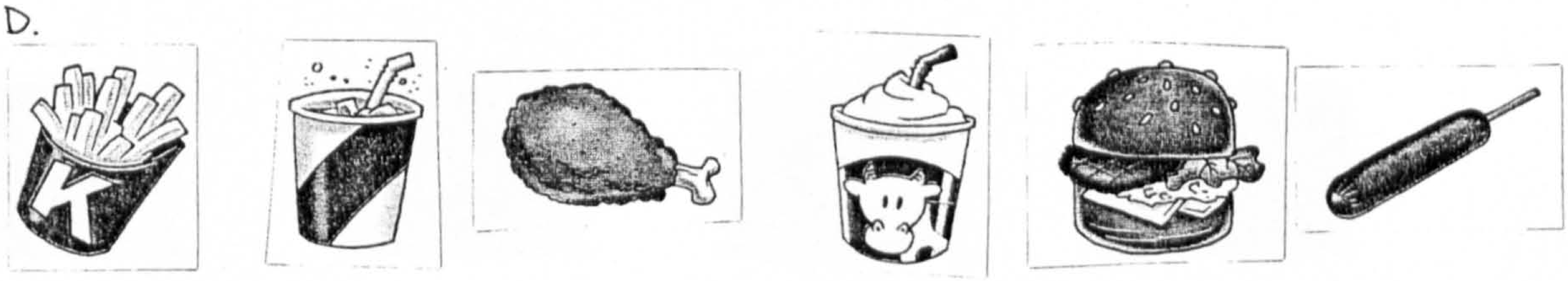
B.



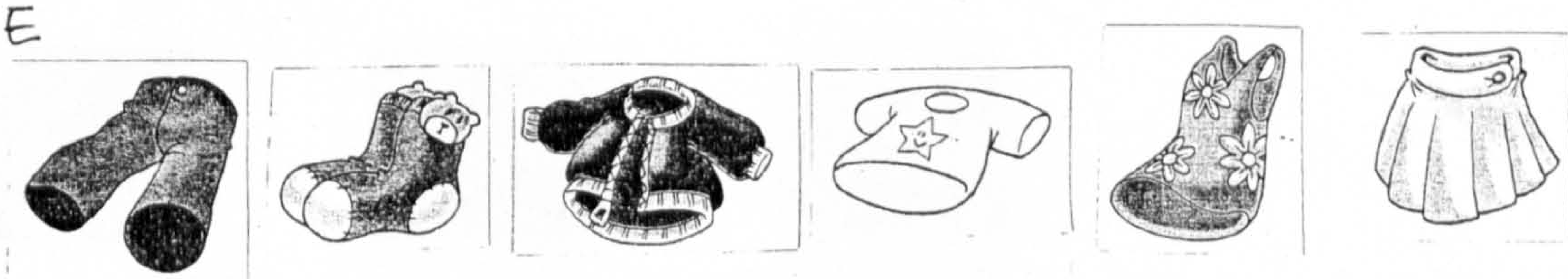
C.



D.



E



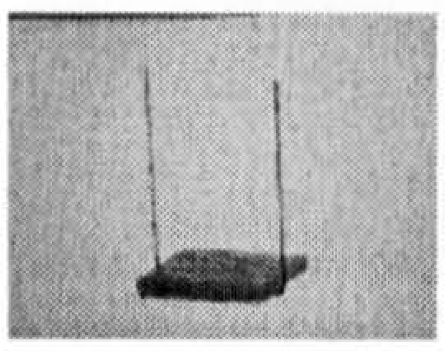
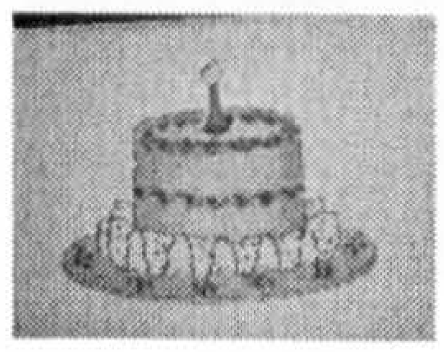
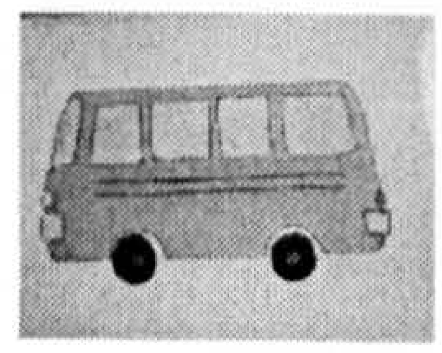
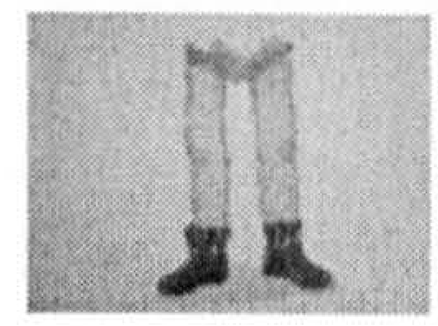
F

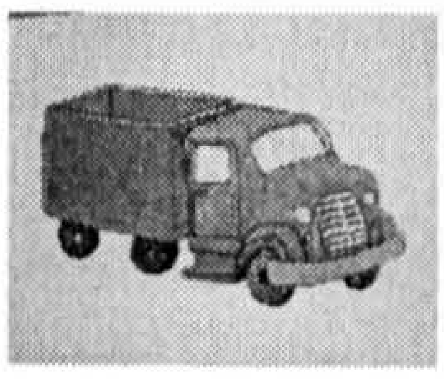
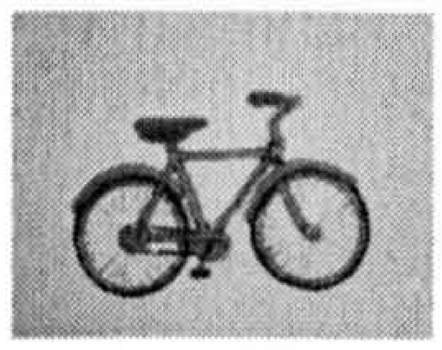
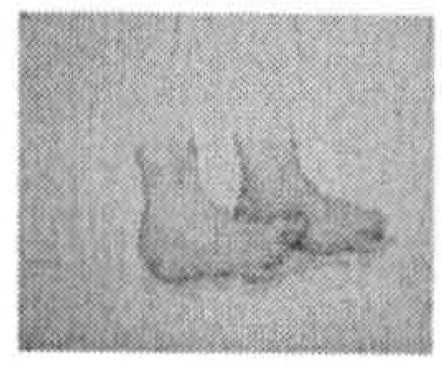
G

Appendix H.ii

Word Review II

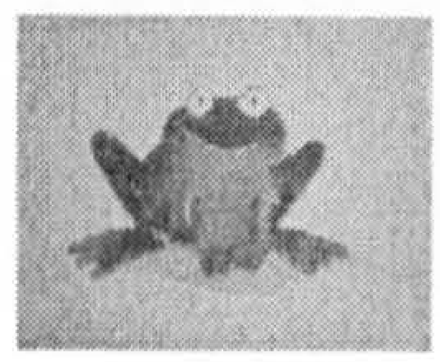
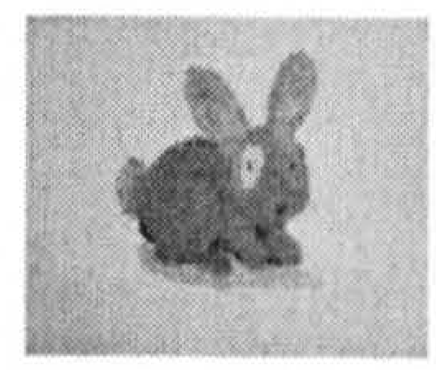
A.

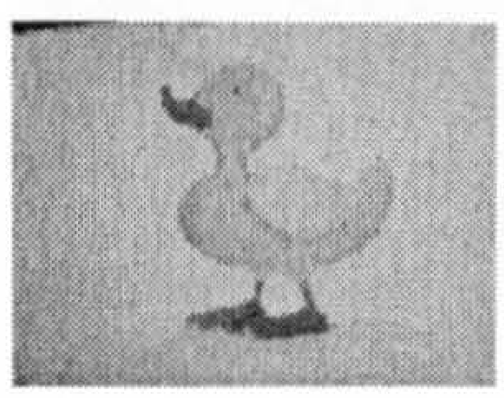
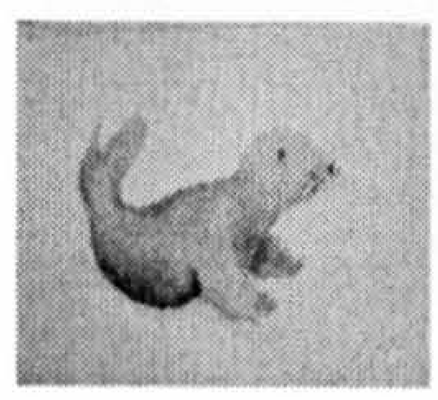





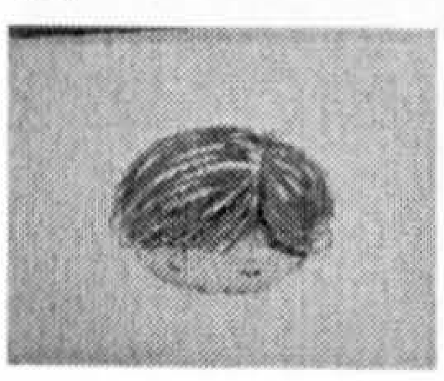
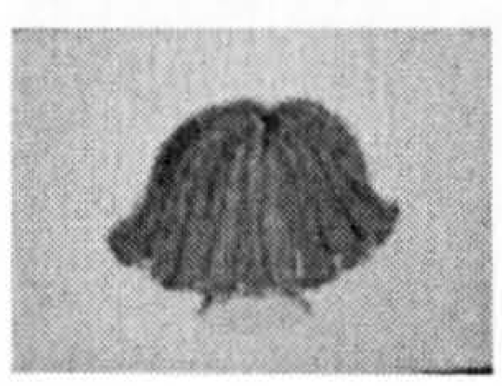
B.

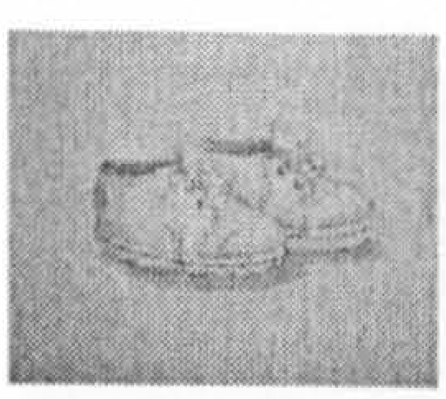
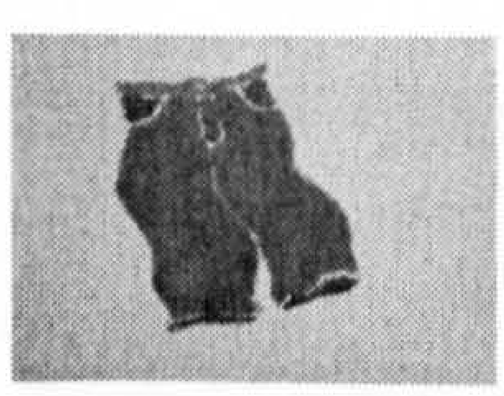
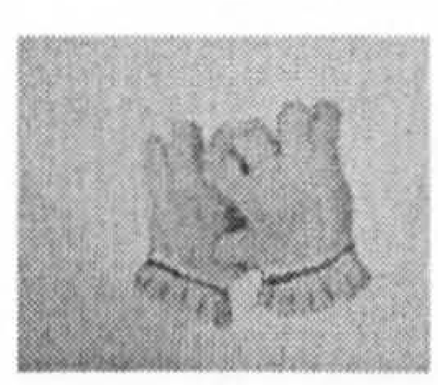


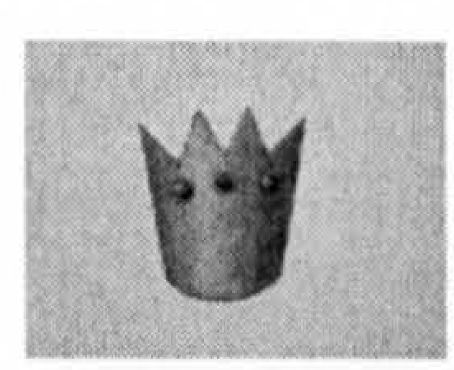
C.

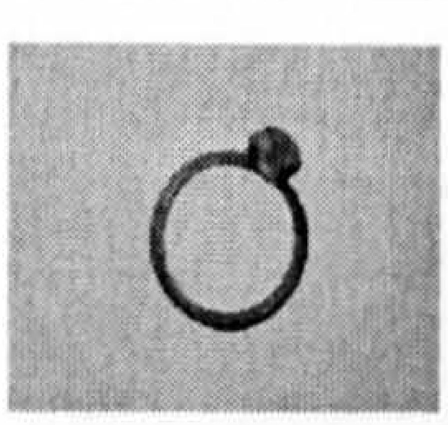
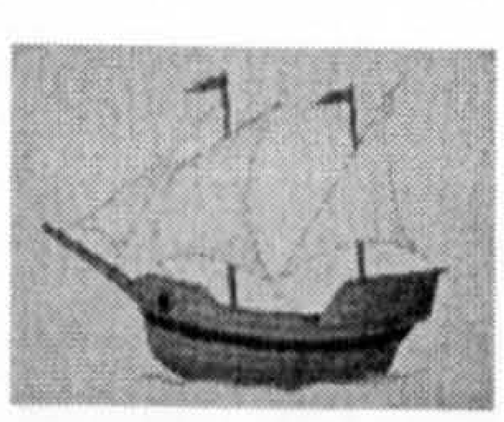



D.



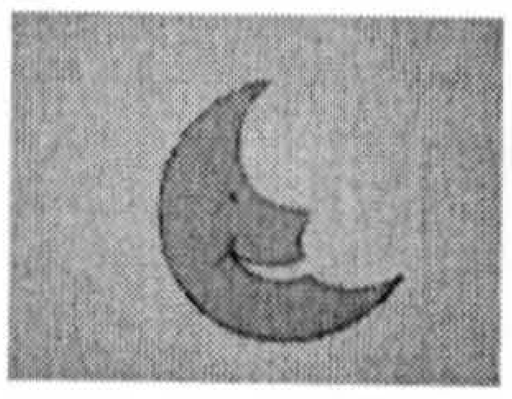
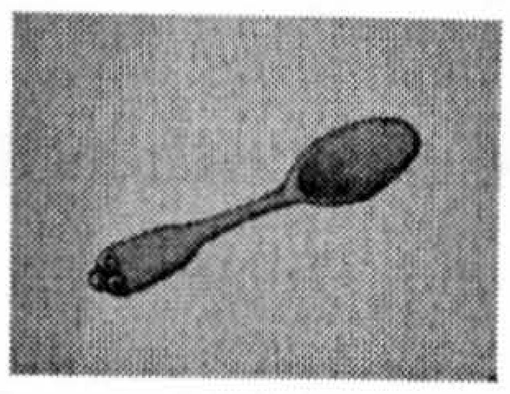



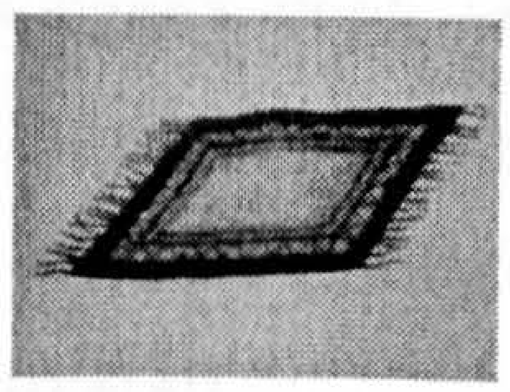
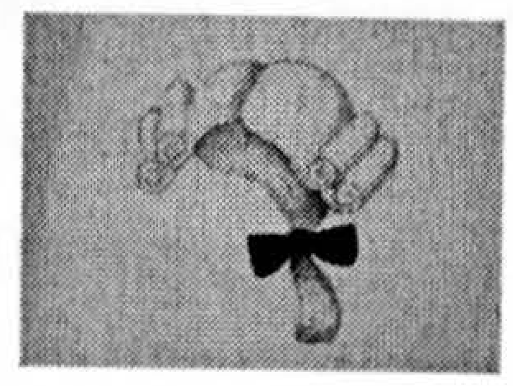

(Appendix H.ii continues)

(Appendix H.ii
continued)

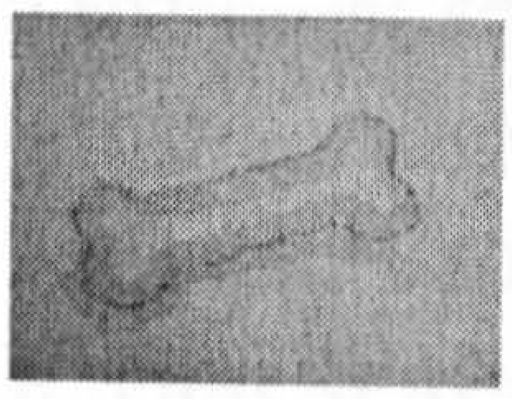
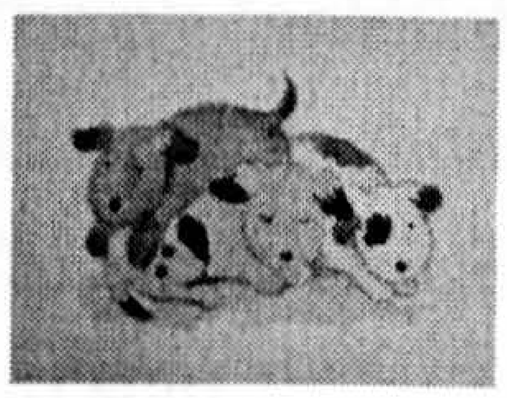
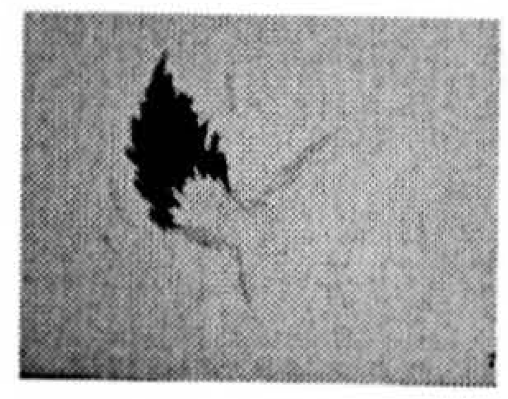
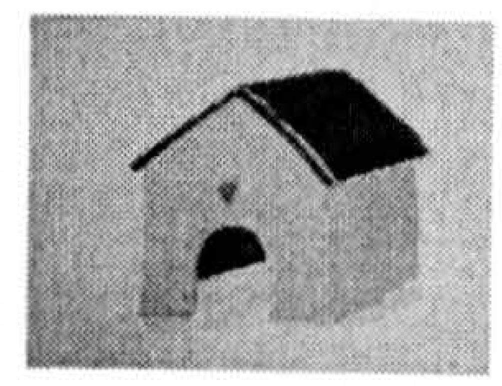
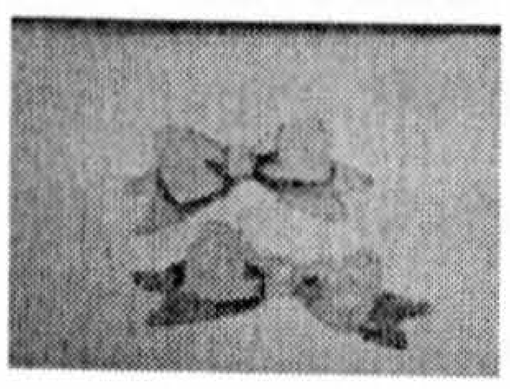
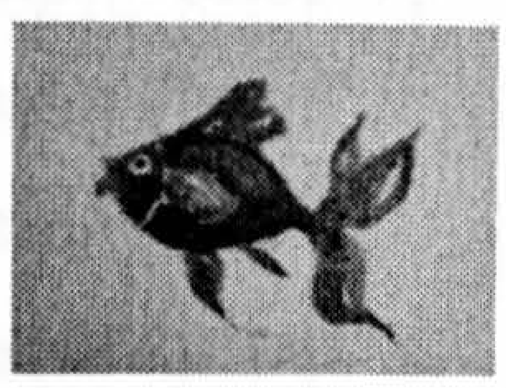
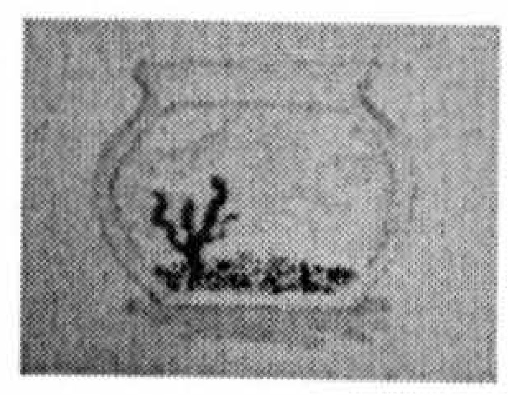
E.

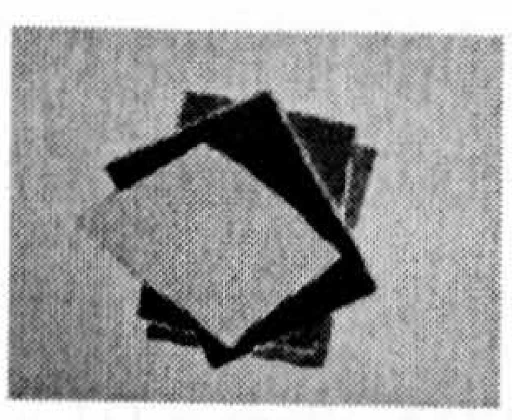
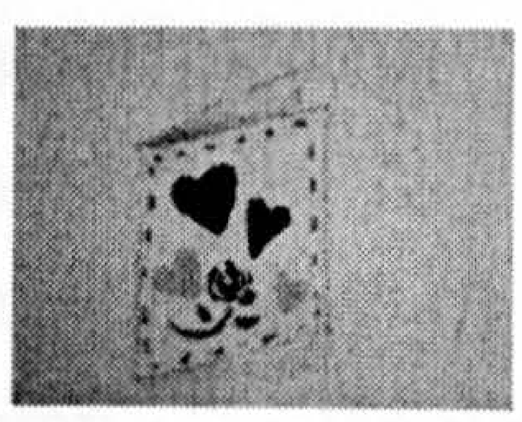
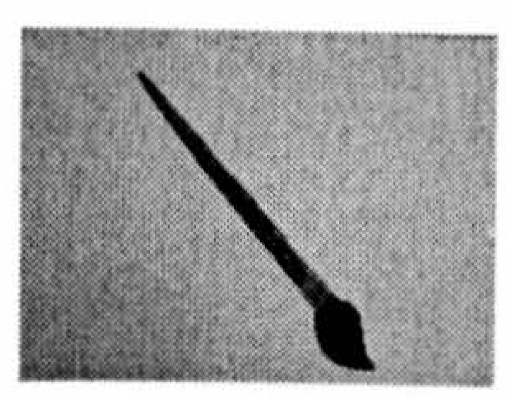
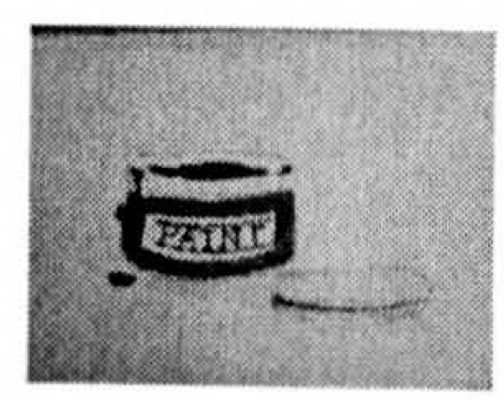


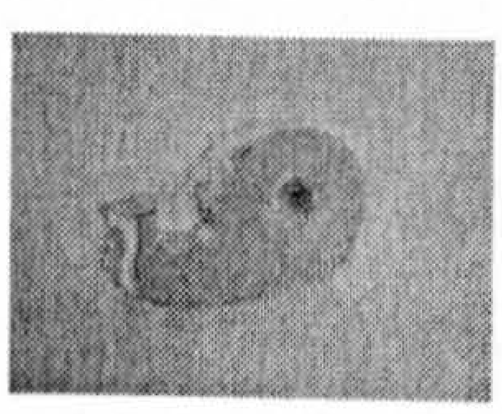
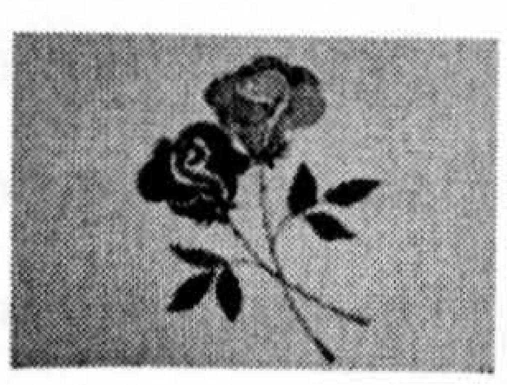
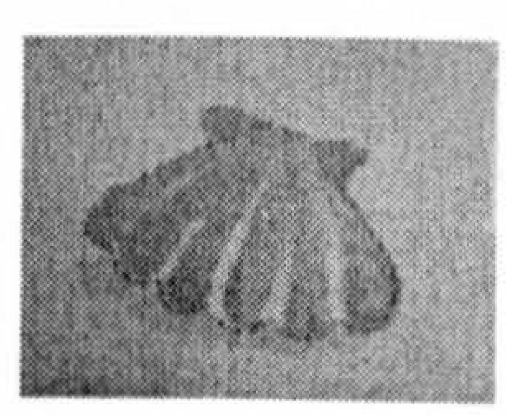



F.

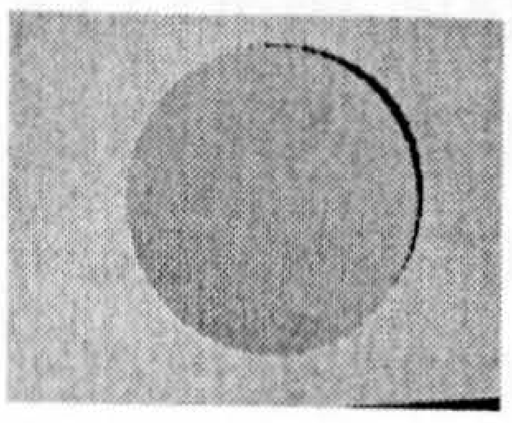
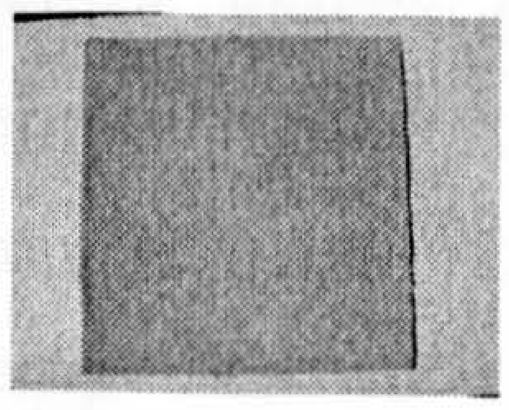
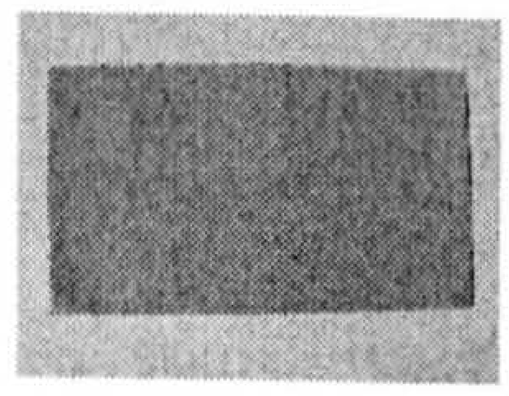
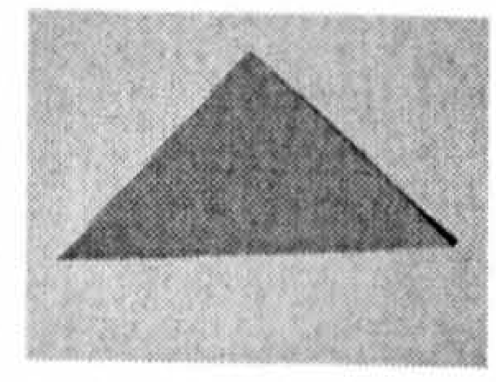
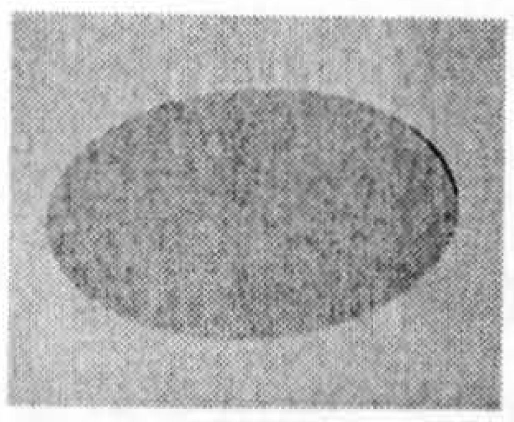
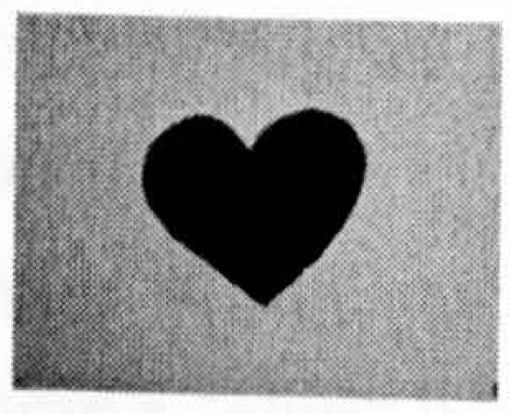








G.

H.

Appendix H.iii

Script of Productive Vocabulary Test for Picture-Naming

```
<ep> <n 50> <cr> <nfb> <fd 180> <t 6000> <id "keyboard"> <mr +Space> <id #mouse>
<umnr> <umpr> <zil> <zor> <vzk +#3> <vzk +#4> <mnr +#3> <mpr +#4> <id
"DigitalVox"> <vzk +DigitalVox> <mpr +DigitalVox> <id RecordVocal 1500> <dbc
255255255> <dwc 0> <vm 800,600,600,16,0> <eop>
0 "Press SPACEBAR to start.";
+101 <ms% 500> /* <bmp> "fishbowl" /;
+102 <ms% 500> /* <bmp> "goldfish" /;
+103 <ms% 500> /* <bmp> "doghouse" /;
+104 <ms% 500> /* <bmp> "triangle" /;
0 "Press SPACEBAR to start.";
+1 <ms% 500> /* <bmp> "boot" /;
+2 <ms% 500> /* <bmp> "boat" /;
+3 <ms% 500> /* <bmp> "tail" /;
+4 <ms% 500> /* <bmp> "jump" /;
+5 <ms% 500> /* <bmp> "hair" /;
+6 <ms% 500> /* <bmp> "brush" /;
+7 <ms% 500> /* <bmp> "paint" /;
+8 <ms% 500> /* <bmp> "tape" /;
+9 <ms% 500> /* <bmp> "round" /;
+10 <ms% 500> /* <bmp> "rectangle" /;
+11 <ms% 500> /* <bmp> "heart" /;
+12 <ms% 500> /* <bmp> "roses" /;
+13 <ms% 500> /* <bmp> "square" /;
+14 <ms% 500> /* <bmp> "swing" /;
+15 <ms% 500> /* <bmp> "sweater" /;
+16 <ms% 500> /* <bmp> "woollen hat" /;
+17 <ms% 500> /* <bmp> "gloves" /;
+18 <ms% 500> /* <bmp> "castle" /;
+19 <ms% 500> /* <bmp> "queen" /;
+20 <ms% 500> /* <bmp> "crown" /;
+21 <ms% 500> /* <bmp> "ring" /;
+22 <ms% 500> /* <bmp> "rug" /;
+23 <ms% 500> /* <bmp> "wig" /;
+24 <ms% 500> /* <bmp> "bug" /;
+25 <ms% 500> /* <bmp> "clap" /;
```

```
+26 <ms% 500> /* <bmp> "flag" /;
+27 <ms% 500> /* <bmp> "feet" /;
+28 <ms% 500> /* <bmp> "spoon" /;
+29 <ms% 500> /* <bmp> "paper" /;
+30 <ms% 500> /* <bmp> "oval" /;
+31 <ms% 500> /* <bmp> "puppies" /;
+32 <ms% 500> /* <bmp> "sneakers" /;
+33 <ms% 500> /* <bmp> "shell" /;
+34 <ms% 500> /* <bmp> "truck" /;
+35 <ms% 500> /* <bmp> "king" /;
+36 <ms% 500> /* <bmp> "paper" /;
+37 <ms% 500> /* <bmp> "bike" /;
+38 <ms% 500> /* <bmp> "bone" /;
+39 <ms% 500> /* <bmp> "legs" /;
+40 <ms% 500> /* <bmp> "hop" /;
+41 <ms% 500> /* <bmp> "head" /;
+42 <ms% 500> /* <bmp> "cup" /;
+43 <ms% 500> /* <bmp> "card" /;
+44 <ms% 500> /* <bmp> "seal" /;
+45 <ms% 500> /* <bmp> "frog" /;
+46 <ms% 500> /* <bmp> "bunny" /;
0"The end.  Thank you." <end>;
```

Appendix I.i

Letter of Consent for the Pilot Study

Dear Parents,

I am a PhD student at Newcastle University and will be working in your child's school. I am trying to investigate how children process and acquire English vocabulary, via the practice of a storytelling programme. I will need to video-record the classroom and administer assessments to the pupils. The videotapes and the test results will be used for research purposes only.

I would appreciate it very much if you would agree to my video-recording the classroom and administering assessments to your child. If you do not want your child to be video-recorded or would like to have your child excluded from the assessments, please indicate by ticking the box below. I will accordingly not include your child in the research project.

Yours sincerely,

Yu-cheng Sieh

Name of Student ______________________

☐ Yes, I agree to my child's participation in the research project.

☐ No, I do not agree to my child's participation in the research project.

Signature ______________________

Appendix I.ii

Letter of Consent for the Main Study

Dear Parents,

I am a PhD student at Newcastle University and will be working in your child's school for three months. I am trying to investigate how children process and acquire English vocabulary, via the practice of a storytelling programme. I will need to video-record the classroom and administer some assessments to the pupils. The videotapes and the test results will be used for research purposes only.

I would appreciate it very much if you would agree to my video-recording the classroom and administering of assessments to your child. If you do not want your child to be video-recorded or would like to have your child excluded from the assessments, please indicate by ticking the box below. I will accordingly not include your child in the research project.

Yours sincerely,

Yu-cheng Sieh

Name of Student ______________________

☐ Yes, I agree to my child's participation in the research project.

☐ No, I do not agree to my child's participation in the research project.

Signature ________________________

Appendix J.i—Intercorrelations Between Vocabulary Assessment Scores Within the Same Category and Each Story

Measure	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
1. WR I	—	.631**													
2. WR II		—													
Story 1															
3. Pre-test			—	.654**	.522**										
4. Post-test					.791**										
5. Follow-up					—										
Story 2															
6. Pre-test						—	.643**	.615**	.592**	.459**					
7. In-class test								.773**	.824**	.740**					
8. Post-test									.787**	.613**					
9. Follow-up										.621**					
10. Online test										—					
Story 3															
11. Pre-test											—	.694**	.559**	.522**	.292*
12. In-class test													.776**	.807**	.560**
13. Post-test														.848**	.629**
14. Follow-up															.619**
15. Online test															—

Note. WR = Word Review.

*$p < .05$. **$p < .01$.

Appendix J.ii

Intercorrelations Between Chinese Phonological Sensitivity Task Scores

Chinese PS task	1	2	3	4	5
1. Rhyme detection	—	.569**	.487**	.541**	.208
2. Head detection		—	.412**	.461**	.054
3. Rhyme & head detection			—	.508**	.339**
4. Rhyme & head production				—	.280*
5. Initial sound isolation					—

Note. PS = phonological sensitivity.

*$p < .05$. **$p < .01$.

Appendix J.iii

Intercorrelations Between English Phonological Sensitivity Task Scores

English PS task	1	2	3	4	5
1. Rime detection	—	.597**	.592**	.659**	.296*
2. Head detection		—	.582**	.600**	.403**
3. Rime & head detection			—	.691**	.430**
4. Rime & head production				—	.446**
5. Initial consonant isolation					—

Note. PS = phonological sensitivity.
**p* < .05. ***p* < .01.

Appendix K

Correlations Between Young Learners' Overall and Individual Chinese Phonological Sensitivity Performances and Vocabulary Assessment Scores

Vocabulary assessment \ PS task	Overall	Rhyme detection	Head detection	Rhyme & head detection	Rhyme & head production	Initial consonant isolation
Word Review I	.259*	.286*	.054	.234	.276*	.050
Word Review II	.523**	.438**	.309*	.438**	.531**	.096
Story 1						
Pretest	.193	.356**	.142	.155	.231	-.178
Posttest	.329**	.359**	.181	.319*	.367**	-.100
Follow-up test	.476**	.391**	.369**	.408**	.444**	.052
Story 2						
Pretest	.283*	.327**	.069	.211	.265*	.151
In-class test	.401**	.452**	.275*	.253*	.436**	.030
Posttest	.435**	.389**	.257*	.404**	.383**	.086
Follow-up test	.462**	.401**	.261*	.359**	.484**	.105
Online test	.377**	.362**	.154	.304*	.425**	.053
RT	-.056	-.150	.014	.098	-.186	-.003
Story 3						
Pretest	.250*	.389**	.189	.159	.209	.012
In-class test	.407**	.459**	.276*	.325**	.406**	-.014
Posttest	.372**	.396**	.305*	.325**	.376**	-.094
Follow-up test	.450**	.391**	.282*	.373**	.472**	.040
Online test	.377**	.328**	.273*	.350**	.346**	.013
RT	-.033	-.159	.051	-.058	-.055	.100

Note. PS = phonological sensitivity.

**p* < .05. ***p* < .01.

Appendix L—Intercorrelations Between Phonological Sensitivity Task Scores of the Two Languages

	2	3	4	5	6	7	8	9	10	11	12
1. English PS overall	.745**	.756**	.822**	.868**	.726**	.679**	.543**	.441**	.571**	.613**	.211
2. Rhyme Detection		.597**	.592**	.659**	.296*	.477**	.403**	.355**	.335**	.423**	.198
3. Head Detection			.582**	.600**	.403**	.552**	.344**	.380**	.542**	.459**	.166
4. Rhyme & Head Detection				.691**	.430**	.659**	.500**	.352**	.539**	.606**	.304*
5. Rhyme & Head Production					.446**	.568**	.480**	.282*	.473**	.515**	.246
6. Initial Consonant Isolation						.440**	.383**	.382**	.382**	.411**	-.015
7. Chinese PS overall							.735**	.666**	.809**	.816**	.517**
8. Rhyme Detection								.569**	.487**	.541**	.208
9. Head Detection									.412**	.461**	.054
10. Rhyme & Head Detection										.508**	.339**
11. Rhyme & Head Production											.280*
12. Initial Sound Isolation											—

Note. PS = phonological sensitivity.

*$p < .05$. **$p < .01$.

Appendix M

Summary of Stepwise Regression Analysis Results for Best Predictors of Young Learners' Vocabulary Scores in Each Vocabulary Assessment

Vocabulary assessment	Predictor	R^2 change	F-ratio
Story 1			
Pretest	Word Review I	.525	67.350***
Posttest	Word Review I	.557	76.782***
	English PS	.121	63.158***
Follow-up test	English PS	.427	45.407***
	Word Review I	.175	45.391***
	L2 instruction length	.029	33.619***
Story 2			
Pretest	Word Review I	.468	53.751***
	English PS	.039	30.888***
In-class test	Word Review I	.498	60.513***
	English PS	.106	45.719***
Posttest	Word Review I	.383	37.936***
	English PS	.153	34.760***
Follow-up test	English PS	.405	41.597***
	Word Review I	.111	32.002***
Online test	English PS	.292	25.179***
	Word Review I	.119	20.914***
Story 3			
Pretest	Word Review I	.276	23.266***
	Word Review I	.383	37.874***
In-class test	English PS	.146	33.658***
	English PS	.442	48.330***
Posttest	Word Review I	.078	32.529***
	English PS	.464	52.828***
Follow-up test	L2 instruction length	.088	36.977***
	Word Review I	.035	28.003***
Online test	English PS	.325	29.398***
Word Review II	English PS	.532	69.403***
	Word Review I	.137	60.562***
	L2 instruction length	.026	44.733***

Note. PS = phonological sensitivity.
***$p < .001$.

Appendix N

Parallel Hierarchical Multiple Regression Analysis for Phonological Processing Skills Predicting Young Learners' Vocabulary Scores as a Function of Determining Unique Contributions of PM and English PS to Vocabulary Scores

Step	R^2 change	F change	Step	R^2 change	F change
Story 1 Posttest					
5 PM	.004	0.592	5 PM	.019	2.912
6 PS	.104	20.087***	6 PS	.088	17.078***
5 PS	.107	21.055***	5 PS	.107	21.055***
6 PM	.001	0.140	6 PM	.000	0.092
Story 1 Follow-up test					
5 PM	.007	0.982	5 PM	.013	1.793
6 PS	.069	11.144**	6 PS	.063	10.242**
5 PS	.076	12.503***	5 PS	.076	12.503***
6 PM	.000	0.014	6 PM	.000	0.034
Story 2 Pretest					
5 PM	.018	2.070	5 PM	.051	6.317*
6 PS	.016	1.908	6 PS	.007	0.853
5 PS	.025	2.977	5 PS	.025	2.977
6 PM	.009	1.032	6 PM	.032	4.019*
Story 2 In-class test					
5 PM	.028	3.901	5 PM	.039	5.638*
6 PS	.035	5.276*	6 PS	.027	4.155*
5 PS	.051	7.571**	5 PS	.051	7.571**
6 PM	.012	1.793	6 PM	.015	2.355
Story 2 Posttest					
5 PM	.003	0.309	5 PM	.001	0.136
6 PS	.083	11.122**	6 PS	.072	9.378**
5 PS	.069	9.056**	5 PS	.069	9.056**
6 PM	.017	2.233	6 PM	.004	0.550
Story 2 Follow-up test					
5 PM	.035	3.764	5 PM	.054	6.100*
6 PS	.069	8.456**	6 PS	.055	6.875*
5 PS	.092	11.224***	5 PS	.092	11.224***
6 PM	.011	1.407	6 PM	.017	2.153
Story 2 Online test					
5 PM	.013	1.143	5 PM	.026	2.423
6 PS	.053	5.106*	6 PS	.043	4.178*
5 PS	.064	6.222*	5 PS	.064	6.222*
6 PM	.002	0.198	6 PM	.006	0.562

(Appendix N continues)

(Appendix N continued)

Story 3 In-class test					
5 PM	.022	2.552	5 PM	.033	3.944
6 PS	.062	8.064**	6 PS	.053	6.891*
5 PS	.078	10.260**	5 PS	.078	10.260**
6 PM	.006	0.722	6 PM	.008	0.985
Story 3 Posttest					
5 PM	.072	7.093**	5 PM	.077	7.682**
6 PS	.144	18.638***	6 PS	.131	16.719***
5 PS	.192	24.012***	5 PS	.192	24.012***
6 PM	.024	3.050	6 PM	.016	2.060
Story 3 Follow-up test					
5 PM	.035	3.786	5 PM	.030	3.230
6 PS	.118	16.287***	6 PS	.117	15.916***
5 PS	.146	20.087***	5 PS	.146	20.087***
6 PM	.007	1.000	6 PM	.001	0.189
Story 3 Online test					
5 PM	.012	0.960	5 PM	.035	2.892
6 PS	.115	10.772**	6 PS	.096	9.023**
5 PS	.127	12.092***	5 PS	.127	12.092***
6 PM	.000	0.014	6 PM	.004	0.365
Word Review II					
5 PM	.034	4.981*	5 PM	.045	6.767*
6 PS	.102	19.876***	6 PS	.091	17.684***
5 PS	.128	24.689***	5 PS	.128	24.689***
6 PM	.008	1.571	6 PM	.008	1.518

Note. PM= phonological memory; PS = phonological sensitivity.

The variation in the parallel hierarchical multiple regression analyses is the entry of the score of nonword repetition task 1 in the *left* panel and that of the mean score of the two nonword repetition tasks in the *right* panel.

$*p < .05$. $**p < .01$. $***p < .001$.

Appendix O

Summary of Stepwise Regression Analysis Results for Best Task Predictors of Young Learners' Vocabulary Scores in Each Vocabulary Assessment

Measure	Task Predictor	R^2 change	F change	F-ratio
Story 1				
Pretest	Chinese rhyme detection	.126	8.827**	8.827**
	Chinese initial sound isolation	.067	4.952*	7.175**
Posttest	English rhyme & head production	.310	27.382***	27.382***
	English initial consonant isolation	.088	8.813**	19.851***
	Chinese initial sound isolation	.042	4.405*	15.454***
Follow-up test	English rhyme & head production	.368	35.485***	35.485***
	English initial consonant isolation	.091	10.122**	25.457***
Story 2				
Pretest	English rhyme & head production	.249	20.253***	20.253***
	NWR 1 & 2	.071	6.304*	14.159***
In-class test	English rhyme & head production	.304	26.589***	26.589***
	English initial consonant isolation	.082	8.008**	18.826***
	NWR 1	.040	4.116*	14.574***
Posttest	English rhyme & head production	.375	36.594***	36.594***
Follow-up test	English rhyme & head production	.375	36.652***	36.652***
	NWR 1 & 2	.072	7.812**	24.279***
Online test	English initial consonant isolation	.291	24.993***	24.993***
	English rhyme & head production	.055	5.085*	15.876***
Story 3				
Pretest	English rhyme & head production	.171	12.567***	12.567***
In-class test	English rhyme & head production	.368	35.473***	35.473***
	NWR 1	.045	4.550*	21.044***

(Appendix O continues)

(Appendix O continued)

Posttest	English rhyme & head production	.382	37.721***	37.721***
	English initial consonant isolation	.127	15.578***	31.157***
	Chinese initial sound isolation	.037	4.863*	23.729***
	NWR 1	.038	5.324*	20.432***
Follow-up test	English rhyme & head production	.430	45.977***	45.977***
	English initial consonant isolation	.101	12.976***	33.990***
Online test	English initial consonant isolation	.389	38.758***	38.758***
Word Review I	English rhyme & head production	.243	19.624***	19.624***
Word Review II	English rhyme & head production	.469	53.975***	53.975***
	English initial consonant isolation	.100	13.904***	39.649***

Note. NWR 1 = nonword repetition 1; NWR 1 & 2 = mean score of nonword repetition 1 & 2.

*$p < .05$. **$p < .01$. ***$p < .001$.