Project Title: Dialogical reading programme on Chinese language and emergent literacy learning in ethnic minority preschool children in Hong Kong

Grantee: The Hong Kong Institute of Education
(The Education University of Hong Kong)

Principal Investigator: ZHOU Yanling
Department of Early Childhood Education
The Education University of Hong Kong

Co-investigators: CHUNG Kien-hoa, Kevin
Department of Early Childhood Education
The Education University of Hong Kong

Catherine MCBRIDE
Department of Psychology
The Chinese University of Hong Kong

Final Report

by

Principal Investigator

1. Abstract

This study examined the effectiveness of the dialogic reading (DR) training for non-native Chinese ethnic minority kindergarten children to learn Chinese. Besides, we also investigated how a DR program with a component of multi-sensory learning (MS) of Chinese character differed from the DR program with morphological awareness (MA) training and DR program solely in terms of Chinese literacy skills acquisition. A total of 167 three-to-five years old ethnic minority kindergarteners in Hong Kong local kindergartens were recruited for the study. All of them are non-native Chinese speaking children. They were randomly divided into four groups: the DR+MS (42), DR+MA (44), DR (41), and control group (40). Only the DR+MS, DR+MA and DR group received a twelve weeks intervention program. Before and after a twelve weeks training program, all participating children were tested on Chinese word reading, writing and vocabulary skills, phonological processing skills, morphological awareness skills, orthographic skills, narrative skills, hand-eye coordination skills, and copying skills. The DR+MS group showed the strongest growth in Chinese word reading and character writing. The DR+MS and DR+MA group also showed a significant improvement in

Chinese orthographic discrimination than the control group. We further conducted a follow-up assessment with three experimental group after three months. The results showed that there was a significant interaction effect between time and group in Chinese word reading, character writing and rapid number naming. These results may have an impact on pedagogical design for Chinese learning among L2 kindergarten ethnic minority children in Hong Kong.

2. Keywords

Chinese learning, dialogic reading, ethnic minority children, intervention program, language acquisition, morphological awareness training, multisensory learning

3. Introduction

Cantonese, Putonghua and English are commonly used in school and community in Hong Kong. The majority of ethnic minority children used English and their heritage language at home (e.g. Ku, Chan, & Sandhu, 2005). Therefore, they may start to learn Cantonese and develop related literacy skills from the kindergarten since they start formal instruction in learning to read and to write at this stage. However, most of the local Hong Kong kindergartens used Cantonese as a medium of instruction. Children with ethnic minority backgrounds may have great difficulties in learning Cantonese, which may make

them lag behind other native Chinese-speaking children. Consequently, it may further hinder the literacy development and academic performance of ethnic minority children. There is an urgent need for developing a specialized teaching strategies and curriculum to help non-native Chinese speaking (L2) children to learn Chinese and adapt the local education system. The current study examined the effectiveness of dialogic reading (DR) with an additional multi-sensory training by comparing with other two groups, a DR group with verbal games tapping on morphological awareness skills and a control group with the DR program solely.

4. Review of literature of the project

4.1 Dialogic reading

Dialogic reading is a shared reading process which was developed for one-on-one reading time between an adult and a child (Whitehurst et al., 1994). Through the shared reading process, the adult can scaffold the child's language and early literacy skills through prompted dialogues about the book. This method places great emphasis on how adult can best encourage and respond to the child's participation through joint attention conversations, language feedback, and adult-child discourse within the context of shared book reading (Farrant & Zubrick, 2011; Zevenbergen & Whitehurst, 2003). Open-ended question prompt is usually used in dialogic training program and children are encouraged to repeat, recall, expand and ask questions in relation to either the storybook or the conversations they have with the adult. Previous research showed that the dialogic reading program helped improve

the vocabulary skills and early related literacy skills in second language learners (Chow, McBride, & Cheung, 2010; Matera, Armas, & Lavandenz, 2016; Tysbina & Eriks-Brophy, 2010). Nevertheless, most of the dialogic reading intervention studies focused on home setting (e.g. Hindman, Connor, Jewkes, & Morrison, 2008; Lonigan, Anthony, Bloomfield, Dyer, & Samwel, 1999; Rahn, Cooggle & Stories, 2016), similar advantages of the dialogic reading program at school setting are hypothesized for the present study.

4.2 Total physical response

With a relatively low proficiency in Chinese for the L2 children, they may have difficulties in understanding the basic instruction and daily vocabulary. Total physical response (TPR) facilitates non-native learners to acquire new vocabularies, phrases and sentience structures in foreign language via engaging learner's whole-body participation (Asher, 1969). TPR is regarded as an effective way to teach non-native speakers foreign language because TPR makes use of learning principal that listening skill usually comes before other language skills. This strategy employs play-oriented activities, and learners will focus on listening and comprehending the target language items prior the verbal production (Asher, 1972). There is a clear evidence showing that TPR had a significant effect for preschool children learning foreign language vocabulary learning for both adults and children (e.g. Asher & Prince, 1967; Mavilidi, Okely, Chandler, Cliff, & Paas, 2015; Toumpaniari, Loyens, Mavilidi, & Paas, 2015).

4.3 Chinese learning and related literacy skills

As previous said, many ethnic minority children often use English and their heritage languages in daily life. However, the Chinese language has its own special features such that learning Chinese differs from learning other kinds of language. Unlike alphabetic language system, there is a lack of phonetic system to assist children to learn Chinese since only around 20% of the Chinese compound character can be identified by the phonological cues provided by the phonetic radicals. However, it does not mean that the phonetic cue is not crucial for Chinese character recognition. Ho and Bryant (1997) revealed the fact that Hong Kong primary school children rely on the phonetic aids to learn new characters. It implied that children are required to acquire a large number of Chinese characters before they are able to make use of the phonetic cues to recognize Chinese characters. Previous research reported that kindergarteners in Hong Kong could able to recognize 204 Chinese word in average before entering primary school. Therefore, the major literacy task for both L1 and L2 children is acquiring as many Chinese characters as they can. If not, their language development may lag behind their peers which hinders their academic development in primary school.

Chinese morphology mostly emphasizes on lexical compounds and homophones. Several studies have revealed that the morphological awareness training can strengthen children's vocabulary development and Chinese word recognition performance (Wang & McBride, 2017; Zhou, Mc-Bride Chang, Fong, Wong, & Cheung, 2012). Chinese language

has a high prevalence of homophones (Li, Anderson, Nagy, & Zhang, 2002). One spoken syllable may map to different characters and have different meanings. Consequently, it is essential for children to understand the Chinese word formation. Understanding the Chinese word formation is beneficial for children to acquire new Chinese vocabulary. Liu and colleagues (2013) proposed that children can make use of their knowledge about morphemes at least in two aspects: First, they can make use of their existing understanding on the rule of morpheme structure to encounter the novel words. Second, they can decode and infer the meaning of the new words based on the acquired words.

A particular important element in learning Chinese is orthographic sensitivity. Previous research found that sensitivity to Chinese radicals is a strong predictor for Chinese reading and writing performance for L2 children (e.g. Hong, Wu, Chen, Chang, & Chang, 2016; Shen & Ke, 2007; Tong & Yip, 2015). Not just only to the radicals, general Chinese orthographic skills are also related to children's reading performance (Shu & Anderson, 1997; Wang et al., 2015). Zhou and McBride (2015) revealed that visual skills of L1 and L2 children were similar, however, L1 Chinese speaking children showed better performance in Chinese recognition ability and orthographic memory. These pieces of evidence showed that being sensitive to the patterns and structures of Chinese characters are the advantages in learning Chinese. Hoosain (1991) reported that there are approximately 200 semantic and 800 phonetic radicals in Chinese language. There are certain degrees of consistency in forming

the character, but, experiencing individual character is the means for preschoolers to learn Chinese character in Hong Kong.

4.4 Copying skills in Chinese learning

Pure copying involves copying the geometric figures and unfamiliar prints in other languages. Rather than the Chinese orthographic knowledge, pure copying emphasizes on visual-motor integration skills which has a close association to visual-spatial and Chinese writing skills (Barnhardt, Borsting, Deland, Pham, & Vu, 2005; Daly, Kelley, & Krauss, 2003). A recent study also demonstrates that pure copying has a strong positive association with Chinese character writing for young Chinese learners. Not just only writing performance, there are also evidences showing the association between pure copying and Chinese character recognition ability (Wang, McBride-Chang, & Chan, 2014). The study of Chinese dyslexia children from McBride-Chang and her colleagues (2011) demonstrated the association between copying skills and Chinese word reading performance. On top of this, copying skills were an indicator to differentiate children with or without dyslexia after controlling the phonological and morphological awareness. Kalindi et al. (2015) found the unique predictive power of copying skills in Chinese dyslexia children only, but not in those with English reading difficulties. Pure copying involves visual planning, visual-spatial and fine motor skills. So, it is superior to hand-writing and hand-eye coordination skills. We can be viewed the activity of copying geometric figures and unfamiliar prints as a kind of multi-

sensory learning.

5. Theoretical and conceptual framework of the project

5.1 Early intervention studies for young learners of Chinese

Recent studies have investiged the effectiveness of copying skills intervention on Chinese language learning (Lam & McBride-Chang, 2013; Lin et al., 2009; Wang & McBride, 2017). Lin and her colleagues (2009) suggested that parents should take a role in facilitating children's Chinese language learning by encouraging writing activities at home. Lam and McBride-Chang (2013) found that writing practice focusing on radical knowledge was the most effective means to improve the writing skills for young children. The parents can scaffold their children to perform cahracter copying or other writing activities. Wang and McBride (2017) performed an intervention studies in Chinese kindergartens and provided an evidence showing that pinyin training along with copying practice brougt additional benefit on Chinese early literacy skills than pinyin training solely. The effectiveness of copying practise also applied to L2 children. A recent experiemental research has investigated the copying practices for L2 young children (Wang, McBride, Zhou, Joshi, & Farver, 2017). The results provided an evidenct that copying practice is also beneficial for L2 young children to learn Chinese. Therefore, copying characters become a common practice for children to master Chinese characters at school. However, it is worth noting that children may feel bored with character copying tasks and this will hinder the motivation for them to learn Chinese

characters.

As previous mentioned, most of the dialogic intervention studies take place in home setting in which both parent and child are native Chinese speakers. However, little support from parents is received for L2 children because their parents may also not master the Chinese language well. Hence, children who have Chinese as an L2 can only have limited exposure to Chinese before normal schooling. They often have low proficiency in Chinese language and limited vocabularies for practicing the metalinguistic skills. So, it is a challenge for them to acquire Chinese characters with L1 peers in the kindergarten because they will need to develop both verbal language and early literacy skills concurrently. Previous intervention studies for L2 speakers to learn Chinese focused on adult or older children, and the results suggested that orthography learning and pinyin learning can facilitate the vocabulary development and Chinese language learning for the participants (Lü, 2017; Nguyen, Zhang, Li, Wu, & Cheng, 2017; Taft & Chung, 1999; Tong & Yip, 2015). Nevertheless, unlike experienced learners, L2 kindergarteners lack of opportunity to expose to Chinese. Therefore, dialogic reading training may be an effective way for training L2 kindergarteners' Chinese vocabulary knowledge and general oral language skills as well as emergent literacy skills.

Dialogic reading program can benefit foreign language learning, but research found that dialogic reading paired with morphological training brought outstanding effect in children's

vocabulary knowledge and early literacy skills when compared with using dialogic reading along (Chow, McBride-Chang, Cheung, & Chow, 2008). Zhou and her colleagues (2012) reported the emergent literacy acquistion of Hong Kong kindergarteners was correlated their Chinese morphological awareness skills, especially Chinese lexical compounding and homophone awareness skills. These results demonstrated the effect of dialogic reading together with morphological awareness training are beneficial for Chinese language learning for young children. Beisdes, children can also enjoy the shared reading process by having interactions and communication with adult.

5.2 Multi-sensory learning for early literacy skills

Multi-sensory learning is the learning strategy that engaging children's two or more sense including visual, auditory, kinaestheic, and touching during the learning process. Goswami (2008) suggested that human's learning is multi-sensory based on our neurological makeup. Therefore, multi-sensory methods is an effective means for children to acqurie early literacy skills (Neumann, Hyde, Neumann, Hood, & Ford, 2012). Using multiple senses for learning definitely can strengthen our memory since the various brain regions will leave a mark for the learning activities. However, Chinese characters are the symbols used for human communication and the characters are frequently appeared in 2D images on th paper. In this case, reading and copying the characters are just using their visual sense rather than a intergration of multiple senses. However, there were empiriacal evidences to support that

multi-sensory activities especially for viasual and haptic senses played a role in acquiring reading and writing for early stage of language learning (Bara, Gentaz, Colé, & Sprenger-Charoll, 2004; Bara & Gentaz, 2011; Labat, Ecalle, Baldy, & Magnan, 2014; Neumann, 2014).

Previous researches investigated the effectiveness of visual training and visual-haptic training on letter recognition in French kindergarten children (Bara et al., 2004; Bara & Gentaz, 2011). Results demonstrated that young children with both visual and haptic training brought additional advantages on letter recognition when compared with visual training solely. Furthermore, Labat and colleagues (2014) have run a similar experiment with French kindergarten children by comparing the effect of visual-auditory-haptic and visual-auditory exposure. It is unsurprised that children in visual-auditory-haptic group had better performance in letter recognition and writing task. These evidences supported the view that the knowledge will be more memorable when more senses are engaged in the learning process. However, most of the evidences for multi-sensory learning strengthens early literacy skills are conducted in alphabetic languages. Little attention has been given to how multi-sensory learning affects children's early literacy acquisition. Therefore, the current study aimed to examine the effectiveness of multi-sensory learning in early Chinese literacy learning.

5.3 The present study

Our study aims to compare the both immediate and long-term impact on Chinese literacy learning for L2 young children in Hong Kong by comparing three teaching strategies: dialogic reading (DR) approach, dialogic reading coupled with morphological awareness training (DR+MA), and dialogic reading coupled with multi-sensory learning (DR+MS). We adopted a dialogic reading program involving total physical response technique to facilitate children's Chinese vocabulary growth. Total physical response is a strategy that help foreign language learners to acquire new vocabularies (Asher, 1965). All three groups received the same training materials and instructions for the dialogic reading session as well as total physical response training. On top of the dialogic reading training, the two experimental groups received additional training materials to let children in DR+MS condition have multi-sensory learning exposures, and children in DR+MA condition have games for morphological awareness training. Besides, a control group without the implementation of the dialogic reading was also involved in our study. Therefore, the results of our study could determine effective teaching and intervention strategies which could be adopted for kindergarten teachers to promote the Chinese learning for L2 children.

6. Methodology

6.1 Participants

At the pre-test stage, we have recruited 127 first and second year non-Chinese native speaking (L2) children from 8 kindergartens of a non-profit kindergarten organization in

Hong Kong. The 8 kindergartens share the same curriculum and all of the participants are come from the Chinese stream which used Cantonese as a medium of instruction in the class. We combined two grades together for the training due to the limited numbers of participants from K1 and K2. Seventeen different mother tongues were spoken by participants, including English, Urdu, Nepali, Punjabi, Tagalog, Hindi, Korean, Chinese, Bahasa Indonesia, Bengali, Dutch, Esan, Ewe, French, Kinyarwanda, Russian, and Tamil. It is interesting that 3 parents reported their children are native Chinese speakers since one of the parents for each of the three families spoke Chinese as L1. However, their school teachers confirmed that although these three children showed better Chinese language proficiency than other L2 peers, the teachers regarded them as L2 speakers when comparing them with other L1 children in the same class. The participants were randomly divided into three groups based on the location of the kindergarten. The three groups included dialogic reading paired with multi-sensory training (DR+MS) group; dialogic reading paired with morphological training (DR+MA) group; and the dialogic reading (DR) group.

The attrition rate was 7% and 17% for the post-test examined immediate effect (T2) and long-term effect (T3) respectively. In total, 9 children were dropped out in T2 including 3 children from the DR+MS group, 4 from the DR+MA group, and 2 from the DR group. Therefore, there were 118 participants included in the data collection and analysis in T2. In T3, 12 more participants were dropped out from the study, including 5 children from the

DR+MS group, 4 children from the DR+MA group, and 3 children for the DR group. To sum up, there were 106 participants completed all the stages of our study including 34 children from the DR+MS group, and 36 children from the DR+MA group and the control group, respectively.

A year later, we have recruited 40 non-native Chinese kindergarteners from the same non-profit kindergarten organization. They are also come from the Chinese stream which used Cantonese as a medium of instruction in the class. Ten different mother tongues were spoken by our participants, including English, Urdu, Nepali, Tagalog, Bahasa Indonesian, Amharic, Bengali, Chinese, Luganda and Spanish. There was a participant did not specify the mother tongue and reported Chinese as L1 in the questionnaire respectively. Nevertheless, we have confirmed with the teachers that both children were treated as L2 Chinese speakers when comparing with their L1 peers. All 40 participants were included in the control group and the control group also followed the assessment scheduled with the experimental groups. However, we only assessed them for two-time points and there was no intervention program in between. Therefore, the posttest was commenced three months later after the pretest ended. The attrition rate for the control group was 10% in the posttest as 4 of the participants were quitted in our study.

6.2 Measures

Chinese vocabulary knowledge. The Chinese vocabulary knowledge task was comprised

of both receptive vocabulary and expressive vocabulary items. All items were extracted from the Peabody Picture Vocabulary Test-Third Edition (PPTV-III; Dunn & Dunn, 1997) for the pre-test including 35 receptive vocabulary items and 30 expressive vocabulary items. Four pictures were shown to children for receptive vocabulary item. Children were informed to select one picture which is consistent with the verbal description from the tester. For expressive vocabulary item, children were required to answer the target vocabulary based on the picture shown. A point was given for the correct response and the ceiling rule would be reached for five consecutive incorrect answers for each task. The items in T2 and T3 were slightly different from pretest. For receptive vocabulary section, we have added 5 more questions from the PPTV-III in T2 and T3. However, the administration of the task was completely same as pre-test. On top of the PPTV-III, 12 questions were added for the expressive vocabulary items. The target vocabulary for the added items were derived from the words in storybooks which used in the intervention program and no ceiling rule was applied to the newly added expressive vocabulary items. One point was allotted for each correct response. Therefore, there were 40 items and 42 items for the receptive vocabulary and expressive vocabulary sections respectively in T2 and T3.

Chinese rapid number naming. The rapid number naming test from the Comprehensive Test of Phonological Processing (CTOPP; Wagner, Torgeson, & Rashotte, 1999) was used to assess the rapid number naming speed of the participants. There were 36 digits printed on an

A4 paper and children were informed to use their fastest speed to name all the digits in sequence for two times. The time for each trial was measured in seconds and average time for two trials was computed. However, the performance of children who made more than four mistakes in either one of the trial was recorded as 300 seconds.

Chinese phonological awareness. A syllable deletion task was commonly used in previous studies in evaluating children's phonological awareness (Chow, McBride-Chang, & Burgess, 2005; Zhou et al., 2017) and we also adapted similar measure to measure the phonological awareness performance of our participants. Tester presented a three-syllable word to children and they were informed to delete one of the syllable. For example, /fo2 ce1 zaam6/ (train station) without /fo2/ would become /ce1 zaam6/. The task comprised of 15 real word and 5 non-word conditions. There were 5 questions required participants to delete the initial, middle, and last syllable respectively. In contrast, only the first syllable was required to drop out for the non-word conditions after considering the Chinese language proficiency of the participants. One point was awarded to each correct response while the task would be ended for children incorrectly answer more than six items within the same level of items.

Chinese dictation task. Children were required to write down the word that verbally presented by the tester in this task. There were four two-character words in the pretest. All the chosen words were the words that commonly used in the preschool Chinese storybooks in Hong Kong. Children could receive one point when they correctly put down the word.

However, we noticed that only a child was able to get one point for this task in the pretest. So, we have adjusted the items and scoring procedure for the T2 and T3 assessment to prevent the floor effect. There were eight single character words and two two-character words in T2 and T3 assessment such that there were 12 characters in total. Six of the characters were sane as the pretest while the remaining characters were come from the storybooks used in the intervention program. A visual aid was presented to children when they did not write down the anything after hearing the target character. Two points and one point were given to children who could write down the correct character after hearing the verbal presentation from the tester and seeing the picture aid respectively.

Chinese orthographic discrimination. Previous studies used an orthographic component search test to evaluate the orthographic processing skills of children (Ho, 2006; Siok & Fletcher, 2001). We adapted a similar test used by Ho (2006) to evaluate children's Chinese orthographic discrimination ability. A Chinese radical "□" (mouth) was shown to children to ensure they can recognize the radical. Children were informed to figure out all the characters that contained the target radical "□" on an A4 paper which contained 80 different Chinese characters. One point was given to each correct identification.

Visual-perception skills. Hand-eye coordination task and pure copying task were used to tap on children's visual perception skills. All 5 items of the hand-eye coordination task were chosen from the Developmental Test of Visual Perception-Third Edition (DTVP-3; Hammill,

Pearson, & Voress, 2014). Children were informed to draw a continuous line from the starting point to the destination, and the line should be stayed within the grey path. The path was further divided into several one-inch segments. For the first two items, one point was allotted when the continuous line did not draw beyond the grey path within each segment. However, additional lines were added to the path and the grey area became narrower starting from the third items. Four points were awarded to children who could draw the continuous line within the grey path within each segment. Children were received one to three points when the unbroken line fell into the unshaded segment according to the deviation from the grey path. No point was rewarded for the broken line or line drew beyond the path for each segment.

The pure copying section consisted of geometric figure copying and unfamiliar print copying. All 18 figure copying items were selected from the Copying sub-test in the DTVP-3 (Hammil et al., 2014). Children were informed to copy the figure into the designated box and their works scored with a 4-point scale: three points were allotted for the work which was drawn approximately as the same as the stimulus; two points were awarded for the work which was similar to the stimulus, but with minor problems (e.g. the drawing touch the boundaries of the box provided); one point was rewarded to children whose drawing showed the basic idea of copying the stimulus, but with a number of inadequacies; however, no point was given to missing item and the drawing which was different from the stimulus. This part would be terminated when children received three consecutive zero points. Besides the figure

copying, children were also informed to copy the unfamiliar prints. The unfamiliar prints were extracted from the previous study (McBride-Chang, Chung, & Tong, 2011) and comprised of 3 Hebrew prints and Korean prints respectively. Children were required to complete all the items in this section. For Hebrew copying, the words were composed of either three or four segments. One point was awarded to each correct segment copying. When children correctly copying all the segments, they had chances to receive one additional point for proper sequence and proper alignment of the word respectively. For Korean print copying, the word in each item consisted of three or four components. Two separate three-point scales were used to code children's coping of each component in terms of shape and position.

Non-verbal intelligence. The Raven's Coloured Progressed Matrices (Raven, Raven, & Court, 1995) was used to evaluate the non-verbal intelligence of children. A figure with a missing part was presented to children and they were asked to find out the missing part to make the figure completed. There were six choices for each item and children were needed to choose one answer to complete the figure. This task consisted of 36 items and one point was given to each correct response.

Chinese homophone awareness. After considering the Chinese oral language proficiency of L2 young children, we adopted the task from Tong and colleagues (2011) which focused listening comprehension to assess children's morphological awareness skills. The task consisted of 14 items, and four pictures with corresponding verbal presentation were

delivered to children in each item (e.g. wooden box 木箱; hurt 受傷; shop 商店; department store 百貨公司). The child heard a targeted morpheme item with an accompanying picture and was asked to select which one of the four morpheme items resembled the morpheme in the targeted items. (e.g. shopping mall 商場). "Shop 商店" is the correct answer for our example. One point was given to each correct response in this task.

Chinese word recognition. All the items in this task were selected from the children's storybooks or textbooks in Hong Kong. There were eighty-eight items including thirty-one single character words and fifty-seven two-character words. The items were sorted in ascending level and children were informed to read all the words in sequence. The task would be terminated when children mispronounce ten words consecutively.

Chinese narrative skills. The Edmonton Narrative Norms Instrument (ENNI; Schneider, Dubé, & Hayward, 2005) was used to tap on the Chinese narrative skills of children. Picture Book A1 from the ENNI was adapted after considering the Chinese oral language ability of the L2 kindergarteners. Picture Book A1 consisted of five pages, and there was a black-and-white picture on each page. The story was relatively simpler in Picture Book A1 as it only contained two characters in the story. Children were informed to tell the story in Chinese. All the pages were shown to children prior they started to tell the story. Their response was videotaped for coding purpose. We evaluated children's performance based on the story grammer scoring sheet from the test developer (Schneider et al., 2005). One or two points

were given to children whose story provided a description to the following elements: characters, setting, initiating event, internal response or plan of characters, purpose of character's action, or characters' feelings.

6.3 Intervention

Children in the intervention groups were randomly divided into three conditions based on the location of the kindergarten. We further divided each condition into ten subgroups such that each subgroup involved three to five children. There were four subgroups and six subgroups in the morning class and afternoon class for every condition respectively. In general, more than 96% of our participants attended eight or more intervention sessions. The intervention program consisted of twelve lessons and the program was delivered once a week over three months. Each session lasted for around forty-five minutes and all the sessions were conducted in children's own kindergarten within the school time.

Three trained university educated kindergarten teacher were employed to deliver the intervention program under the guidance and supervisions from the principal investigator. The teachers were trained on how to perform dialogic reading (Whitehurst et al., 1994) and use of TPR strategy (Asher, 1965) to teach L2 children. Individual trainings on multi-sensory learning and morphological awareness skills were given to the teacher who was responsible for the DR+MS and DR+MA condition respectively. All the lessons were videotaped, and lesson observations were made by the principal investigator periodically. On top of these,

weekly meetings were conducted to monitor the teaching and children's learning progress to ensure standardize teaching across three groups as well as the effectiveness of the intervention session. All three groups received the same training materials and teaching preparation for the dialogic reading training content. Twelve Chinese storybooks were selected on the basis that we judged the storybooks were culture-free and appropriate for children aged from three to four. The storybooks list for the intervention program was shown in Appendix A.

Every lesson consisted of three sections: Part 1 involved a TPR game to preview the target vocabularies for the storybook. This part lased for approximately 10 minutes and five novel vocabularies were introduced in each lesson. TPR required children's complete immersion in listening comprehension without speaking. Children were encouraged to perform the action associated with the vocabulary, or even speak the phrases after they completely comprehended the meaning of the vocabulary (Asher, 1969).

Part 2 involved the storytelling section which employed the dialogic reading strategy. The dialogic reading section lasted for approximately 20 minutes. We used the PEER principal (i.e. prompt, evaluation, expansion, and repetition) to deliver the story to children. Teacher encouraged children to express their ideas to the story, the teacher then evaluated children's opinions and expanding the response into a complete sentence with more complex sentence structure. Children were invited to repeat the sentence presented by the teacher

eventually. Five prompts including completion prompts, recall prompts, open-ended prompts, Wh-prompts, and distancing prompts were used to help children to recall the vocabularies learnt from the TPR section. Previous research showed that this repetition strategy can help L2 speakers to acquire the vocabularies in foreign language (Lever & Sénéchal, 2011).

The contents of Part 3 were varied across three conditions. This section lasted for 10 to 15 minutes. For the DR group, the teacher extended the conversations about the story with children because no additional training was provided. For the DR+MS condition, children had opportunities to experience the target character(s) in a multi-sensory way. Simple structured and high frequency characters were chosen from the storybooks as the target characters. The teacher showed the evolvement of the shapes of the character in history at the beginning of this section. Prior to copy the character on paper, children were given a model of character which made up by either paper cut, clay, or 3D printing to feel the structure of the character. For the DR+MA condition, a compound word (e.g. apple tree /蘋果樹/) from the storybook was selected the principal investigator in each lesson. The teacher explained to children that the tree with apples was called apple tree /蘋果樹/. One more example was presented and explained to children, such as the tree with watermelon was called watermelon tree /西瓜樹/. After the illustrations, children were informed to form a compound word based on the meaning provided (e.g. How do we call a tree with banana?). The answer for the example was banana tree /香蕉樹/.

The Leuven Involvement and Well-Being Scales (Laevers, 2005) was used to measure children's well-being and involvement level in the class from lesson 9 to lesson 12. The results showed that there was no group difference between children's well-being and involvement across the group. Therefore, the learning environment in the intervention program of children was assumed to be equivalent across three groups.

7. Data collection and analysis

Upon the signed informed consent form was received from the parents, a questionnaire surveying their family socioeconomic status (e.g. parental educational levels and family income) and home language/literacy environment. Besides, children were assessed on a range of Chinese language- and early literacy-related tasks. Except the non-verbal intelligence task, all the measures were assessed at all three-time point. The pretest data collection started from mid-February to mid-March in 2017 and 2018 for the experimental groups and control group respectively. After the pretest data collection was completed, three experimental groups started the intervention program immediately. The intervention program lasted for 12 weeks from mid-March to mid-June. The posttest was started right after the completion of the intervention program and lasted until mid-July to compare the effectiveness of the intervention program. The third data collection was administered in between mid-September to mid-October for viewing the effectiveness of the intervention program in long-term.

Data screening, like outlier identification and treating missing data, was conducted after

data processing. Before comparing the group improvement, analyses were conducted to ensure the demographic information and pre-test performances were statistically similar across groups.

One-way analysis of variance (ANOVA) was used to compare the demographic information between groups, and the results suggested that both mothers' educational levels were similar across group. In contrast, there was a statistical difference in fathers' educational level. The post hoc comparison showed that fathers' educational level in the control group was significantly better than the DR+MA and DR group. Moreover, the non-verbal intelligence was similar across group, but there was significant age difference among groups. The post comparison revealed that children in the DR+MS group was relatively older than children in the DR group and children in the control group was younger than all three experimental groups. The possible age difference between the groups may accounted by the unequal distribution of K1 and K2 children. All the demographic information for the participants is displayed on Appendix B. We further compared the frequency of children using Chinese for daily activities at home and children's home literacy environment. Insignificant group effect was found on children's use of Chinese for daily activities at home and home literacy activities.

We also conducted ANCOVAs to compare the pretest scores on Chinese vocabulary knowledge, rapid number naming, phonological awareness, character dictation, orthographic

discrimination, word recognition, narrative skills, and visual skills after controlling grade and non-verbal intelligence. Results suggested children's performances on all measures were similar across three groups, except for Chinese narrative skills. The pairwise comparisons revealed that children in the DR+MA group had better performance on narrative skills than children in the DR+MS group. To sum up, we statistically controlled grade, non-verbal intelligence and pre-test scores for the subsequent analyses in comparing the group improvement to minimize the influence of the confounding.

Appendix C shows the means and standard deviations for the measures among all three time points. It was noted that the homophone awareness task was excluded in our study because our analysis showed an unacceptable reliability of the task, and the group improvement as well as the pretest performance were statistically similar among three groups. It might be too challenging for L2 kindergarten children to hold four items in mind prior making a choice of a targeted morpheme, although they were not required to speak Chinese words. Consequently, we did not include this task in subsequent analyses.

8. Results and discussions

8.1 Immediate effect

After putting grade, non-verbal intelligence and pre-test score as covariates, nine separate analysis of covariances (ANCOVAs) were conducted for Chinese vocabulary knowledge, rapid number naming speed, phonological awareness, character dictation,

orthographic discrimination, word recognition, narrative skills, hand-eye coordination and pure copying. A significant group effect was found in Chinese word reading, dictation task and Chinese orthographic discrimination. Specifically, the pairwise comparison of estimated marginal means showed that the improvement for the DR+MS group on Chinese word reading was significant larger than all the groups while no other significant difference was found. The results suggested that children in the DR+MS group enjoyed the improvement for Chinese character recognition most. It is noted that the estimated marginal means for DR+MA and DR group on both Chinese word reading and dictation were also higher than the control group, regardless of the insignificant result.

Similarly, a significant group improvement difference was also found for the Chinese character dictation task. Therefore, the pairwise comparison of estimated marginal means was conducted and the results revealed that the improvement for the DR+MS group was significantly larger than the improvement for the remaining groups. Although the performance for DR+MA and DR group in character dictation was better than the control group, the results were insignificant. Hence, it implied that the DR+MS training was the most effective way to improve L2 children's Chinese character writing ability when compared with other two experimental conditions. Nevertheless, it is worth noting that there was only one child from the DR group could receive a point by writing "□" (mouth) correctly in the pretest.

We also conducted the pairwise comparison for the Chinese orthographic discrimination task to examine the group improvement difference since there was a significant group effect. The results showed that the DR+MA and DR+MS group had larger improvement than the control group. Despite the insignificant difference between the DR group and control group, the DR group had higher estimated marginal mean than the control group. So, the results suggested that the dialogic reading was beneficial to L2 children Chinese orthographic discrimination ability development.

Regardless the insignificant difference for Chinese vocabulary knowledge task, the estimated marginal mean suggested that the DR+MA group had better performance in Chinese vocabulary knowledge task when compared with the DR+MS group, DR group and control group. This result was consistent with previous finding that morphological awareness training can facilitate young children's vocabulary development. Meanwhile, the estimated marginal means also suggested that the DR+MS group had the fastest speed on rapid digit naming task and better performance in Chinese phonological awareness, despite the insignificant differences for these tasks. Furthermore, all three experimental groups demonstrated better narrative skills than the control group, although the ANCOVA result was insignificant.

8.2. Long-term effect

Our study also investigated long-term effect of the intervention program. We have

conducted nine separate mixed ANCOVAs to examine the children's growth in Chinese related literacy skills among the three experimental groups. After statistically controlling grade and non-verbal intelligence, the interaction effect between time and group was found in Chinese word reading, rapid number naming and character dictation task.

Appendix D is the graphical representation for the estimated marginal means of Chinese word reading scores for three experimental groups under all three time points. In general, the DR+MS group had the largest improvement among the experimental groups and a sharply improvement was found in T2, although there was a slightly decrease in T3. Similar pattern was also found in the DR group, but the growth rate was relatively less salient. In contrast, the DR+MA group showed a relatively steady improvement across T2 and T3. The results suggested that the school can practice the dialogic reading with multisensory learning continuously to optimize L2 children character reading when compared with other two training methods.

For the character writing, all the experimental groups showed the continuous improvement in T2 and T3. Nevertheless, it is noting that the growth for the DR+MS in T2 was much more than other two groups. Although the growth rates were similar across the groups in T3, the rapid improvement immediately after the intervention program made the DR+MS group children enjoyed writing more characters in T3. This evidence revealed that the DR+MS training was beneficial for L2 children to learn to write Chinese character in both

immediately and long-term when compared with the DR+MA training and DR training solely. The estimated marginal means for the dictation task were displayed in Appendix E.

It is interesting that children in the DR+MS group improve most in Chinese rapid number naming in T2, but, their performance was slightly decline in T3. Similar pattern was found for the DR group, regardless of the less enhancement in T2. The results showed that the DR+MA group had improvement consecutively. In T2, the DR+MA group did not improve as much as the DR+MS group. However, the DR+MA group improve noticeably in T3, even children in the DR+MA group performance in Chinese rapid number naming was better than children in the DR+MS group in T3. It reflected that the DR+MS training brought largest immediate improvement and the DR+MA training had relatively long-lasting effect in phonological processing speed among groups. The graph plotted the estimated marginal means was shown in Appendix F.

Appendix A

Storybooks List for the Twelve-week Intervention Program

Week	Author	Title	Publisher
1	Herve Tullet *Translated by*: 蒲蒲蘭	Press Here 點點點	21st Century Publishing Group 二十一世纪出版社
2	Eric Carle *Translated by*: 林良	From Head to Toe 從頭動到腳	Tomorrow Publishing House 明天出版社
3	Yoshio Nakae *Translated by*: 趙靜、文紀子	Little Mouse Wants an Apple 想吃蘋果的鼠小弟	Nan Hai Publishing Co. 南海出版公司
4	Jon Klassen *Translated by*: 楊玲玲、彭懿	I Want My Hat Back 我要把我的帽子找回来	Tomorrow Publishing House 明天出版社
5	Mao Xiao (蕭袤)	The Frog and the Boy 青蛙與男孩	Petrel Publishing House 海燕出版社
6	Graeme Base *Translated by*: 影子	Eye to Eye 眼靈靈 心靈靈	Changjiang Children's Publishing Group 長江少年兒童出版社
7	Bill Martin Jr. and Eric Carle *Translated by*: 李坤珊	Brown Bear, Brown Bear, What Do You See? 棕色的熊，棕色的熊，你在看什麼？	Tomorrow Publishing House 明天出版社
8	Camilla Reid *Translated by*: 黃筱茵	Lulu Loves Colours 露露愛顏色	Sun Color Culture Publishing Co., Ltd. 三采文化有限公司
9	Eric Carle *Translated by*: 鄭明進	The Very Hungry Caterpillar 好餓的毛毛蟲	Tomorrow Publishing House 明天出版社
10	Tatsuya Miyanishi *Translated by*: 彭懿	Harapeko Hebikun 好餓的小蛇	21st Century Publishing Group 二十一世纪出版社
11	Camilla Reid *Translated by*: 黃筱茵	Lulu Loves Shapes 露露愛形狀	Sun Color Culture Publishing Co., Ltd. 三采文化有限公司

Appendix A

Storybooks List for the Twelve-week Intervention Program

Week	Author	Title	Publisher
12	Camilla Reid *Translated by*: 黃筱茵	Lulu Loves Noices 露露愛聲音	Sun Color Culture Publishing Co., Ltd. 三采文化有限公司

Appendix B

Demographic Information for the Participants

Characteristic	DR+MS group (%) (*n* = 42)	DR+MA group (%) (*n* = 44)	DR group (%) (*n* = 41)	Control group (%) (*n* = 40)
	n (%)	*n* (%)	*n* (%)	*n* (%)
Grade				
K1	11 (26.20)	27 (61.40)	26 (63.40)	22 (55.00)
K2	31 (73.80)	17 (38.60)	15 (36.60)	18 (45.00)
Gender				
Boy	24 (57.10)	13 (29.50)	20 (48.80)	21 (52.50)
Girl	18 (42.90)	31 (70.50)	21 (51.20)	19 (47.50)
Fathers' educational level				
Primary 3 or under	2 (4.80)	2 (4.50)	0 (0.00)	1 (2.50)
Primary 4 to 6	6 (14.30)	5 (11.40)	9 (22.00)	1 (2.50)
Secondary 1 to 5	10 (23.80)	14 (31.80)	11 (26.80)	4 (10.00)
Secondary 6 to 7	1 (2.40)	5 (11.40)	5 (12.20)	4 (10.00)

Appendix B

Demographic Information for the Participants

Characteristic	DR+MS group (%) (*n* = 42)	DR+MA group (%) (*n* = 44)	DR group (%) (*n* = 41)	Control group (%) (*n* = 40)
	n (%)	*n* (%)	*n* (%)	*n* (%)
Higher diploma	7 (16.70)	4 (9.10)	6 (14.60)	10 (25.00)
Bachelor's degree	4 (9.50)	4 (9.10)	4 (9.80)	10 (25.00)
Master's degree or above	9 (21.40)	4 (9.10)	3 (7.30)	8 (20.00)
Missing	3 (7.10)	6 (13.60)	3 (7.30)	2 (5.00)
Mothers' education level				
Primary 3 or under	3 (7.10)	0 (0.00)	0 (0.00)	1 (2.50)
Primary 4 to 6	3 (7.10)	12 (27.30)	7 (17.10)	0 (0.00)
Secondary 1 to 5	16 (38.10)	13 (29.50)	13 (31.70)	14 (35.00)
Secondary 6 to 7	1 (2.40)	6 (13.60)	5 (12.20)	4 (10.00)
Higher diploma	4 (9.50)	3 (6.80)	4 (9.80)	6 (15.00)
Bachelor's degree	9 (21.40)	9 (20.50)	10 (24.40)	9 (22.50)

Appendix B

Demographic Information for the Participants

Characteristic	DR+MS group (%) (*n* = 42)	DR+MA group (%) (*n* = 44)	DR group (%) (*n* = 41)	Control group (%) (*n* = 40)
	n (%)	*n* (%)	*n* (%)	*n* (%)
Master's degree or above	5 (11.90)	0 (0.00)	1 (2.40)	5 (12.5)
Missing	1 (2.40)	1 (2.30)	1 (2.40)	1 (2.50)

Appendix C

Means and Standard Deviations for the Measures Across All the Time Points

Measure	DR+MS group	DR+MA group	DR group	Control group
	M (*SD*)	*M* (*SD*)	*M* (*SD*)	*M* (*SD*)
Chinese vocabulary knowledge				
Time 1	9.57 (8.70)	8.20 (7.10)	6.56 (6.75)	6.28 (5.66)
Time 2	19.52 (9.44)	18.97 (10.65)	15.24 (12.71)	13.72 (9.33)
Time 3	19.01 (11.63)	18.56 (10.90)	16.03 (11.26)	N/A
Chinese rapid number naming				
Time 1	247.91 (101.19)	270.57 (76.58)	281.83 (65.58)	268.11 (78.00)
Time 2	194.03 (124.70)	255.23 (96.20)	256.56 (97.09)	251.08 (99.45)
Time 3	192.88 (125.97)	209.77 (121.06)	262.51 (91.95)	N/A
Chinese phonological awareness				
Time 1	2.38 (4.43)	1.09 (3.02)	1.10 (2.82)	.38 (1.13)
Time 2	3.96 (5.87)	1.76 (3.75)	1.83 (4.14)	1.67 (2.57)

Appendix C

Means and Standard Deviations for the Measures Across All the Time Points

Measure	DR+MS group	DR+MA group	DR group	Control group
	M (*SD*)	*M* (*SD*)	*M* (*SD*)	*M* (*SD*)
Time 3	5.26 (5.91)	3.78 (5.44)	3.25 (4.51)	N/A
Chinese dictation task				
Time 1	.00 (.00)	.00 (.00)	.02 (.16)	.00 (.00)
Time 2	1.94 (2.85)	.14 (67)	.55 (2.09)	.00 (.00)
Time 3	2.28 (3.32)	.70 (1.93)	.64 (1.82)	N/A
Chinese orthographic discrimination				
Time 1	44.40 (5.88)	42.16 (5.14)	41.48 (4.33)	40.10 (7.84)
Time 2	48.98 (7.21)	44.20 (7.39)	42.80 (8.21)	40.81 (7.77)
Time 3	47.84 (9.06)	45.69 (9.29)	45.55 (8.41)	N/A
Chinese word reading				
Time 1	1.95 (3.99)	.82 (2.42)	2.20 (10.54)	.75 (2.43)

Appendix C

Means and Standard Deviations for the Measures Across All the Time Points

Measure	DR+MS group	DR+MA group	DR group	Control group
	M (*SD*)	*M* (*SD*)	*M* (*SD*)	*M* (*SD*)
Time 2	5.93 (7.96)	1.30 (3.39)	3.12 (11.61)	1.28 (3.57)
Time 3	6.00 (8.91)	2.58 (5.76)	2.68 (10.34)	N/A
Chinese narrative skills				
Time 1	.00 (.00)	.32 (.98)	.10 (.49)	.10 (.50)
Time 2	1.00 (1.84)	.66 (1.55)	.67 (1.77)	.25 (.53)
Time 3	.69 (1.23)	.59 (1.16)	.52 (.92)	N/A
Hand-eye coordination				
Time 1	96.26 (37.11)	94.36 (38.31)	98.54 (29.36)	100.58 (37.65)
Time 2	120.18 (26.65)	114.83 (26.60)	115.26 (26.52)	115.17 (24.52)
Time 3	130.73 (26.74)	128.11 (17.57)	132.11 (22.34)	N/A
Pure copying				

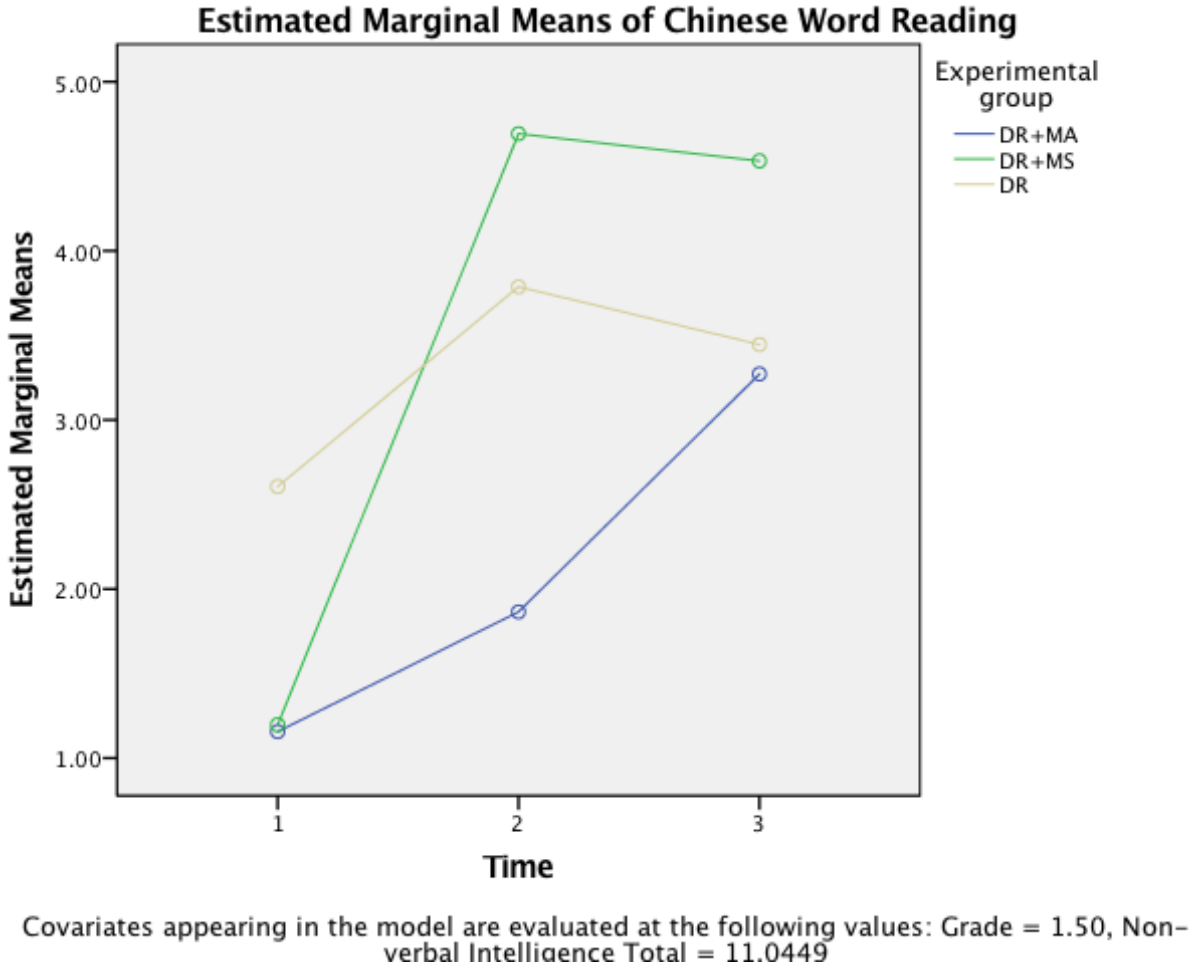
Appendix C

Means and Standard Deviations for the Measures Across All the Time Points

Measure	DR+MS group	DR+MA group	DR group	Control group
	M (*SD*)	*M* (*SD*)	*M* (*SD*)	*M* (*SD*)
Time 1	30.69 (23.63)	20.86 (21.75)	23.46 (24.92)	28.95 (25.68)
Time 2	41.41 (23.32)	27.75 (22.32)	30.08 (23.33)	36.47 (25.17)
Time 3	40.61 (20.51)	37.32 (19.06)	35.57 (19.31)	N/A
Non-verbal intelligence				
Time 1	12.07 (4.37)	10.48 (3.85)	10.60 (4.73)	9.95 (4.30)

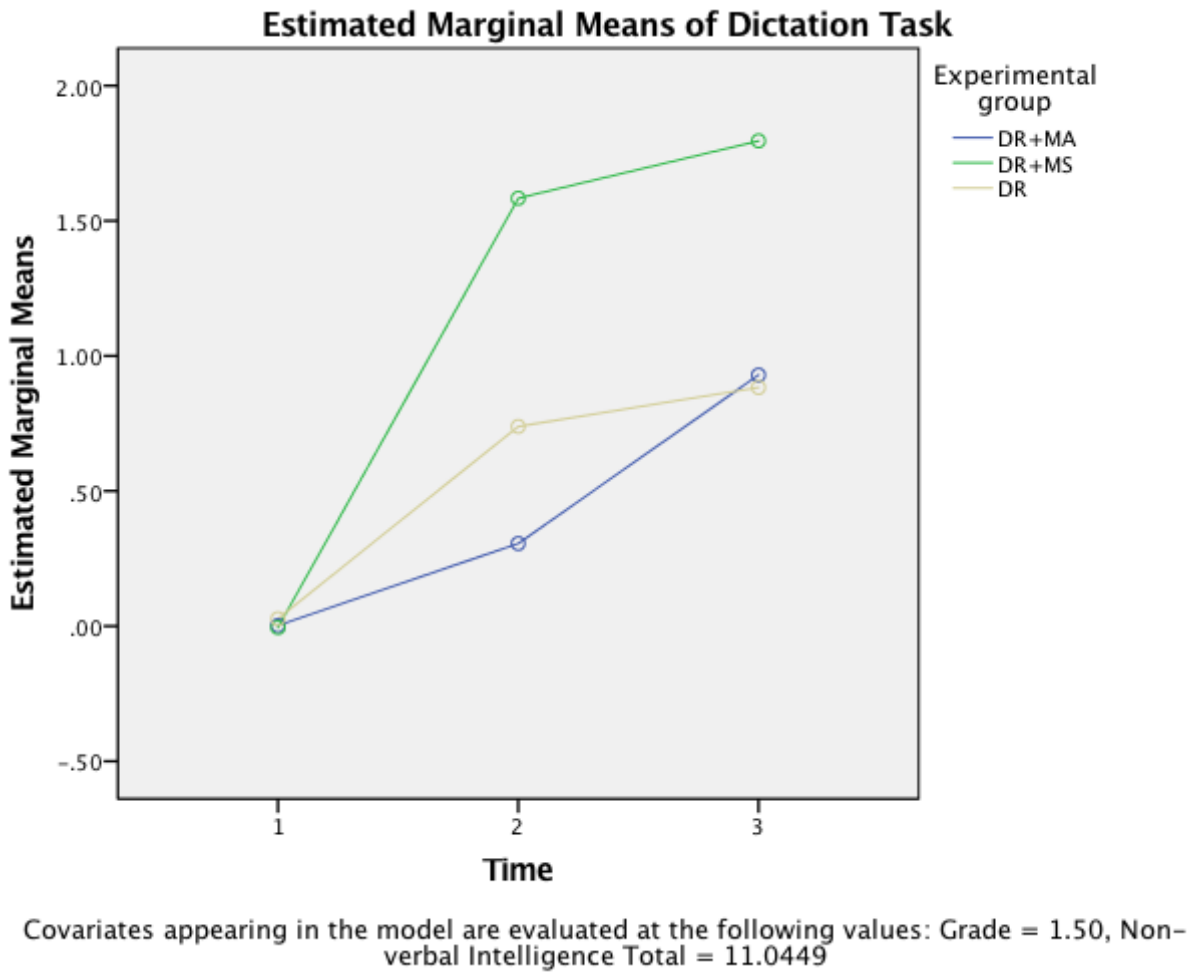
Appendix D

Mean Scores of Chinese Word Reading for All the Groups Across All Three Time Points



Appendix E

Mean Scores of Chinese Dictation Task for All the Groups Across All Three Time Points



Appendix F

Mean Scores of Chinese Rapid Number Naming for All the Groups Across All Three Time Points