INTRODUCTION

Comparing with research into child language development in English in the West, research into child language development in Mandarin in Malaysia is relatively under-explored. As it is clear from the literature, young children acquiring first words by one year, and combining words to form sentences using morphemes (e.g. prepositions in, on) from end second year (Owen, 2016). However, some children may not follow this typical language development due to reasons such as specific language impairment (Reed, 2012). It is essential to identify children at risk for language disorder at an early age and to provide remediation accordingly- the role of speech-language therapists (SLTs). Currently the local SLTs are facing challenges of a lack of clinical assessment tools to diagnose language disorder in Chinese pre-school children. Following the establishment of the first academic programme- the Speech Sciences Programme at *Universiti Kebangsaan Malaysia (UKM),* Malaysia, in 1994, there is a need to develop local child language tests.

The present research represents a pilot study to adapt a contemporary published British English child language test- The New Reynell Developmental Language Assessment Scales (NRDLS)(Edwards, Letts & Sinka, 2011) to Mandarin. NRDLS provides a multilingual toolkit to aid cross-linguistic adaptation. The aim of the present study is to propose adaptation of NRDLS to Mandarin. The present research also aims at providing preliminary language norms for the local Chinese children. The ultimate goal of the present pilot study is to lay the potential for future standardization using a larger sample size and to develop a standardized child language test.

LITERATURE REVIEW

Unlike English, in Mandarin, it is not inflections but particles and word order which control the grammar and meaning of a sentence (Fung, 2009). Mandarin has aspect markers, not tense markers; Mandarin verbs do not express tenses, temporal properties are expressed through temporal terms, aspect markers and contexts (Chen & Shirai, 2010). The four aspect markers in Mandarin are perfective *-le0* and experiential *-guo4* (perfective form c.f. *-ed* in English), progressive *-zai4* and durative *-zhe4* (imperfective form c.f. *-ing* in English)(Duff & Li, 2002).

Another striking difference between the grammar of English and Mandarin is that Mandarin has numeral classifiers. In English, a number is attached directly to a noun without a classifier (e.g. four cars). In contrast, in Mandarin, a classifier must be attached to a noun e.g. *liang4* as in *si4 liang4 che1 (four liang4 cars).*These classifiers classify noun referents based on perceptual or conceptual dimensions: shape, size and function. For instance, *liang4* is used for transports such as car. This type of classifiers is known as sortal classifiers. A second type of classifiers is mensural classifiers which are used to indicate quantity, for instance, *pai2* as in *san1pai2shu4 (three rows of trees)*(Li, Huang, Hsiao, 2010).

Existing literature of Mandarin child language shows that early words, prepositions, temporal terms and classifiers are amongst the most commonly researched areas. In the following section, a pioneer review of relevant local research in these areas will be given.

A PIONEER REVIEW OF STUDIES ON ASSESSING MANDARIN CHILD LANGUAGE SKILLS IN MALAYSIA

PAST STUDIES USING STRUCTURED LANGUAGE TESTS

Since 2000s, about a dozen of cross-sectional clinical linguistic studies of Mandarin child language development amongst the local typically-developing Chinese children have emerged in Malaysia. These are pioneer studies conducted by the first author and her students in the Speech Sciences Programme at *Universiti Kebangsaan Malaysia*. All these small-scale preliminary normative studies have used informal self-devised structured language tests drawing upon the framework of published works in English in the West; or published works in Mandarin in mainland China, Hong Kong or Taiwan. Adaptations were needed because of linguistic and cultural differences that exist between the different populations of children acquiring the same or different languages. Factors influencing test performance such as age, gender, socio-economic status (SES) were investigated in some of these studies since existing research has suggested an association between these factors and child language development globally.

Four studies have been conducted on word development (Chok, 2001; Ooi, 2003; Teng, 2003; Tan, 2004). Chok (2001) investigated acquisition of first words amongst 30 young children (15 boys & 15 girls) aged 10-20 months. She interviewed parents of these children based on a first word checklist. She reported a developmental trend in the acquisition of words. The children acquired approximately 80 receptive words and 150 expressive words by 20 months. Receptive words are found to have emerged before expressive words. Examples of early words: kinship terms *ma1ma0, ba4ba0 (mummy, daddy),* verbs *gei3 (to give)* and toys *qiu2 (ball)*. Girls were found to have outperformed boys. Children with mothers having higher education (Form Six secondary school education known as STPM)(c.f. A-level) and above were found showing comparable language skills with children with mothers having lower education background (below STPM). Despite some discrepancies in the age of acquisition of words (e.g. verbs) between the children and the English-speaking children in the Western studies (e.g. Fenson, Dale, Reznick, Thal, Bates, Hartung, Pethick & Reilly, 1993), the findings of a positive age effect for vocabulary development are consistent with virtually all past studies reported in the literature.

Teng (2003) studied development of Mandarin prepositions by 36 children aged 2-5. The children were asked to carry out action commands for the comprehension of preposition task *(e.g. Put the ball outside the cup).* For production of prepositions, they were asked to answer questions in regards to actions performed by the tester *(e.g. Where is the ball? Under the chair).* Teng reported receptive prepositions were acquired before expressive prepositions, both of which were fully acquired by 4;07-5;00. Overall, *li2mian4, shang4mian4, xia4mian4 and hou4mian4 (in, on, under, behind)* were acquired before *qian2mian4 (in front of).* Teng claimed that his findings of order of acquisition of prepositions were consistent with Zhang's (1985) previous findings on Chinese children in China.

Ooi (2003) examined comprehension of Mandarin temporal terms amongst 36 children (18 boys & 18 girls) aged 3;00-8;11. Amongst others, the temporal terms *xian1… guo4hou4, yi3qian2, yi3hou4 (first… then, before, after)* were tested. The children were asked to carry out a few action commands with toy items *e.g. aeroplane flies first, car then moves.* The *t*emporal term *xian1…guo4hou4 (first….then)* appeared earliest in the children (3;00). The children acquired *yi3hou4 (after)* prior to *yi3qian2 (before).* Ooi's claimed that her finding of a more advanced acquisition of *after* than *before* is consistent with the findings for Mandarin (Zhu, Wu, Ying, Zhu & Zhuang, 1982), but it is inconsistent with the findings for English (Clark, 1973).

Tan (2004) looked into understanding of Mandarin sortal classifiers amongst the 3-6 year-olds. All 40 children (20 boys and 20 girls) in her study were tested for comprehension of classifiers via a picture selection task. All children demonstrated understanding of classifiers tested between 3;00-4;00. On the whole, classifiers with three dimensions (3Ds) were found to have acquired earliest (e.g. *liang4 for lorry),* followed by those with 1D (e.g. *tiao2 for rope),* and finally those with 2Ds *(e.g. ke1 for tree).* Tan claimed that her findings of age and order of acquisition for classifiers were consistent with the previous findings by Chang (1992) on Chinese children in Taiwan.

One study was done on Mandarin sentence development (Phoon, 2003). Phoon investigated children's comprehension of six wh-questions: what-, who-, where-, how-, why-, when-. 48 children (24 boys & 24 girls) aged 2-6 were asked questions about four scenes elicited via composite pictures (e.g. breakfast time). Older children were found to have outperformed younger children. There was no significant gender effect found on the test performance. What- and who-questions were found to have mastered before how-, where-, when- and why- questions. Despite some discrepancies in terms of age and order of acquisition of wh- questions, the findings of an earlier acquisition of what- question comparing to why- and when- questions are claimed to be consistent with the previous findings on children in Taiwan and China and English-speaking children in the West.

The review so far implicates cross-linguistic similarities and differences among different populations of children acquiring Mandarin or Mandarin and English, empirical investigation into local language per se is therefore needed. Influencing factors in child language acquisition namely age, gender and maternal education have not been investigated consistently in the above studies. None of these studies have illustrated a comprehensive developmental profile from words to sentences. Future effort in compiling the scattered pieces of information about Mandarin child language is highly recommended to illustrate a proper developmental profile for the local Chinese children. This developmental profile may well be translated in a structured child language test which reflects current trend of Mandarin child language development.

PAST STUDY USING SPONTANEOUS LANGUAGE SAMPLING

In Malaysia, only one standardised test has been published for Mandarin child language: the Chinese Language Assessment, Remediation and Screening Procedure (C-LARSP)(Jin, Rogayah & Oh, 2012). C-LARSP has been adapted from the British English LARSP (Crystal, Fletcher & Garman, 1989). The study of C-LRASP utilised spontaneous language sampling: e.g. free conversation and story-telling. A total of 130 Malaysian Chinese children aged 1;00-6;11 were employed. A developmental trend was discerned in the children's language skills. The youngest children (1;00-1;06) were found to have produced limited utterances, most of which were symbolic noises, vocatives, nouns and verbs. Short phrases (e.g. SV) were observed to have emerged from 1;06-1;11. A spurt of clause and phrase structures were detected at 2;00-2;05 including use of classifiers. Compound conjunctions e.g. temporal terms *ran2hou4 (after)* were used from 2;06-3;00. More complex utterances were used from 3;00. Spontaneous language sampling can be commended for collecting data in child's most naturalistic environment. However, collecting spontaneous data is effortful and time consuming, and yet desired language structures might not be captured. The authors of C-LASRP have acknowledged that the absence of passive marker *bei4* from the entire speech corpus for instance, might be a consequence of the nature of language sampling namely conversational topics involved, rather than its non-existence in children's grammar system. Letts, Edwards, Schaefer & Sinka (2013) recommends the use of a mixture of

spontaneous language sampling and structured language testing alongside case history interviewing for efficient evaluation of child language skills.

This review indicates that a time-efficient structured language test which is capable in capturing fundamental features in Mandarin is needed. Structured language tests are particularly desirable since the Malaysian SLTs are commonly struggling with heavy daily clinical case load. The present study aimed at adapting NDRLS to Mandarin. The present research also aimed at providing preliminary norms for early language development (comprehension and production) for the Malaysian Chinese children. Using a cross-sectional study approach, the present research aimed to describe language acquisition by 40 Malaysian Chinese children aged 2;00-6;11 based on their performance on the adapted test. The questions addressed were: how does early language comprehension and production develop from 2;00-6;11? Do age, gender and socio-economic status (SES)(maternal education) affect early child language development?

THE NEW REYNELL DEVELOPMENTAL LANGUAGE SCALES (NRDLS)(EDWARDS, LETTS & SINKA, 2011)

TEST CONSTRUCTS

NRDLS is made up of two language scales: Comprehension and Production (Table 1). Each scale comprised six parallel constructs:1. Early first words (nouns). 2. Sentences with two nouns/a locative preposition. 3. Intransitive verbs. 4. Simple sentences with transitive and intransitive verbs. 5. Tenses (present/past contrast). 6. Complex sentences with relative clauses and passive sentences including wh-questions. Two constructs were added to the Comprehension scale only: Inferencing (questions about a scene at a fast food restaurant) and Pronouns. On the other hand, one construct was added to the Production scale: Grammaticality Judgement (yes/no questions about the correctness of grammar in a sentence). The test materials involved in NRDLS were: three key animal figures *(Monkey, Teddy & Rabbit),* other child familiar toy objects *(e.g. ball)* and a colourful test picture book which consist of pictures of the similar toys and objects. In general, for comprehension tasks, the child is required to either perform some action commands with the toy objects, or to point to the target picture in the picture book. For production tasks, the child is required to answer questions verbally.

TABLE 1.English NRDLS- Sections in the Comprehension Scale and Production Scale.

Comprehension Scale	Production Scale
Section A: Selecting Objects. Task: Understand single nouns. *(e.g. ball).*	Section A: Naming Objects. Task: Name nouns. *(e.g. ball).*
Section Bi-ii: Relating Two Objects. Task: Understand simple commands containing two nouns/a preposition. *(e.g. Teddy and Rabbit).* *(e.g. ... in front of lorry).*	Section Bi-ii: Relating Two Objects. Task: Produce two nouns/a preposition. *(e.g. Teddy and Rabbit).* *(e.g. …. in front of lorry).*
Section Ci-ii: Verbs. Task: Understand intransitive verbs. *(e.g. Monkey sit).* *(e.g. Monkey is reading).*	Section Ci-ii: Verbs. Task: Name actions. *(e.g. sit).* *(e.g. read).*
Section Di-ii: Sentence Building. Task: Understand simple sentences containing transitive/intransitive verbs. *(e.g. Rabbit walk).* *(e.g. Rabbit is eating an apple).*	Section Di-ii: Sentence Building. Task: Describe actions using simple sentences containing transitive/intransitive verbs. *(e.g. Rabbit walk).* *(e.g. Rabbit is eating an apple).*
Section E: Verb Morphology. Task: Understand perfect/imperfect aspect markers. *(e.g. … the girl who drinks).* *(e.g. … the boy who ran).*	Section E: Verb Morphology. Task: Describe actions using perfect/imperfect aspect markers. *(e.g. she drinks).* *(e.g. he ran).*
Section F: Pronouns. Task: Understand reflexive/non-reflexive pronouns. *(e.g. is the father feeding himself?)* *(e.g. is the father washing him?)*	Section F: Complex Sentences. Task: Describe pictures using complex sentences containing relative clauses/passive markers. *(e.g. the cat is bitten by the dog).* *(e.g. the boy who is carrying the elephant is smiling).*
Section G: Complex Sentences. Task: Understand complex sentences containing relative clauses/passive markers. *(e.g. the cat is bitten by the dog).* *(e.g. the girl who is wearing a hat is running).*	Section G: Grammaticality Judgement. Task: Indicate whether a sentence is grammatically well-formed by answering yes or no. *(e.g. the rabbit the ball kicked).*
Section H: Inferencing. Task: Understand inferred meaning. *(e.g. who is feeling very happy?)*	

Adapted from Letts, Edwards, Schaefer & Sinka (2013:3).

FACTORS INFLUENCING TEST PERFORMANCE

A total of 1,266 British-English speaking children (626 boys & 640 girls) aged between 2;00-7;06 were recruited in the study of NRDLS (Letts, Edwards, Sinka, Shaefer & Gibbons 2013). These children were divided into 11 six-monthly age bands. Factors influencing test performance: age, gender, SES (measured by years of maternal education and indices of deprivation) were examined. Age is found to affect children's language skills. A mild gender effect is found with girls outperforming boys on the test. Maternal education is found to affect children's language skills only when comparing children with mothers with statutory minimum education (leaving full time education at16) and mothers with higher education (A-levels, diploma, degree and postgraduate) up till the age of 3;06. Poverty is found to have no effect on children's language skills.

METHODOLOGY

RESEARCH APPROACH AND LOCATION

The present study consists of a cross-sectional group study of early language acquisition. The children were recruited from nurseries and via personal contacts (friends, relatives). Data collection took place in the Chinese nurseries and children's home in the Penang island, by the second author as the single tester. The second author is a final year student in the Speech Sciences Programme at *Universiti Kebangsaan Malaysia* who speaks fluent Mandarin. Both parental consent form and head teacher consent form were distributed and collected prior to testing date.

PARTICIPANTS

The children were randomly selected based on the following criteria and were divided into ten six-month age bands (Table 2):

1. Malaysian Chinese ethnic origin, defined as having Malaysian Chinese parents.
2. No reported mental and physical disorders, syndromic disorders or hearing disorders.
3. Dominant in Mandarin as reported by parents (and teachers) and as observed by the tester.
4. Have mother with secondary education qualification (SPM) or further education (diploma and degree).

The children were representative of Malaysian Chinese children. They used Mandarin as their dominant home language. All children aged 3;00 and above (70%) were attending nursery at the time of data collection. The youngest children aged between 2;00-3;00 (30%) who had not attended nursery were taken care by parents or grandparents at home. Mandarin was used as the medium of instruction in the nursery. The children were also learning English and Malay subjects in the nursery. Some of the children had received exposure to Chinese dialects at home (e.g. Hokkien- 38%; Hakka-3%). As none of the parents of the children in the urban area where the present study was conducted had achieved lower than secondary education i.e. primary school education, in the present study, SES was approached by incorporating correlates of statutory minimum education till 17 i.e. Form Five secondary education known as SPM (c.f. O-level) and further education (diploma and degree) only (Table 2). SES measuring by deprivation indices (e.g. income) was excluded since it is not available in Malaysia.

TABLE 2. Number of participants in each age group.

Age group	SES (maternal education)				Total
	Secondary school education		Further education		
	Boys	Girls	Boys	Girls	
2;00 – 2;05	1	1	1	1	4
2;06 – 2;11	1	1	1	1	4
3;00 – 3;05	1	1	1	1	4
3;06 – 3;11	1	1	1	1	4
4;00 – 4;05	1	1	1	1	4
4;06 – 4;11	1	1	1	1	4
5;00 – 5;05	1	1	1	1	4
5;06 – 5;11	1	1	1	1	4
6;00 – 6;05	1	1	1	1	4
6;06 – 6;11	1	1	1	1	4
Total	10	10	10	10	40

TEST MATERIALS

CHARACTERISTICS OF ADAPTATION

In this initial effort of adapting the NRDLS to Mandarin (thereafter NRDLS-M), careful consideration were given to the linguistic and cultural differences that exist between the two populations. The primary sources of the words, phrases or sentences used for translation or modification are summarized below:

1. Consultation with the existing local literature.
2. General intuition, knowledge and observation of Mandarin language used by the local Chinese children.
3. Extensive years of academic and clinical experience namely observation and analysis of child language in Mandarin by the first author.

In general, changes to the original test materials including toys and pictures were kept to a minimum for two main reasons. First, consideration of consistency factors such as style of drawing pictures, materials of drawing; sizes and characteristic of toys. Second, constraints of time and manpower namely, given a single tester over a time span of 9 months. The entire procedure of adaptation involved is summarized below and detailed in the following sections.

PROCEDURE OF ADAPTATION

First phase

1. An initial translation and modification of test items from English to Mandarin.
2. A face validity check on the adapted test by an expert in Mandarin (see further).
3. A pilot study of 5 children to confirm the appropriateness of the adapted test.

Second phase

1. A final revision of the adapted test based on the findings of the pilot study.
2. A second face validity check on the finalized adapted test by an expert in Mandarin (see further).

DETAILS OF ADAPTATION

Examples of adaptation of the test including rationale for modifications were provided in Appendix A. A summary of the test sections of the adapted NRDLS in Mandarin is presented in Appendix B. NRDLS-M comprised 85 items for Comprehension and 75 items for Production. In general, items containing words that were culturally or linguistically inappropriate for the local children were either replaced with more appropriate words or deleted (e.g. sledge, badge)(Appendix A). The present vs. past tense markers (Table 1) were replaced with two Mandarin aspect markers: progressive *-zai4 vs.* perfective *-le0* described earlier. Pronouns *(e.g. him, her, himself, herself)*(Table 1) were replaced with Mandarin reflexive pronoun *zi4ji3 (self).* Due to the potential ambiguity involving with testing of the third personal singular non-reflexive pronoun (anaphor), *ta1(him/her)* was not tested, but was replaced with *zi4ji3 (self).* It is worth noting that, in spoken Mandarin, *ta1* is gender neutral. It is also worth noting that, elsewhere, the pronouns *the boy, the girl, the baby* are less familiar in spoken Mandarin. They were replaced with *di4di0 (younger brother)* for both "boy" & "baby"; and *mei4mei0 (younger sister)* for "girl" (Qi, 2010). As code-mixing is widely used by the local speakers (Lim, Wells & Howard, 2015), translation equivalent terms in English for the three main toy figures of the test: Monkey, Rabbit & Teddy were allowed in the test. For passive sentences (Table 1), the passive marker *-bei4* described earlier was used. Because of linguistic differences, most of the items in Grammaticality Judgement (Table 1) which requires a child to judge whether a sentence heard is grammatically well-formed or ill-formed were

modified by incorporating target structures commonly found in early Mandarin child language e.g. SVO, SVC, AVO (c.f. Jin et al. 2012; Tsang & Stokes, 2001)(Appendix C).

FACE VALIDITY

The face validity check of the adapted test was conducted twice, first before the pilot study and second after the pilot study but prior to the main study. The panel of experts consisted of a Chinese Linguistic lecturer from one other local university in Malaysia, an experienced local Chinese SLT, and a Chinese primary school teacher, all three of whom are fluent in Mandarin and English. Two changes were recommended by the panel: 1. Comprehension of Verbs in Section Ci: The target intransitive verb "wave" was switched to "stand" for *Make monkey wave*. The precise translation equivalent term for "wave" in Mandarin *hui1shou3* is formal and unfamiliar to the local children. The English translation equivalent term "bye-bye" (good bye) is commonly used to signify "wave" hands and so the initial proposal was to modify this item to *bai4bai0 (bye-bye).* However, confusion arose with the logography of *bai4bai4* having an ambiguous meaning with "prayer" in colloquial Mandarin on the scoring form. 2. Use of translation equivalent term for the original English word "Make…" as in *Make monkey wave* across the Instruction Sections. Initially the term *gei3 (give)* was used to replace "make". However, *gei3* is more of a Chinese dialectal term in this context, and therefore *rang4 (let)* a more appropriate Mandarin term was used to replace *gei3* across the Instruction Sections.

PILOT STUDIES

In the first pilot study, a total of five children were asked to do the newly adapted test to confirm the appropriateness of the test items. The results of the pilot study have generally implicated a low test sensitivity level given children across age groups obtaining rather similar test scores.

Careful thought was then paid on the constructs of the adapted test. One obvious reason would be negligence of other important aspects of Mandarin not found in the original English version such as temporal terms and classifiers reviewed earlier in this article. A revision to the adapted test was then carried out by introducing new test components of Mandarin Temporal Terms and Classifiers (Appendices D-E). Three temporal terms were tested on the Comprehension scale: *xian1…. ran2hou4; yi3qian2, yi3hou4 (first…then, before, after)*. Child was required to carry out action commands using the three key animal figures i.e. Monkey, Teddy and Rabbit *e.g. hide rabbit in the box first, then hide monkey.* On the other hand, ten classifiers were tested on both scales. Both sortal classifiers *(e.g. ben3 for book)* and mensural classifiers *(e.g. pai2 for trees)* were incorporated. Child was required to point to one of four pictures in the Comprehension task but to name pictures in the Production task.

TESTING PROCEDURE

The children were assessed individually in a quiet room at home or in the nursery. The adapted New Reynell Developmental Language Scales- Mandarin (NDRLS-M) took approximately 30-40 minutes. All test sessions were recorded using a high quality video recorder (Sony Cyber-shot DSC-W350). These recordings were used for post-hoc scoring accuracy checking purposes. Prior to testing, a brief warm-up free play session (e.g. colouring) was conducted to build up rapport with the children. The children were asked to do the test following the original testing procedure in NRDLS. In each section of the test, clear instructions with trial items were given. No cues

were given to the children and so only spontaneous data was taken into analysis. The children were rewarded with a sticker at the end of the test.

SCORING PROCEDURE

The scoring form of NRDLS-M was devised drawing upon careful translation and modification of the English NRDLS (Edwards et al., 2011). Similar scoring procedure of NRDLS was employed in NRDLS-M. One mark was given for correct responses and a zero mark was given for wrong or nil responses; these marks were entered on the scoring form. Code-switching and code mixing data (e.g. Mandarin-English) were noted in the scoring form.

INTER-RATER RELIABILITY

The data of five children (10%) was independently scored by a local Chinese speech-language therapist who is fluent in Mandarin. The degree of scoring agreement was calculated based on percentage of agreement. Overall, the transcription reliability was high (98% for Comprehension; 96% for Production).

TEST-RETEST RELIABILITY

Five children (10%) were asked to repeat the test within a time span of 7-14 days. Pearson's correlation coefficient revealed a high test-retest correlation (0.991 for Comprehension and 0.994 for Production).

RESULTS

Both quantitative and qualitative analyses were used to address the research questions of the present study.

QUANTITATIVE ANALYSIS

As ceiling effects plus heterogeneity of variance were present in the present data corpus, non-parametric Kruskal-Wallis One-Way Analysis of Variance test (Howell, 2002; Siegel & Costellan, 1988) was used.

DOES AGE AFFECT EARLY CHILD LANGUAGE DEVELOPMENT?

Table 3 shows that there were improvement with age for both Comprehension and Production scales. Statistical analysis confirmed that there were significant age effects on both Comprehension and Production scales (*Kruskal-Wallis* chi^2=32.72, *df* = 9, *p*=0.000 for comprehension; *Kruskal-Wallis* chi^2=33.78, *df* = 9, *p*=0.000 for production). Hence, it can be concluded that age affects early child language development.

Table 3. Scores for Comprehension and Production scales (mean percent correct and standard deviation) by age

	2;00-2;05 (n=4)	2;06-2;11 (n=4)	3;00-3;05 (n=4)	3;06-3;11 (n=4)	4;00-4;05 (n=4)	4;06-4;11 (n=4)	5;00-5;05 (n=4)	5;06-5;11 (n=4)	6;00-6;05 (n=4)	6;06-6;11 (n=4)
Comp. (n=84)	20.24 (1.69)	39.88 (5.37)	51.19 (16.41)	52.38 (2.92)	59.82 (11.19)	70.83 (15.57)	80.06 (3.13)	82.44 (3.69)	81.55 (3.15)	85.12 (2.83)
Prod. (n=74)	1.35 (1.56)	12.84 (11.17)	33.78 (9.03)	43.58 (13.02)	43.24 (17.55)	55.74 (25.67)	75.68 (6.71)	80.41 (2.59)	77.70 (7.84)	92.91 (3.19)

DOES GENDER AFFECT EARLY CHILD LANGUAGE DEVELOPMENT?

Table 4 shows that boys seemed to have outperformed girls for Comprehension scale. Conversely, girls seemed to have outperformed boys for Production scale. Statistical analysis confirmed otherwise that there were no significant differences in the scores for both scales by both boys and girls *(p >0.5).* Hence, it can be concluded that gender does not affect early child language development.

Table 4. Scores for Comprehension and Production scales (mean percent correct and standard deviation) by boys and girls

	Boys (n=20)	Girls (n=20)
Comprehension (n=84)	60.48 (22.79)	64.23 (21.63)
Production(n=74)	52.97 (31.90)	50.47 (30.94)

DOES SOCIO-ECONOMIC STATUS (MATERNAL EDUCATION) AFFECT EARLY CHILD LANGUAGE DEVELOPMENT?

Table 5 shows that children with mothers having high SES (further education) seemed to have outperformed children with mothers having middle SES (secondary school education) for both scales. Statistical analysis confirmed otherwise that there were no significant differences in the scores for both scales by children with both groups of mothers *(p>0.5).* Hence, it can be concluded that SES (maternal education) does not affect early child language development.

Table5. Scores for Comprehension and Production scales (mean percent correct and standard deviation) by children with mothers having middle SES and high SES

	Children with mothers having middle SES (secondary school education)(n=20)	Children with mothers having high SES (further education) (n=20)
Comprehension (n=84)	61.37 (20.70)	63.33 (23.74)
Production (n=74)	50.07 (31.24)	53.38 (31.56)

QUALITATIVE ANALYSIS

In this section, children's acquisition of language structures (from single words through complex sentences) based on NRDLS-M will be discussed in terms of age and order of acquisition, both contribute to preliminary norms (Appendix F). The "age of acquisition" is defined as the age when a target structure was scored correctly by at least 75% of children in an age group (Appendix F). The "order of acquisition" is derived from the most to the least number of children in an age group

scoring a target structure correctly (Appendix F). Some of the ways in which the children approached the task of responding to the test battery including simplification strategies (errors) will also be discussed.

Vocabulary- Early nouns (section Ai-ii). Children started acquiring nouns from 2;00-2;05 with a spurt of growth detected at 2;06-3;05. Several children, particularly those below 3;06, were noticed to have named objects in English *(e.g.bei1zi0→ cup).* Code-switching and code-mixing were scored as wrong responses. This has resulted in the late mastery of nouns in the present study i.e. scoring 100% only by 5;00 (Appendix G). Common errors noted among children below 5;00 were: replacing target words with close-semantic words or associated functions/objects *e.g. zhuo1zi0→yi2zi0 (table→chair); qian1bi3→xie3zi4 or bi3he2 (pencil→ to write or pencil box).*

Phrases containing two nouns/a preposition (section Bi-ii). Children started acquiring phrases containing two nouns from 2;00-2;05; with a full mastery by 5;00 (Appendix G). Common errors discerned were: deleting one of the two objects given *e.g. ping2guo3 he2 chuang2→ chuang2 (apple and bed→bed).* One surprising error made by several children for *hou4mian4 (behind)* was putting Teddy at the back portion (loading space) of the lorry instead of behind the entire lorry. Children demonstrated a more superior comprehension than production for prepositions. Common errors noticed were: giving no responses or replacing target prepositions with *zhe4li3 (here).*

Vocabulary-Early transitive verbs (section Ci-ii). Children showed a steady growth in the development of verbs from 2;00-2;05, with a spurt of growth at 3;00-3.05. It was interesting to note that virtually children across all age groups scored higher in this section C (verbs) compared from the previous section B (nouns)(Appendix G).

*Sentences (section Di-ii).*Children started acquiring short sentences from 2;00-2;05, following by a rapid development from 3;00-3;05 (Appendix F). They started acquiring longer SVO sentences from 4;06-4;11 (Appendix F). The younger children showed a tendency to fail longer sentences (SVO) than shorter sentences (SV).Their errors commonly involved nouns than verbs in a sentence e.g. *Rabbit beat Teddy→Rabbit beat Monkey* rather than performing an inaccurate action for *beat.*

*Aspect markers (section E).*Children started acquiring aspect markers from 2;00-2;05. They showed a more superior comprehension than production of aspect markers (Appendix G). Both non-perfective and perfective aspect markers *zai4* (c.f. *-ing)* and *le0* (c.f. *-ed*) were fully acquired by 4;00-5;05 (Appendix F). Common errors made were: 1. No responses by children aged 2;00-2;11. 2. Omission by children aged 3;06-3;11.

Pronouns (section F on Comprehension scale only). The pronoun *zi4ji3* was acquired by 5;06-5;11 (Appendix F). The younger children below 4;00-4;05 were observed to show a tendency of opting for a positive answer by answering "yes" to questions about pronouns when they were uncertain of the answers.

Complex sentences (section G on Comprehension scale & sections Fi-iii on Production scale). Relative clauses, passive sentences and wh- (who- & which-) questions were acquired late (5;00-6;11)(Appendix F). Common errors made were: 1. Use of short and incomplete sentences for relative clauses 2. Substitution of active voice for passive voice 3. Partial or full repetition of stimulus given for passive sentences or wh- questions.

Temporal terms (section H on Comprehension scale only). As with classifiers, temporal terms were mastered late (5;06-5;11)(Appendix F). The temporal terms *xian1… ran2hou4 (first …then*) and *yi3hou4 (after)(4;06-4;11)* were found to have emerged before *yi3qian2 (before)(5;06-5;11)*(Appendix F). Younger children showed a tendency to miss out one item in a

command *e.g. hide Teddy in the box first, then hide Monkey → hide Monkey*. While older children swapped around the two items in a command *e.g. hide Teddy in the box first, then hide Monkey → hide Monkey in the box first, then hide Teddy.* Virtually all children were observed to have performed worse for Temporal terms in section H than Classifiers in the following section I (Appendix G).

Classifiers (section I on Comprehension scale and section G on Production scale). Classifiers were also acquired late (6;06-6;11)(Appendix F). In general, children showed a more advanced acquisition of sortal classifiers (from 4;06-4;11) comparing to mensural classifiers (from 5;00-5;05). There were no consistent error patterns noted on the Comprehension scale. However, on the Production scale, *ge4 (classifier used for person)* was most frequently used as an erroneous replacement (default) for all target classifiers.

*Grammaticality judgement (section H on Production scale only).*The youngest children (2;00-3;05) could not respond to this task, namely to answer whether sentences heard were grammatically well-formed or ill-formed by saying "yes" or "no". From 3;06-3;06, children showed a steady progress in grammaticality judgement with a significant improvement detected from 4;06-4;11 (Appendix G). Younger children made random errors; they gave more incorrect responses than correct responses. Older children on the other hand made more sporadic errors; they gave more correct responses than incorrect responses. It is worth noting that children's performance for Grammaticality judgement in section H was better than Classifiers in section G (Appendix G).

Inferencing (section J on Comprehension scale only). Children aged 4;00-4;05 and below found the task difficult, namely to answer questions based on a fast food restaurant scene depicted via a composite picture. Even the oldest children (6;06-6;11) scored only about half of this section correctly (50%-60%)(Appendix G). Indirect question such as *who does not come here to buy food to eat?* which requires inferencing skills posed challenges for the children than the more direct questions *e.g. who is feeling very happy?* with cues readily picked up from the picture (e.g. smiling faces). Items involving who- questions were observed to have acquired earlier (5;06-5;11) than items involving whose- questions (not acquired by 6;06-6;11)(Appendix F).

DISCUSSION AND IMPLICATIONS

Effects of age, gender and maternal education on child language development were examined in the present study since existing literature suggests that both internal (biological) and external factors (environment) have an impact on children's language acquisition. The present finding of a positive age effect on language performance concurs with the previous local findings of Mandarin child language (e.g. Teng, 2003) and the previous findings of NRDLS on British English child language (Letts, Edwards, Sinka, Schaefer & Gibbons, 2013). This implicates that language acquisition is universal.

Gender effect, though generally reported to be weaker than age, and sometimes found absent, has consistently been pointing to a more superior language performance of girls than boys (Zhang, Jin, Shen, Zhang & Hoff, 2008). Gender effect was not found in the present study. The study of NRDLS on British children has reported a mild gender effect with girls outperforming boys slightly on the language test. Of the two pilot studies reviewed earlier in this article, one found non-significant gender differences on acquisition of wh- questions (Phoon, 2003), whilst the other found a significant gender effect with girls outperforming boys for the acquisition of first words (Chok, 2001). One explanation for these mixed findings by the three local studies would be

discrepancies of age ranges under investigation. The subjects in the present study and the study of Phoon are older (2 years & above) than the subjects in the study of Chok (2001)(below 2).

Maternal education level does not affect language performance in the present study. In the past study of NRDLS on British English-speaking children, Letts, Edwards, Sinka, Schaefer & Gibbons (2013) found a mild maternal education effect on language performance up till 3;06 only when comparing children with mothers having least education (statutory minimum education by 16) and mothers having further education, higher education and post-graduate education. The input receiving by the older children at the centres is thought to have associated with the diminished effect of maternal education on language acquisition of these children. Supportive language teaching devices such as eliciting conversation, using picture cards and telling stories have been reported to have a positive impact on language development in children (Zhang et al., 2008). This explains the previous findings of Letts et al. and the present findings to some extent. On the whole, the majority of children (70%) in the present study were receiving robust input from the nurseries which were very academic bias for 5 hours a day and 5 days a week. Nonetheless, the correlation of age and maternal education was not examined in the present study.

The present data indicates an overall more superior language comprehension than language production on virtually all test components. The children in the present study exhibited a rapid growth of vocabularies and short phrases and sentences from 2;00-2;11. These findings are consistent with what have commonly been reported in the local child language studies (e.g. Chok 2001; Jin et al., 2012) reviewed earlier in this article and beyond (e.g. Owen, 2016). These findings implicate the influences of language universality, not surprising given that all children are subject to language acquisition device, cognition and world knowledge (Genesee, 2003). Their use of close-semantic words (e.g. table→chair) or associated functions of object (e.g. pencil→ to write or pencil box) indicates partial understanding towards the target words. This observation implicates acquisition of words in young children is gradual. Some children had code-switched Mandarin vocabularies with English, resulting in a late mastery (100%) of vocabularies (by 5;00). This indicates a bilingual Mandarin-English test version is needed to reflect the bilingual or multilingual repertoire of these children, otherwise inaccurate diagnosis of child's vocabulary skills may be made (Lim et al. 2015). Language specific effect seems to govern child language development in the present study as well. Unlike English (Letts et al., 2011), prepositions posed more challenges in Mandarin than verbs, as evidenced in a better test score for verbs in the present study. Despite some discrepancies in the age and order of acquisition for prepositions in the present study and the previous local study (Teng, 2003), both studies have indicated a relatively late development of prepositions (from 3;00 onwards) than verbs (from 2;00 onwards). This result calls for a proposal to swap around Section B (prepositions) with Section C (verbs) in NRDLS-M in order to reflect the developmental approach to the test. Further, since confusion arose with testing the preposition *-behind* using the toy truck, a change of stimulus is recommended. One plausible explanation for which would be: relating the front portion (driver seating area) commonly known as *qian2mian4 (front)* in spoken Mandarin with the back portion (loading area) commonly known as *hou4mian4 (back)* in spoken Mandarin.

One common criticism about structured language test is the small number of stimuli involved for each target structure. However, this test allows comparison of word and grammatical knowledge alongside underlying processes used by children within similar age ranges who had under gone the same test. These processes, manifested through their error patterns, implicate difficulties with memory, processing, auditory discrimination, social cognition and inferencing (c.f. Bishop, 2003). In the present study, the younger children gave more incorrect responses and

no responses than the older children. Whenever in doubt, the younger children showed a tendency to simply respond by answering "yes" or to repeat the whole or part of the test stimuli presented. The younger children deleted one noun in a two noun phrase *(e.g. apple and bed→ bed)* or in a longer SVO sentence *(e.g. Rabbit beat Teddy→ Teddy)* whilst the older children were observed to swap around the two nouns in the SVO sentence *(e.g. Rabbit beat Teddy→Rabbit beat Monkey)*, both of which might implicate a consequence of memory or processing overloading with older children doing slightly better than the younger children.

Both aspect markers perfective *–le0* (c.f. -ed) and imperfective *-zai4 (c.f. -ing)* were acquired later than single nouns and verbs in the present study (4;06-5;05). Chen & Shirai (2010) reported early emergence of *–le0* and *-zai4* in the spontaneous language sampling of four Chinese children subjects with a tender age (1;04-3;05). Li & Bowerman (1998) reported increasing correct responses in the 3 to 6 year-olds deriving from their picture identification tasks of four aspect markers including *-le0 and -zai4* with six types of verbs. The few existing studies have consistently reported a frequent use of *–le0* with achievement, and a frequent use of *-zai4* with activity and statives verbs (Li & Bowerman, 1998; Chen & Shirai, 2010).In the present study, *-le4* was tested with telic verbs (e.g. achievement) as in *e.g. shu1tou2le0 (brushed hair)* while *-zai4* was tested with atelic verbs (e.g. activities) in e*.g. ta1 zai4 he1shui3 (she drinks)* implicating the present measures of aspect markers are developmentally appropriate.

The reflexive pronoun *-zi4ji3 (self)* was acquired by 5;06-5;11 in the present study. The existing studies (e.g. Chien & Wexler, 1987; Hao, Sheng & Gao, 2014) have shown that as with English-speaking children, Mandarin-speaking children achieved a full competency in the comprehension test of reflexive *-zi4ji3* by 5. In the present study, *-ta1 (him/her)* was not tested due to differences of the pronoun system that exist between English and Mandarin. For example, the pronoun her as in *e.g. is the mother painting her?*(a picture showing a mother is painting a picture of a girl who is standing near her) the translation equivalent in Mandarin for which *ma1mao shi4 bu1 shi4 zai4 hua4zhe0 ta1?* can be inferred as:1. The mother is painting a picture of herself . 2. The mother is painting a picture of the girl (reflexive anaphors)(Wang, 2011). To avoid this ambiguity, the target stimulus *-ta1* was changed to *zi4ji3 e.g. is the mother painting her?→ is the mother painting herself?* (*ma1ma0 shi4 bu1 shi4 zai4 hua4zhe0 zi4ji3?).* A second look at the literature, to avoid this ambiguity, researchers have been incorporating for instance, more than one mother in a picture *e.g. is every mother painting her?* (a picture showing three mothers are painting a picture of a girl who is standing near them)(Hao et al., 2014). This strategy is also included in the trial item of the original NRDLS: *is every grandfather painting him?* (a picture showing three grandfathers are painting a picture of a boy who is standing near them). A revision to the section of Pronouns is therefore recommended by incorporating *ta1* using this kind of strategy.

The relatively late acquisition of complex sentences: relative clauses, passive sentences and wh- questions in the present study is consistent with the existing findings for cross languages including English and Mandarin. Examples of studies with comparable findings in Mandarin are, for relative clauses: Hu, Gavarro & Guasto, 2016; for passive sentences: Zeng, Mao & Duan, 2016; for wh- questions: Fahn, 2003; Phoon, 2003. However, as this section provides small sample of relative clauses, passive sentences and wh- questions, an expansion of this section to include more structures is recommended. For example, the grammatical structure such as the order of who- and which- in a wh- question is reported to have an influence on children's understanding of wh-questions namely, object wh- questions vs. subject wh- questions (Fahn, 2003). And likewise, subject relative clauses vs. object relative clauses (Hu et al., 2016). And for passive sentences, whether the agent in a passive sentence is preverbal (Fahn, 2003). In the present study, the passive

marker *bei4* was acquired by 5;00-6;11 on both scales. This finding highlights the strength of a structured language test. This passive marker was found absent in the language sampling study of the local Chinese children (Jin et al., 2012).

Temporal terms were acquired late across languages (e.g. for English: Owens, 2016). Consistent with the local study reviewed earlier (e.g. Ooi, 2003), the temporal term *xian2…ran2hou4 (first… then)* was acquired before *yi3qian2 (before)* in the present study. But this finding is contrary to one of the classic study in English by Clark (1973), implicating cross-linguistic differences. In the present study, children were found to make more errors in the section of Temporal terms than Classifiers. A revision to swap around these two sections in NRDLS-M is recommended.

Classifiers were also late acquired features in the present study. The present study has shown inconclusive results to support the dimensional theory with regards to the rate of acquisition of classifiers (3Ds→1D→2Ds). The present findings on the other hand support the previous findings of an earlier acquisition of sortal classifiers (e.g. shape) than mensural classifiers (quantity)(e.g. Li, Huang & Hsiao, 2010). In the present study, both mensural classifiers tested namely *dui2 (for rocks)* and *pai2 (for trees)* were acquired late (5;00-6;11) on both scales.

The children in the present study have shown awareness about ambient syntactical structures (grammatical judgement) from 4;06 onwards. They scored higher in this test section than the previous Classifier section. A second look at the proposed items for Grammaticality Judgement (Appendix C), there was only one item in this section that had been used to test awareness about morphological violation. All other items had been used to test word order violation. This explains the good performance for this section. As cited in the literature, for Chinese languages, children scored higher on sentences involving word order changes than morphological errors (Tsang & Stokes, 2001). The present findings contribute further knowledge to the literature of Mandarin child language by confirming existing findings of an advantage of word order changes over morphological violation in Mandarin syntactical awareness tasks. A revision to include more items with morphological violation in NRDLS-M is highly recommended.

Items involving who- questions were observed to have acquired earlier than items involving whose- questions in the section of inferencing skills. The present findings support that inferencing skills have involved integration of abilities to understand language with non-linguistic structures such as world knowledge and experience (Edwards, Letts & Sinka, 2011).

CONCLUSION

The present findings suggest that the revised NRDLS-M is developmentally sensitive, although further revisions to the test are required. The present study has demonstrated that age affects language development in the local children aged between 2;00-6;11. Language abilities were found to develop with age. In general, language comprehension was mastered before language production. Gender and maternal education on the other hand were found to have no impact on child language development. In addition to its general contribution to the study of Mandarin child language, the present study provides useful preliminary normative information for professionals such as speech-language therapists dealing with the local children.

LIMITATIONS AND RECOMMENDATIONS

The sample size used in the present study was relatively small due to constraints of time and manpower. Future study using a larger sample size is recommended. Future investigation on the psychometric properties of the revised NRDLS-M such as reliability and validity is desired. The present study has focused on one of the local children's dominant languages i.e. Mandarin only. Future effort in expanding the test to include their one other dominant language i.e. English, namely a bilingual Mandarin-English version is highly recommended.