

**Spoken Mandarin Competence of Chinese Children from Different Language-speaking Homes:
Implications for Mandarin education**

Hock Huan, GOH, Yongbing LIU, Chunsheng, ZHAO

National Institute of Education, CRPP

Nanyang Technological University

Abstract

Bilingualism has been the language education focus of Singapore, but concern has arisen for the increasing imbalance of bilingual abilities among Chinese children, due to shift in the use of home-language from Mandarin to English. In order to understand the difference in Mandarin competence among Singaporean Chinese children, this study analyzes syntactic complexity of Chinese preschool children from different language-speaking homes. This study finds that children from different home-language backgrounds bear differences in utterance types, Mean-Length-Utterance and clause relations, but do not vary much in phrase structure types, clause voice and clause forms. This study supports that Mandarin curriculum developers and educators should take children's different language competences into consideration in their curricular and pedagogic innovation.

Keyword: *Language Competence, Syntactic Complexity, Mandarin Education, Home-language; Singapore*

1. Introduction

The language and education policy of Singapore since her independence in 1965 is based on bilingualism. A problem with this policy is that it focuses on English (the nation's working language) and

relegates the respective ethnic languages (i.e. Mandarin, Malay, Tamil) of its people to the status of “second language” in society and at school (Gopinathan, 1998; Guo, 2004). In the case of Chinese community in Singapore, this policy has resulted in an imbalance in the bilingual development of Mandarin and English languages in young Chinese Singaporean. It is evidently shown that many Chinese families have given up their ethnic mother tongue (Mandarin) and use English as a home-language in the recent years (Gupta, 1997, Dixon, 2004, Guo, 2004). This home-language shift among Chinese families raised several curriculum and pedagogical concerns for learners, teachers and academics (Cheah, 2003; Goh, 2004; CLCPRC [Chinese Language Curriculum and Pedagogy Review Committee], 2004) as it may lead to a situation where Chinese Singaporeans are unable to use Mandarin competently. To address this issue, the Singaporean government engaged Chinese elites and education professionals to form review committee in 1992, 1999 and 2004 to review the Mandarin curriculum and make recommendations for the teaching of Mandarin.

In the recent report of this committee (i.e. CLCPRC, 2004), it was recommended that new curriculum and pedagogical approach be implemented with consideration of the varied Mandarin competence of children with different home-language exposure (particularly English-speaking and Mandarin-speaking homes). Though this recommendation is critical to the revival of the above-illustrated situation, its implementation is not evidence-based from real-time data to understand the differences in Mandarin competence among children of different home-language backgrounds. In order to provide some understanding on the difference, this study draws data from the Singapore Children Spoken Mandarin Corpus for a preliminary analysis on Mandarin competence with the following objectives:

- to analyze and compare the oral linguistic competency differences of children from different home-language backgrounds in terms of their syntactic complexity in Mandarin;
- to provide pedagogical implications for Chinese curriculum developers and educators in Singapore as well as other Chinese communities.

Specifically, this study hopes to address the following questions about Mandarin competence of Singapore Chinese children in term of syntactic complexity:

1. How complex is the syntax of Mandarin spoken by Chinese preschoolers in terms of utterance types, phrasal types, clausal types/relation and utterance length?
2. Does this complexity of syntax vary among Chinese preschoolers with different home-language exposures?

3. What syntactic structures have the Chinese preschoolers mastered and what syntactic structures need to be developed in future curricular attainment?

These questions are indeed important to the curriculum development for Mandarin education in Singapore, but they are yet to be discussed among academics in the field of Mandarin education research.

2. Literature Review

Competence in a language can be broadly divided into linguistic and communicative competences (Gleason, 2001). Linguistic competence refers to one's knowledge of phonology, morphology, semantics and syntax of a particular language, while communicative competence refers to one's knowledge of pragmatics, including social rules/norms for the use of the language. As highlighted in our objective, this study will only focus on one aspects of language competence, i.e. Syntactic Complexity (SC). To explore this aspect of competence in Mandarin spoken by our sample and look for relationship between competency and home-language background of our samples, we will briefly look into some prior research related to syntactic complexity (SC) and syntactic development of Chinese children. Firstly, we discuss the notion of SC and its measurement in the field of child language studies. Secondly, we look at some past findings on syntactic features in Mandarin of Chinese children in China to provide some benchmark reference. Lastly, as there is no specific research on SC and syntactic development on Mandarin spoken by Singaporean children, we will highlight some relevant research that analyze the Mandarin of children in the Singapore bilingual context.

2.1 Syntactic Complexity (SC) and its Measurement

Syntactic Complexity (SC), also known as syntactic maturity or linguistic complexity, has a long history in language acquisition, language learning and language proficiency studies (Iwashita, 2006). SC, in general, refers to the range of forms that surface in language production and the degree of sophistication of such forms (Ortega, 2003: 492). The main assumption for SC is that the language form produced by children/learners will become more complex syntactically as they advance in the language over time. In other words, as a child grows, he/she is expected to generate more syntactically complex language units that reflect his/her growth in language competence. Among research that dealt with SC of children, many of them used Mean Length Utterance (MLU, also commonly known as Sentence Length) as a measurement (see Ingram (1989) for a list of

studies that look into MLU). These studies explored the MLU based on different measurement unit, such as words, syllables or intonation units in a sentence/utterance produced. Other than differences in the measurement unit, these studies also varied in their definition for the boundary of computation. Some research set the boundary using the notion of clause while others define the boundary of utterance using the notion of sentence. No matter which measurement unit or boundary a research used, this approach to compute MLU do face certain limitations. Firstly, such measurement for length can only show how far an utterance has been extended physically, but it does not describe how this utterance is constructed (Szmrecsányi, 2004). Secondly, the measurement of MLU seem to reach a ceiling effect for children beyond the age of three, and fail to effectively distinguish different levels of SC (Segae et al., 2005). Lastly, there is difficulty defining the boundary for computation of MLU, especially for children speech data as their utterance is often incomplete or loosely-conjoined. Though with these drawbacks, MLU is a popular way of analysing SC due to its ease in computation (Szmrecsányi, 2004; Segae, 2005).

Other than computing MLU, some studies have also explored other measurement of SC, such as the counting of phrasal nodes in phrase structure trees (e.g. Johnson 1966; Ferreira 1991; Rickford et al., 1995), and the scoring of language data to compute the Index of Productive Syntax (Scarborough, 1990). Apart from these two measurements, the annotation of language data is another explored approach. Studies using this approach generally develop annotation schemes which consist of Grammatical Categories (e.g. Nouns, Verbs, and Adjectives etc) and Grammatical Functions (e.g. Subject, Verb, Object, Complement etc.). These schemes were used to annotate for grammatical features of language data and the frequency of the annotated features is computed and analyzed for prominent linguistic trends/properties which are then co-related to the specific phenomenon under research. Yaruss (1999) is one of the cited studies that used such approach. In the study, Yaruss aimed to analyze the relations between Sentence Length (SL), SC and stuttering of children (which was often deemed to be caused by extensive length and complexity of sentence). He developed a comprehensive syntactic annotation scheme to describe SC. With this scheme, he annotated 75 utterances from conversational speech samples of 12 boys (aged 40 to 66 months) who stuttered during their 30 minutes free-play interactions with their mothers, and he computed for the SL and SC of these utterances of each child. With these processed data, Yaruss found that both SL and SC showed significant differences between fluent and stuttered utterances produced by the children. It was found that both SL and SC were not predictive for individual child, but SL seemed to be a prominent predictor for the occurrence of stuttering when logistic regression was performed.

Generally speaking, SC can be measured via word-counting (i.e. MLU), phrase-structure node-counting, scoring or annotating. Among these methods, word-counting measurement is straightforward and relatively easy to administer, whereas the other three measurements are more complex and sometimes tedious to administer when the data size for analysis is relatively large. In the comparison of word-counting, node-counting and scoring methodologies, Szmrecsányi (2004) pointed out that results obtained via these methods were rather comparable and he believed that MLU is a less hazardous and most time-effective method to compute SC. However, as mentioned above, MLU can only provide a rough sketch on SC in terms of length, more informative descriptive methods (like annotation) have to be used simultaneously to capture a more holistic view of SC.

2.2 Syntactic Features in Mandarin of Chinese Children

In a review by Li (2002), child language studies in China developed dynamically after the Culture Revolution and these studies were largely of psycholinguistic nature. One of the most important researches till the present is a large-scale study carried out by Zhu and colleagues (Zhu, 1989), entitled “Zhongguo Ertong Qingshaonian Xinli Fazhan Tedian Yu Jiaoyu (Special Characteristics of Psychological Development and Education of Children and Teenagers in China)”, or sometimes simply referred to as the “Shishengshi Yanjiu (Ten Province-and-City Study)”. This project involved about 8,000 informants from 10 provinces and cities of China over a period of 6 years (1983 to 1989). Its main objective is to understand the psychological development of Chinese children and teenagers. Among many sub-studies in this large scale research, there was a study by Shi (1989) which explored the language development of Chinese children at age 3 to 6. Shi's study analyzed the speech data of more than 2,000 Chinese children sampled from the 10 provinces and cities. This study describes various developmental aspects of Mandarin of its informants, including phonological, lexical, syntactic and pragmatic development. For her analysis on syntactic development, Shi further sampled 500 children from the above-mentioned informant size and obtained over 7,000 sentences produced by these 500 children during a task of picture elicitation and recall-description. By analysing these sentences, she found that the sentence length of her informants showed correlation with their age, i.e. at age 3 to 4, the SL is between 4-6 words per sentence; at age 4 to 5, the SL is between 7-10 words per sentence; and by age 5 to 6, the SL is also at an average of 7 to 10 words per sentence but some children having 11-16 words per sentence. When analysing sentence types, she further found that, at any given age, the number of simple sentences (i.e. sentence with one clause) is more than complex sentences (i.e. sentence with multi clauses), but the difference between the number

of simple and complex sentence decreases over age growth, and by age 6 to 7, the number of simple and complex sentences are almost equal (simple sentence – 53.8%, complex sentence – 46.2%). Apart from sentence length and sentence types, Shi also analyzed the simple sentences using three pairs of syntactic constructions (i.e. ‘Subject-Predicate Construction’ vs. ‘Mon-Subject-Predicate Construction’; ‘Subject-Predicate-Object Construction’ vs. ‘Non Subject-Predicate-Object Construction’; and ‘Conjoined Verb Construction’ vs. ‘Noun-Sharing Construction’). She found that more complex constructions (like ‘Subject-Predicate Construction’, ‘Subject-Predicate-Object Construction’, ‘Conjoined Verb Construction’ and ‘Noun-Sharing Construction’) increases over age growth of children while less complex syntactic constructions (like ‘Mon-Subject-Predicate Construction’ and ‘Non Subject-Predicate-Object Construction’) decreases over age growth of children. Besides syntactic construction, Shi also looked at the function of simple sentence in terms of descriptive sentence and declarative sentence. She found that children at all age showed advancement in descriptive sentences than declarative sentences, and she believed that this is related to the daily needs of children to describe things around them. Other than analysing simple sentences, Shi also analyzed the clause relations of the complex sentences (in terms of coordinate and subordinate relation). She found that children at all ages generally have more coordinate complex sentences rather than subordinate complex sentences. She highlighted that these complex sentences observed are mostly of recurring structure and low in psychological quality as there is obvious lack of sequencing, reasoning and other higher order psychological processes.

Generally speaking, Shi’s findings illustrated above are very comprehensive and informative about the syntactic features and its development in Mandarin of children in different age group. Her correlation of these features with age growth provided important benchmarks for analysis of language competency in children of the similar age group. However, there are some setbacks in her study. Firstly, there is no detailed demographic description (like gender, social class etc.) on her sampling of the 2,000 or 500 children. Secondly, she used age interval to group her informants, i.e. age 3-4, 4-5, 5-6, 6-7. These intervals seem to have overlapping points and we are not sure how she grouped the “borderline” children. The number of children in each group is also not provided. Lastly, this study lacks clear description of methodology on data collection and data analysis. Fundamental methodological issues like definition of sentences measurement boundary, annotation categories etc. are not detailed. Though with these setbacks, Shi’s findings on the syntactical features and its development in spoken Mandarin of children still served an important benchmark for present and future studies.

2.3 Relevant Research in Singapore

Studies on child language in Singapore are diverse in the research perspectives and often complex in research methodologies employed. These studies include large-scale longitudinal study, sociolinguistic survey study, psychometric tests study and case study, which employed methodologies like questionnaire survey, psychometric test, language data collection and transcript analysis etc. However, among these studies, specific studies on Mandarin of children are rare to be found, studies with some findings about Mandarin of Singaporean children are often found as subsidiary analysis in the above diverse studies. In the diverse studies mentioned above, the large-scale longitudinal study by Institute of Education and Bernard van Leer Foundation, commonly known as the IE-BvLF study, stands out significantly as one study that researched on some aspects of Mandarin in Singaporean children. This nine year (1983 - 1992) longitudinal project is a multi-disciplinary research on broad areas of language, cognition and socialization of children, which aimed to investigate the cognitive and social developmental processes of Singapore preschool children between ages 3 to 6 years old. Of the language related research in IE-BvLF study, Sim (1988) is one of the attempts that analyzed specifically on bilingual competence of Singaporean preschool children. In her study, various tests were administered to about 1400 children to test various aspects of language competency (such as word knowledge, listening comprehension, verbal fluency etc.) in both English and Mandarin (also other mother tongue depending on the race of the children). Via these tests, Sim found that more Chinese children were successful in their Mandarin word knowledge test than their English word knowledge test. This finding on word knowledge differences among children correlates with the fact that a majority of these tested children use Mandarin at home. In other words, preschoolers who spoke Mandarin at home are more competent in Mandarin as illustrated by the test results. With this observation, it is believed that the linguistic competence (like word knowledge) had obvious relations with the preferred home language of children;

Among language-related research in the IE-BvLF study, Ong (1993) was the only study that looked into Singaporean children's oral lexicon and grammar in Mandarin. Ong's research, entitled "Xinjiapo Xueqian Ertong Huayu Kouyu Cihui" (The Mandarin Oral Vocabulary of Singapore Preschool Children) was carried out from 1983 to 1986, with a sample of 1,200 preschool children (aged 4 to 6 years) in 40 kindergartens and childcare centres. This research used four pictures to elicit verbal responses from individual children, and these responses were recorded. For this research, Ong looked at the oral vocabulary capacity of preschool children (in Mandarin). On top of this vocabulary capacity, Ong (1993) also reported his observation on two types of sentence, i.e. simple and complex sentence, produced by these children. However, besides listing examples for

each type of sentence, he did not provide any statistical findings regarding the percentage of these two types of sentence; nor did he further describe/analyze these examples in terms of syntactic complexity. Being the first specific research on Mandarin of children in the Singapore context, Ong's research bears various flaws. Firstly, the analysis of the lexical items was too general; there was no comparison made in terms of the gender, social economic status, or home-language background etc., of the children. Secondly, as mentioned earlier, Ong's sentential analysis only provided examples of simple and complex sentences, he did not explain exactly what is meant by simple and complex sentences, and he did not give any statistical description to provide a clear picture of how many simple and complex sentences were observed in his study. Neither did he provide any examination on the grammatical complexity of the sentences, which will be important reference for textbook design and Mandarin teaching. Lastly, this study lacked full description of its sample and methodology. For sampling, this study did not give detailed description on the demographics of its informants; neither did it detail the criteria of sample selection. As for methodology, this study also lack justifications on the validity of its methodology in terms of how the fluency of the informants was decided upon.

To sum this review, we have looked at various methods of measurement of SC and some prior research findings on syntax of Mandarin in preschool children. For the measurement of SC, we found that MLU is a convenient tool to compute SC, but due to its limitation in providing in-depth information about SC, MLU has to be complemented with other SC measurement such as annotation to provide a holistic view of SC. As for syntactic features of Mandarin, we understood from Shi that children at different ages vary in competence in terms of sentence length, syntactic constructions and clause relation in complex sentences. From Sim's study, we also knew that children's language competence has correlation with the different home language they are exposed to.

3. Theoretical Basis

In line with the new initiation of the review committee (i.e. CLCPRC, 2004) and with our understanding from our literature review, this study assumes that the Mandarin competence of Singapore Chinese children is related to their home-language exposure/background. Generally, this assumption anchors its support with the behaviourist and the interactionist perspectives of child language studies. The behaviourists believed that parents' language served as a stimulus to children, and children imitate the parents' language as a

response which is further practiced to perfect language production (Goh & Silver, 2004; Jin, 2004; Li, 1995). As for interactionists, they believed that the language of children develops as a result of the complex interplay between the innate capacities and the environment (especially social environment) that the children live in (Foster-Cohen, 1999; Li, 1995, Lightbrown & Spada, 1999). Other than these two perspectives, with focus on bilingual children, research (De Houwer, 1995; Genesee, 1993; Romaine, 1989) has also highlighted the importance of “input” for language learning of bilingual children. They argued that various linguistic features of bilingual children, like the phenomenon of code-switching, are probably the result of modelling language input from parents or other adults in the children’s social environment. With these theoretical perspectives as a basis, this study takes into consideration the assumption on home-language groups by CLCPRC (2004) and assume that Chinese children in Singapore come from three types of language-speaking homes i.e. Predominantly Mandarin-Speaking Homes (PMSH), both Mandarin and English-speaking Homes (MESH) and Predominantly English-Speaking Homes (PESH). This study hypothesizes that PMSH children develop higher Mandarin competence, whereas PESH children develop lower Mandarin competence, and MESH children achieve moderate Mandarin competence between these two groups.

4. Data and Methods

4.1 Data

The data of this research are obtained from the Singapore Children Spoken Mandarin Corpus (SCSMC). This corpus is built with real-time language data collected from 600 Singapore Chinese preschool children (aged 5 to 6 years old) from different types of kindergartens situated in different parts of Singapore (Liu & Goh, 2005). The data are collected via a 15-minute interview and a 15-minute picture elicitation with each child. The aim of the interview and picture elicitation is to maximally capture the interviewee’s Mandarin production to reveal their competence. These interviews and picture elicitations are digitally audio-taped and fully transcribed using *Transcriber* Version 1.4 with reference to the *CRPP Transcription Standard and Convention*. As part of this project, a 25-item questionnaire is also administered to parents/grandparents of each child before actual data-collection. This questionnaire surveyed for the basic demographic information and home-language exposure factors such as:

- Language that the child uses with his/her parents;

- Language that the child engages via TV/radio programmes & books;
- Language that the parents choose to engage via TV/ radio programmes & books;
- Language that the child uses with his/her peers;

With information obtained from the questionnaire, this project computed the Chinese Exposure Index (CEI) for each child. The interval of this index is between -1 to 1, and, the children will spread along this interval in a continuum. However, as SCSMC assumed that there are three typical home-language backgrounds that the children came from (i.e. PESH, MESH and PMSH), the children was categorised into their respective home-language groups based on this CEI (see Table 1).

Table 1: CEI for Each Home-language Groups

| Home Language Group | CEI |
|---|-------------|
| Predominantly English-Speaking Families (PESFs) | -1 to -0.5 |
| Both Mandarin & English-Speaking Families (MESFs) | -0.4 to 0.4 |
| Predominantly Mandarin-Speaking Families (PMSFs) | 0.5 to 1 |

As our study is a preliminary analysis of language data in SCSMC, in terms of syntactic complexity (SC), our data are obtained via random selection of full transcripts of 18 children from the three home-language block grouping in SCSMC illustrated above. Table 2 describes the data we used in this study.

Table 2: Description of Block Sampling

| Home Lang. Grp | Sample | Avg CEI (Std Dev) | Data Duration | T. Turns |
|----------------|--------|----------------------|---------------|----------|
| PESH | 6 | -0.8 (± 0.1) | 176 mins | 1239 |
| MESH | 6 | 0.1 (± 0.3) | 184 mins | 1430 |
| PMSH | 6 | 0.7 (± 0.1) | 179 mins | 1583 |
| Total | 18 | -0.004 (± 0.6) | 539 mins | 4252 |

In the above table, “T. Turn” refers to the total number of student’s audible conversational turns in the entire transcripts. This study has used conversational turn as the measurement unit to describe SC in the data because such units are clear-cut, and they are also relatively objective in the sense that the boundary of utterance is

determined by the speakers rather than the researcher. And this turn is deemed to have a complete meaning for what the child intended to convey in that particular turn of utterance. And at any point of our following discussion on data analysis (especially in Section 5 - Results), the notion of “utterance” is equivalent to conversational turn, unless otherwise stated.

4.2 Methods

As mentioned above in our literature review, there are in fact various ways to examine SC, e.g. Mean Length Utterance (MLU), Phrasal node Counting, Index of Productive Syntax (IPSyn) and syntactic annotation of language data. Some of these methods are simple in computation but provides less details on SC (e.g. MLU), while some methods are more tedious but provides more insights to SC (e.g. annotation). As we hope to gather as much details of SC in this preliminary study, we have adopted the annotation method, and looked into MLU as an additional indicator of competence for children’s utterances beyond a phrase.

4.2.1 Annotation of SC

For Annotation of SC, this study adapted the annotation scheme of Yaruss (1999) with Mandarin specific syntactic categories listed by Shao (2001) and Shi (1989). This scheme has been design to analyze both Inter-Clausal Complexity (i.e. the complexity of clausal forms and their relation between clauses/sentences in an utterance) and Intra-Clausal Complexity (i.e. complexity of phrase structure types within a clause/sentence). The adapted annotation scheme is hence as follows:

Table 3: Description of Syntactic Structure Annotation Scheme

| Lever 1 | Level 2 |
|----------------------------------|---|
| 1) Single Word Utterance (SWU) | |
| 2) Single Phrase Utterance (SPU) | Noun Phrase (NP) <ul style="list-style-type: none"> ▪ Simple NP ▪ Complex NP Verb Phrase (VP) <ul style="list-style-type: none"> ▪ Simple VP ▪ Complex VP Prepositional phrase (PP) Adjectival Phrase (AP) |
| 3) Single Clause Utterance (SCU) | Clause Voice (CV) |

| | |
|---------------------------------|--|
| | <ul style="list-style-type: none"> ▪ Active ▪ Passive |
| | Clause Form (CF) |
| | <ul style="list-style-type: none"> ▪ Declarative ▪ Imperative ▪ Interrogative |
| 4) Multi-Clause (Utterance MCU) | Clause Relations (CR) |
| | <ul style="list-style-type: none"> ▪ Coordinate ▪ Subordinate ▪ Mixed |

The annotation will follow through the above two levels for each utterance (i.e. conversational turn) of each selected transcript of the children. Firstly, the study identifies the utterance types at Level 1 to understand the proportion of different types of utterance produced by the children; secondly, we further identify the phrasal and clausal utterance types in terms of phrase structure, clause voice, clause forms and clause relations, so as to illustrate the nature of phrases, clauses and clausal relationship. The following are definitions for some categories that may be ambiguous:

| | |
|--------------------------------------|---|
| Single-word Utterance (SWU) | <p>Single-Word Utterance refers to an utterance/turn that merely consists of one word. E.g.</p> <p>爸爸。</p> <p><i>bàba.</i></p> <p>Father.</p> |
| Single-Phrase Utterance (SPU) | <p>Single-Phrase Utterance refers to an utterance/turn that merely consists of one phrase. E.g.</p> <p>吃面。</p> <p><i>chī miàn.</i></p> <p>Eat noodles.</p> <p><i>((I) eat noodles.)</i></p> |
| Single-Clause Utterance (SCU) | <p>Single-Clause Utterance refers to an utterance/turn that merely consists of one clause with a “one subject - one predicate” relationship. E.g.</p> <p>我六岁。</p> <p><i>wǒ liù suì.</i></p> <p>I six years old.</p> <p><i>(I am six years old).</i></p> |
| Multi-Clause Utterance | <p>Multi-Clause Utterance refers to an utterance/turn that merely consists of two or more clauses with structures like “one subject – multiple predicate” or “multiple subjects – multiple predicates” relationships. Such utterance will be conjoined by</p> |

conjunctions or display anaphoric relations between the subject and its predicates.
E.g.

婆婆帶我坐車來學校。

pópo dài wǒ zuò chē lái xuéxiào

Grandmother brings me to take a bus to come to school.

(Grandmother takes me to school by bus).

Simple NP

Simple NP refers to noun phrase that consists of a “modifier-head” relationship, whereby the modifier merely consists of a noun or an adjective. E.g.

黃色的裙子

huángsè de qúnzi

yellow colour *poss.* dress

(Yellow dress).

Complex NP

Complex NP refers to noun phrase that consists of a “modifier-head” relationship, whereby the modifier consist of a noun phrase, verb phrase or an adjectival phrase.
E.g.

吃東西的地方

chī dōngxi de dìfāng

eat things *part.* place

(Place for dining)

Simple VP

Simple VP refers to verb phrase that only consists of a main verb in its verb string.
E.g.

看戲

Kàn xì

Watch show.

((we) watch show.)

Complex VP

Complex VP refers to verb phrase that consists of “adverb-verb” relation in its verb string. E.g.

我不喜歡蝴蝶。

wǒ bù xǐhuān hǔdié

I not like butterfly.

(I do not like butterfly.)

In the above sentence, “bu” is an adverb that modifies “xihuan”, the main verb.

Though the above annotation scheme is drawn up with reference to Yaruss (1999), the definition and criteria of the annotation categories have been adapted or simplified to suit the nature of data and objective of this study. Other than the categories defined above, the definitions of all other categories adhere to the conventional definitions of the respective origin, i.e. Yaruss (1999), and reference will be made to Shao (2001) and Lu (2002) for unique syntactic structure in Mandarin and Singapore Colloquial Mandarin respectively.

4.2.2 Computation of MLU for SCU and MCU

As mentioned earlier in the review, the computation of MLU is an established, straightforward and easy-to-use indicator for SC. In general, MLU indicates complexity by measuring the average length of utterance produced by a speaker. The assumption is that the longer the MLU, the more complex the language production of the speaker. For our study, we take MLU as an additional indicator for complexity of utterances that contain one or more clauses, i.e. SCU and MCU. The MLU of SCU and MCU are computed in terms of number of words per utterance. The notion of ‘word’ for computation of MLU is conceived with reference to the “973 Segmentation and Annotation Guidelines for Contemporary Chinese Text” (For details of this guideline, see Shanxi Daxue, 2004); and the notion of utterance as the boundary for MLU computation is conceived as conversational turn as illustrated below:

| | | | |
|---------------------|--|---|---------------------|
| Interviewer: | 你几岁? | | |
| | <i>nǐ jǐ suì?</i> | | |
| | You how many years (old). | | |
| | <i>(How old are you?)</i> | | |
| Child: | 我五岁。 | | |
| | <i>wǒ wǔ suì.</i> | | |
| | Five years (old). | | |
| | <i>(I am five years old.)</i> | | |
| | | } | Conversational turn |
| | | | <i>MLU = 3</i> |
| Interviewer: | 那你家里有几个人? | | |
| | <i>nà nǐ jiāli yǒu jǐgèrén?</i> | | |
| | (And) that you home got how many people | | |
| | <i>(And how many people are there in your family?)</i> | | |

For this example, the child’s length of utterance will be compute as three. For cases where there are stuttering, the repeated word will be counted as one despite the frequency of repetition caused by the stuttering. The MLU will be computed for all utterances/turn for each child and a mean of his/her length of utterance will be produced for SCU and MCU for the particular child. If there are cases of code-switching (CS) in the utterance, the English word will also be considered in the computation of the length of utterance. For utterances that contain ambiguous English words (as in “alright” or “all right” or “bedsheet” vs “bed sheet”), the word’s dictionary entry in *Cambridge Dictionaries Online* (URL: <http://dictionary.cambridge.org/>) will be consulted. Some specific compound words used in Singapore like hawker centre, food-courts etc. will be counted as one word.

As for CS that involves Malay words like pasar-malam (night market), ice-kajang (a kind of dessert) etc. will be considered as one word as well.

5. Results

5.1 Utterance Types

Via the first level of our annotation, we looked at the different types of utterance and their relative proportion in the total number of audible children's turns in each home-language group. The annotation result is summarized as follows:

Table 4: UT Distribution of children from the three Home-language Groups

| Home Lang. Grp | T. Turns | SWU (%) | SPU (%) | SCU (%) | MCU (%) |
|----------------|----------|-------------|-------------|------------|------------|
| PESH | 1239 | 602 (48.6) | 372 (30.0) | 152 (12.3) | 113 (9.1) |
| MESH | 1430 | 354 (24.7) | 512 (35.8) | 247 (17.3) | 317 (22.2) |
| PMSH | 1583 | 469 (29.6) | 521 (32.9) | 312 (19.7) | 281 (17.8) |
| Total | 4252 | 1425 (33.5) | 1405 (33.1) | 711 (16.7) | 711 (16.7) |

The above result showed that most utterances of children in the three home-language groups are SWU (PESH – 48.6.3%; MESH – 24.7% & PMSH – 29.6%) and SPU (PESH – 30.0%; MESH – 35.8% & PMSH – 32.9%). These two types of utterances constitute at least 60% of the total number of audible utterances among each home-language group (PESH – 78.3%; MESH – 60.5% & PMSH – 62.5%). These single word and single phrase utterances are mainly one-word/phrase answers to elicitation from the interviewer. Excerpt 1 is a typical scenario showing how SWU and SPU occurred:

Excerpt 1:

| ID | Turn No. | Speaker | Utterance |
|----|----------|-------------|---|
| S4 | 40 | Interviewer |那你可以告诉我这个图片里头有谁吗? ... Can you tell me who do you see in this picture? |

| | | | |
|----|----|-------------|---|
| S4 | 41 | Student | (.) 哥哥。 <i>[short pause] Brother.</i> |
| S4 | 42 | Interviewer | 哥哥呵？哥哥在做什么？ <i>Brother? What is this brother doing?</i> |
| S4 | 43 | Student | (.) 玩。 <i>[short pause] Playing.</i> |
| S4 | 44 | Interviewer | 玩什么呢？ <i>What is he playing?</i> |
| S4 | 45 | Student | 飞机。 <i>Aeroplane.)</i> |
| S4 | 46 | Interviewer | 玩飞机。Yinghui 家里有没有飞机？ <i>Playing aeroplane. Yanghui, do you have aeroplane at home?</i> |
| S4 | 47 | Student | (.) 没有。 <i>[short pause] No.</i> |
| S4 | 48 | Interviewer | 没有呵？那 Yinghui 有没有坐过飞机？ <i>No? And, Yanghui, have you ever taken a plane before?</i> |
| S4 | 49 | Student | 没有。 <i>No.</i> |
| S4 | 50 | Interviewer | 没有呵。那哥哥在玩飞机，那，还有谁呢？ <i>No. And, the brother is playing aeroplane, then who else is (in the picture).</i> |
| S4 | 51 | Student | (.) 姐姐。 <i>[short pause] Sister.</i> |
| S4 | 52 | Interviewer | 姐姐在做什么呢？ <i>What is the sister doing?</i> |
| S4 | 53 | Student | 画画。 <i>Drawing a picture.</i> |
| S4 | 54 | Interviewer | 画画啊？她在画什么呢？ |

| | | | |
|----|----|---------|--|
| | | | <i>Drawing a picture? What is she drawing?</i> |
| S4 | 55 | Student | (..) 不知道。 <i>[long pause]Do not know.</i> |

The above excerpt came from a 6-year-old girl from PESH, the interviewer began with a general question with the aim to trigger the child's description of characters in a picture; however, this question was not effectively understood by the child as she gave a one-word answer of "brother", focusing on just one character in the picture. Hence, the interviewer continued from the child's response and elicited what the "brother" was doing? And the child again gave a SWU – "play". And when the interviewer followed up with the question on what the brother is playing, the child still gave a single word answer– "aeroplane". To assist the child further, the interviewer tried to link the picture with the child's personal experience at home, but this attempted failed as well, as the child produce SWU in turn 49 and 47. These SWU responses were sustained by the child until Turn 53 where the child responded that the "sister" in the picture is "drawing a picture". And in respond to the interviewer's question on what the "sister" is drawing, she produced a simple verb phrase - "(I) do not know". From this excerpt, it is obvious that the child actually understood the items and scenes illustrated in the picture, but she somehow was not able to make descriptions (beyond a word/phrase) on her own, and depended on cues from the interviewer. We believe that quite a number of such SWU/SPU may have resulted from the dialogic nature of an interview. However, without denying this interview effect on our data, the high proportion of SWU (48.6%) in PESH children (like the one illustrated in the above excerpt) remains prominent in the above annotated result. This high proportion of SWU to some extent highlights the low competence of PESH children in producing utterance with a sentence/clause.

As for utterances beyond a word/phrase, we found that PESH and PMSH children produced more SCU (12.3%; 19.7%) than MCU (9.1%; 17.8%), whereas MESH children produced more MCU than SCU, i.e. 22.2% (MCU) and 17.3% (SCU) respectively. These results seem to illustrate that children (especially PESH and PMSH children) are generally more capable in producing simpler utterances (i.e. SCU) than complex utterances (i.e. MCU). The following two excerpts show typical examples of SCU and MCU produced:

Excerpt 2 (SCU):

| ID | Turn No. | Speaker | Utterance |
|----|----------|---------|-----------|
|----|----------|---------|-----------|

| | | | |
|-----|-----|-------------|--|
| S16 | 535 | Interviewer | (.)那, ah 你看一下, 这个小弟弟在做什么? <i>Then, ah, you take a look (at this picture), what is this boy doing?</i> |
| S16 | 536 | Student + | 1:(..)他要(.)玩水。2:(..)嗯。 <i>1: He wants to play water. 2: Erm.</i> |
| S16 | 537 | Interviewer | (..)玩水啊, 这样好不好呢? <i>Play water, is this (a) good (deed) or not?</i> |
| S16 | 538 | Student | (..)不好。 <i>No.</i> |
| S16 | 539 | Interviewer | (..)为什么不好呢? <i>Why is it not (a) good (deed)?</i> |
| S16 | 540 | Student | (..)玩水会感冒。 <i>Play (with) water will (cause) flu.</i> |

Excerpt 3 (MCU):

| ID | Turn No. | Speaker | Utterance |
|----|----------|-------------|--|
| S9 | 437 | Student + | 1: [很 long]。2: [那钓] 鱼要放什么呢? |
| | | Interviewer | 1: very long. 2: <i>Then, what do you put (as bait) when fishing?</i> |
| S9 | 438 | Student | 钓(.)钓鱼, 可能你要放鱼吃的东西在那里, 它就以为可以吃。它吃了之后呢, then 你就可以 hook 它上来。可是如果抓不到, 你要调这个下去。如果太下面, 你要转回去上去。这个东西转上来就是你太下去了, 转下去就是你太上去了。 <i>Fishing, maybe you need to put things that the fish eat at there, it then think that (bait) can be eaten. After it eats the bait, then you can hook it up. But if (you) cannot catch (it), you need to regulate this down. If (it) is too low, you need to turn (it) up. When this thing turn up, (it means) you are too</i> |

| | | | |
|----|-----|-------------|--|
| | | | <i>low, (but if it) turn down (it means) you are too high.</i> |
| S9 | 439 | Interviewer | /- 哦 -/ 。 |
| | | | <i>Oh.</i> |

Excerpt 2 came from a 6-year-old girl from PMSH. She was asked to describe the action of a particular character in a picture (Turn 535), which she replied in a clause in Turn 536 – i.e. *He wants to play with water*. When she was asked for a reason why “playing with water” is not a good deed, she also replied in a clause that *playing with water will cause flu*. These two clauses are typical SCU we found in our data and the clauses produced by our sampled children are relatively simple, generally carrying the “subject-predicate” grammatical structure. As for Excerpt 3, it came from a 6-year-old boy from MESH. To answer the interview’s query on how to go about fishing, he gave a very complex description from *how to use the bait* to *how to position the fishing rod using the gauge*. In this MCU, the child used a mixture of coordinate and subordinate clause relations to conjoint his clauses so as to show the sequence of action (e.g. 它吃了之后呢, then 你就可以 hook 它上来) and condition of action (这个东西转上来就是你太下去了).

To sum up this section, as highlighted in shades in Table 4, it is interesting to note that the three groups of children in our data each seem to have their “flair” across utterance types, i.e. PESH children are fond of producing SWU, whereas MESH children are fond of producing SPU, while PMSH children are fond of producing SCU. If we agree with Shi (1989), Li (1995) & Jin (2004) that word, phrase and clause generally represent lower to higher competence in Mandarin, then our current results on SWU, SPU and SCU, support our earlier hypothesis on the three group of children, i.e. PMSH children achieve higher Mandarin competence, whereas PESH children achieve lower Mandarin competence, and MESH children achieve moderate Mandarin competence. However, our result on MCU posted a slight challenge to this hypothesis as MESH children showed additional competence in MCU which is of higher syntactic complexity than SCU. We believe that this exception to some extent highlights the linguistic advancement of balance bilingual children, and this advancement may even supersede the predominantly Mandarin speaking children.

4.2 Phrase Utterance

Other than Utterance Types, we also make further analyses on three utterance types, i.e. SPU, SCU and MCU. SPU is the first utterance type that we examine. To examine SPU, we look at the phrase structure of the phrase in each SPU and Table 5 summarizes our annotation result:

Table 5: PT of children from the three Home-language Groups

| Home Lang. Grp | T. SPU | NP (%) | VP (%) | AP (%) | PP (%) |
|----------------|--------|------------|-------------|----------|---------|
| PESH | 372 | 84 (22.6) | 281 (75.5) | 5 (1.4) | 2 (0.5) |
| MESH | 512 | 142 (27.7) | 368 (71.9) | 1 (0.2) | 1 (0.2) |
| PMSH | 521 | 113 (21.7) | 398 (76.4) | 9 (1.7) | 1 (0.2) |
| Total | 1405 | 339 (24.1) | 1047 (74.5) | 15 (1.1) | 4 (0.3) |

From Table 5, the phrase types in SPU are mostly VPs (74.5%) and NPs (24.1). Other than these two phrase types, APs and PPs are scarce in SPU produced in our data (less than 2%). When viewed across home-language groups, there are no obvious differences in this proportion of phrase types in SPU. SPUs of all three groups of children are mostly VPs (PESH – 75.5%; MESH – 71.9% & PMSH – 76.4%), whereas NPs (PESH – 22.6%; MESH – 27.7% & PMSH – 21.7%) are comparatively less, and APs (PESH – 1.4.6%; MESH – 0.2.7% & PMSH – 1.7%) and PPs (PESH – 0.5%; MESH – 0.2% & PMSH – 0.2%) are rarely observed. The following are some typical examples of the different phrase types.

Example 1 (NP):

| | ID | Turn No. | Utterance | Phrase Structure |
|----|-----|----------|--------------------------------|------------------|
| a) | S7 | 056 | 我奶奶。 <i>My grandmother.</i> | Pronoun + Noun |
| b) | S13 | 554 | 五楼。 <i>Fifth floor.</i> | Numeral + Noun |

Example 2 (VP):

| | ID | Turn No. | Utterance | Phrase Structure |
|----|-----|----------|--|--------------------|
| a) | S3 | 341 | Ah. ((Student is thinking)) 买东西。 <i>Ah, buy things.</i> | Verb + Noun |
| b) | S15 | 88 | (.)变(.)变大只。 <i>(Ultraman) grow bigger.</i> | Verb + Noun Phrase |

Example 3 (AP):

| | ID | Turn No. | Utterance | Phrase Structure |
|----|-----|----------|--|--------------------|
| a) | S13 | 595 | 很臭。 <i>(Smoking) is very smelly.</i> | Adverb + Adjective |
| b) | S2 | 394 | 不对。 <i>(Pushing the girl) is not right.</i> | Adverb + Verb |

The above examples are generally simple in structure. Though simple, these phrases displayed different combinations of phrase elements. Example 1a shows the combination of a pronoun and a head noun to form the possessive NP, while 1b shows the combination of numeral and noun to form the numeric NP. Example 2a shows the conjoining of verb and its object to form a VP, whereas 2b shows the conjoining of verb and its complement as a VP. Example 3a shows the combination of an intensifying adverb and an adjective to form an AP, whereby 3b shows the conjoining of a negating adverb and an adjective to form an AP.

In order to understand the nature of the prominent VP and NP in SPUs, we make further annotation on these two types of SPU and obtained the following results:

Table 6: Types of NP and VP of children from the three Home-language Groups

| Home Lang. Grp | NP | Simple (%) | Complex (%) | VP | Simple (%) | Complex (%) |
|----------------|-----|------------|-------------|------|------------|-------------|
| PESH | 84 | 77 (91.7) | 7 (8.3) | 281 | 16 (5.7) | 265 (94.3) |
| MESH | 142 | 138 (97.2) | 4 (2.8) | 368 | 19 (5.2) | 349 (94.8) |
| PMSH | 113 | 110 (97.3) | 3 (2.7) | 398 | 65 (16.3) | 333 (83.7) |
| Total | 339 | 325 (95.9) | 14 (4.1) | 1047 | 100 (9.6) | 946 (90.4) |

Our further analysis of NPs and VPs in SPU does not indicate obvious differences among the three groups of children as well. Generally, children in our data produce more simple NPs (PESH – 91.7%; MESH – 97.2% & PMSH – 97.3%) rather than complex NPs (PESH – 8.3%; MESH – 2.8% & PESH – 2.7%), and they produce more complex VPs (PESH – 94.3%; MESH – 94.8% & PMSH – 83.7%) rather than simple VPs (PESH – 5.7%; MESH – 5.2% & PMSH – 16.3%). Though there are generally no obvious differences across children of the

three home-language groups, we noticed that PESH children showed some advancement in producing complex NPs (8.3%), while PMSH children showed advancement in producing simple VPs (16.3%). The following are examples extracted from the transcripts to illustrate the different types of complex phrases:

Example 4 (Complex NP):

| | ID | Turn No. | Utterance | Phrase Structure |
|----|-----|----------|---------------------------------------|--|
| a) | S17 | 078 | 做运动的东西。 (Facility) for exercising. | Verb Phrase + Noun |
| b) | S18 | 316 | 读书的。 (Shoes) for schooling | Verb Phrase + Nominalization Particle |

Example 5 (Complex VP):

| | ID | Turn No. | Utterance | Phrase Structure |
|----|----|----------|--|---------------------------|
| a) | S8 | 26 | (.) 起来刷牙。 (I) get up and brush teeth | Verb + Verb Phrase |
| b) | S8 | 28 | (..) 换衣服去学校。 (I) change clothes and go to school. | Verb Phrase + Verb Phrase |

The above examples are complex phrases found in SPU. These phrases contain more complex phrase structures as they usually contain a phrase within another. For Example 1a, it has a VP as modifier of the head noun to form the NP, whereas for Example 1b, the VP was nominalized by the particle “的” to form a NP. For example 2a, a verb and a VP is conjoined to form a typical Conjoined Verb Phrase (连动词组) in Mandarin, whereby in Example 2b, two VPs are conjoined to form the Conjoined Verb Phrase.

To sum up our analysis on SPU, our results suggest that children from different home-language exposure show no obvious differences in phrase types of SPU. Children from different home-language groups generally produce a high proportion of VPs and some NPs while other phrase types are barely observed. This high concentration in the production of complex VPs in despite of home-language groups seems to imply that many children are more developed in constructing verb constructions. As suggested by researchers like Tardif, this advancement in VP production may be language-specific. According to her study (Tardif, 2005), she found that Mandarin-speaking children tends to use more verbs than nouns at early age (8-30 months) as compared to

English-speaking children at similar age. This early acquisition and development in verbs of Mandarin-speaker may have contributed to the above advancement in VP production.

4.3 Clause Utterance

In this study, we highlighted two types of clause utterances under analysis, i.e. SCU and MCU. In section 4.1, we illustrated the proportion of these two types of utterances across the three home-language groups. In this section, we shall make a closer examination of these utterances, in terms of Mean Length Utterance (MLU), Clause types of SCU and Clause Relation in MCU.

4.3.1 Mean Length Utterance (MLU)

As mentioned in our literature review, MLU is a common and easily computed measurement for syntactical complexity. Though some researchers, like Sagae et al (2005), believe that MLU has ceiling effect on language data collected from children beyond certain age, other researchers like Shi (1989) and Yaruss (1999) has find it effective on utterance produced by children till age six/seven. In view of its effectiveness and ease in computation, we computed the MLU of SCU and MCU for each child in each home-language group and the results are summarized as follow:

Table 7: MLU of SCU and MCU

| Home Lang. Grp | SCU MLU (Std Dev) | MCU MLU (Std Dev) |
|----------------|-------------------|--------------------|
| PESH | 4.5 (± 0.9) | 9.2 (± 4.2) |
| MESH | 6.1 (± 1.0) | 16.3 (± 4.4) |
| PMSH | 5.1 (± 0.8) | 12.2 \pm (4.6) |
| Overall | 5.2 (± 0.8) | 12.6 (± 3.6) |

Via our computation, we found that the average MLU for SCU showed variation across the three home-language groups. PESH children have the shortest MLU among the three groups, i.e. 4.5 words per SCU. PMSH children have slightly longer MLU in SCU than PESH children, i.e. about 5.1 words per SCU. And; MESH children have the longest MLU in SCU, i.e. 6.1 words per SCU. Children of each group have an average MLU difference of about one word. As for the average MLU of MCU, it shows similar differences as observed in SCU. PESH children again have the shortest MLU in the MCU they uttered, i.e. 9.2 words per MCU. PMSH

children also have longer MLU in MCU than PESH children, i.e. 12.2 words per MCU. And MESH children once again top the three groups by having 16.3 words per MCU. The average MLU difference of MCU between the three groups is about three/four words. Sharing the assumption that longer MLU signifies higher syntactic complexity which in turn signifies higher language competence, our computation results suggest that PESH children show less competence in Mandarin as they generally produce utterances with shorter MLU; whereas MESH children show more competence in Mandarin as they produce the utterances with longer MLU; while PMSH children show moderate competence in Mandarin as they produce utterances with MLU between the PESH and MESH children. And as MESH children have the longest MLU for both SCU and MCU, it concurs our observation in section 4.1 that more balanced bilingual children seem more capable of producing utterance with higher syntactic complexity.

4.3.2 Clause Voice and Clause Form of SCU

Other than MLU, our study further analyzes SCU in terms of Clause Voice (CV) and Clause Forms (CF) via annotation. For CV, we differentiated for Active Voice and Passive Voice for each SCU, and for CF, we differentiated for Descriptive, Interrogative and Imperative form for each SCU. Our results are as follows:

Table 8: CT of SCU of children from the three home-language groups

| Home Lang. Grp | SCU | Clause Voice | | Clause Forms | | |
|----------------|-----|--------------|---------|--------------|--------------|-------------|
| | | Active | Passive | Descrip. (%) | Interro. (%) | Imperi. (%) |
| PESH | 152 | 152 | 0 | 152 (100.0) | 0 (0.0) | 0 |
| MESH | 247 | 247 | 0 | 244 (98.8) | 3 (1.2) | 0 |
| PMSH | 312 | 312 | 0 | 309 (99.0) | 3 (1.0) | 0 |
| Overall | 711 | 711 | 0 | 705 (99.2) | 6 (0.8) | 0 |

As shown in the table above, we did not observe differences among children of the three home-language groups. We found that all the SCU are in active voice (PESH – 100.0%; MESH – 100.0% & PMSH – 100.0%) and almost all the SCU are in descriptive form (PESH – 100.0%; MESH – 98.8% & PMSH – 99.0%). It is also noted from the above table that children from MESH and PMSH have some interrogative utterance like:

Excerpt 4:

| ID | Turn No. | Speaker | Utterance |
|-----|----------|-------------|--|
| S10 | 260 | Interviewer | ……。ok。这个呢？这个是什么？你猜猜看。 Ok. How about this? What is this? Make a guess. |
| S10 | 261 | Student | 哪一个？ Which one? |

Excerpt 5:

| ID | Turn No. | Speaker | Utterance |
|-----|----------|--------------------------|--|
| S16 | 85 | Interviewer | ## % ……。……？你觉得这是在家里的哪一个部分？ ## %? Which part of the house do you think (this picture) is (about)? |
| S16 | 86 | Student + Interviewer | 1: ## % 部分是什[么]？ 2: [哦]，这在家里的那里？ 1: ## % What is (meant by) “part”? 2: ## % oh, where is this in the house? |

The question in Excerpt 4 is typical of interrogative form in our data. The child, a 6-year-old MESH girl, is not certain which object in the picture the interviewer is referring to, hence she asks “which one?” to clarify her doubts. As for Excerpt 5, the 6-year-old PMSH girl does not understand a particular word used by the interviewer, and she asks about the meaning of “部分(part)”. From these two excerpts, we notice that these two girls have shown some ability in questioning, which is not found in most children of our current data.

Summing this our analysis on CV and CF of SCU, the results shown in Table 8 suggest that children from all three home-language groups are rather advanced in producing active voice and descriptive forms at clausal level. Our finding on CF is in line with Shi’s (1989) observation, that preschool children generally produce more descriptive sentences. She believed this advancement in descriptive clause form is probably due to the need to make description in preschooler’s daily interaction. Due to this descriptive nature of the clause, most clauses are hence in active voice. Though our study concurs to Shi’s observation, we shall also admit that this high concentration on active voice and descriptive clause form may also have taken effect from the nature of interview and elicitation in our data. Not denying such effect at play, this high concentration on active voice

clause and descriptive clause form do post concern for development of passive voice clause, interrogative clause form and imperative clause form in the children language. And such concern will have to be addressed by educators and curriculum developers when these children enter formal education.

4.3.3 Clause Relations

Apart from examining clause types in SCU, we also look into the Clause Relation (CR) in MCU in our data. For CR, we differentiate for three types of relations, namely coordinate relation, subordinate relation and a mixture of both coordinate and subordinate relation. Our annotation result is summarized as in Table 9.

Table 9: CR of MCU of children from the three home-language groups

| Home Lang. Grp | MCU (%) | Coordinate (%) | Subordinate (%) | Mixed (%) |
|----------------|---------|----------------|-----------------|------------|
| PESH | 113 | 72 (63.7) | 9 (8.0) | 32 (28.3) |
| MESH | 317 | 185 (58.4) | 86 (27.1) | 46 (14.5) |
| PMSH | 281 | 217 (77.2) | 21 (7.5) | 43 (15.3) |
| Overall | 711 | 474 (66.7) | 116 (16.3) | 121 (17.0) |

From the above result, when viewed across the three CR, we found that most MCU are in coordinate relation (66.7%), where subordinate and mixed relation are generally on a par with less than 1% difference (i.e. Subordinate – 16.3%; Mixed – 17.0%). When viewed across home-language groups, all three groups have more coordinate MCU as well, i.e. PEMH – 63.7%, MESH – 58.4 and PMSH – 77.2. The difference across the three groups mainly lies in subordinate MCU and Mixed MCU. PESH and PMSH children produced more Mixed MCU (28.3%; 15.3%) than subordinate MCU (i.e. 8.0% & 7.5%). MESH children behaved differently from the other two groups by having more subordinate MCU (27.1) than mixed MCU (14.5). Below are examples of coordinate, subordinate and mixed MCU:

Example 6 (Coordinate CR):

| | ID | Turn No. | Utterance |
|----|----|----------|---|
| a) | S5 | 256 | <p>我开, 我放在里面, 我放我的 paper money, 放在里面。 ^^ 真真 ^^ 的 paper money, then 我 close, then 妈妈不知道。</p> <p><i>I open (the box), I put (money) inside, I put my paper money, put inside. ^^ real ^^ paper money, then I close (the box), then (my)</i></p> |

| | | | |
|----|-----|-----|---|
| | | | <i>mother did not know.</i> |
| b) | S14 | 323 | (..)这个男孩子不小心倒掉水。这个男孩子倒掉饭。 (..)这个男孩子乱跑。 (..)这两个男，男孩子在吵架。 (..) <i>this boy falls into the water accidentally. This boy spills the food. (..) This boy runs around. (..) These two boys are fighting.</i> |

Example 7 (Subordinate CR):

| | ID | Turn No. | Utterance |
|----|-----|----------|---|
| a) | S11 | 12 | 嗯。。。没有人带，因为我是坐校车来的。 <i>Erm... no one bring (me to school), because I take school bus here.</i> |
| b) | S12 | 359 | ()上=上次我朋友，上次我姐姐朋友借老鼠，所以我们养了。它很大了，所以我姐姐还他。 () <i>last time, my friend, last time, my sister's friend lent (us) a rat, hence we reared (it). It (got) very big, hence my sister returned (the rat) to him.</i> |

Example 8 (Mixed CR):

| | ID | Turn No. | Utterance |
|----|----|----------|---|
| a) | S3 | 532 | 因为他(.)怕。因为他不要玩，ah，roller coaster，所以他怕。他没有，ah，try 过。我有 try 过。他，他不知道。Ah，是很好还是不好。 <i>Because he (.) is scared. Because he did not want to play, ah, roller coaster, hence he (is) scared. He has not, ah, tried before. I have tried. He does not know, ah, (if it) is very fun or not fun.</i> |
| b) | S5 | 22 | 我(.)我没有，我(.)我如果 primary one 的时((si2))候妈咪说((suol))每天给我一块。 <i>I do not have. When I (enter) Primary One, mummy says that (she) will give me a dollar (as pocket money) everyday.</i> |

The above are typical examples of CR found in MCU of our data. For coordinate CR, Example 6a illustrates how clauses are coordinately conjoined to sequentially describe how the child kept his money in a box; while Example 6b illustrates how the child coordinates clauses in describing different characters and what they do in a picture. For subordinate CR, Example 7a and 7b illustrate the use of subordinating conjunctions like “因为 (because)” and “所以 (hence)” for reasoning the outcomes like ‘why no one send the child (S11) to school’ and ‘why the child (S12) and his sister reared a rat’. For Mixed CR, the child in Example 8a begins her utterance by using subordinating construction to reason why her brother is not willing to try the roller coaster, she then uses

coordinate construction to illustrate that she has tried the roller coaster but her brother has not, and she further illustrates coordinately that her brother does not know if roller coaster is fun or not. And in Example 8b, the first clause is conjoined to the other two clauses in coordination, but it does not have any relation in meaning to the other two clauses. The second and third clause in 8b are more clearly conjoined in subordination with the conjunction “如果 (if)” to state the condition of when the child’s mother will give him a dollar as pocket money.

To sum up this section, our annotation shows that children from all three home-language groups are pretty advanced in producing MCU with coordinate CR, as more than 50% of their MCU are co-ordinately conjoined. Across the home-language groups, differences are observed among PESH, MESH and PMSH children in terms of proportion of MCU with subordinate and mixed CR. By taking CR and home-language into consideration, we found that each home-language stands out on one CR, i.e. PESH has the most MCU with mixed CR, MESH has the most MCU with subordinate CR and PMSH has the most MCU with coordinate CR. From this correlation of CR and home-language group, we can conclude that MESH children are more syntactically advanced than PMSH children as subordinate construction are always deemed to be more complex than coordinate construction. However, what we cannot affirm is whether children with mixed CR utterance are more advanced than children who only produce either coordinate or subordinate CR in their utterance. From the examples we described earlier, such mixture of CR is rather random and it does not have a clear syntactic purpose for co-using both relations. Such MCU seem more like a hybrid of coordinate construction which illustrates a random clustering of information that the child wishes to utter about a topic in mind. For this reason, we believe such Mixed CR will not indicate high competency in Mandarin.

6. Conclusion and Significance

To summarize our findings, this study has outlined the Mandarin competence of children from different home-language groups in terms of syntactic complexity (SC) with the following observations:

Firstly, utterance types showed variation across children from the three home-language groups. The three groups of children in our data each showed prominent production for at least one utterance types, i.e. PESH children produced the most SWU, whereas MESH children produced the most SPU and MCU, while PMSH children produced the most SCU.

Secondly, phrasal types of SPU did not show variation across the three home-language groups. SPUs of all three groups of children are mostly VPs whereas NPs are comparatively less, and APs & PPs are rarely observed. Further analysis of NPs and VPs in SPU did not show differences among the three groups of children as well.

Thirdly, we found that the MLU for SCU and MCU showed variation across children of the three home-language groups. PESH children produced utterances with shortest MLU; whereas MESH children produced utterances with the longest MLU; while PMSH children is somewhere between PESH and MESH children as they produced moderate MLU

Fourthly, our analysis on clause types (i.e. clause voice and clause forms) of SCU showed that all the SCU are in active voice while almost all of SCU are in descriptive form. There are only a few interrogative forms produced by children from MESH and PMSH.

Lastly, in our analysis on clause relations (CR) on MCU, most MCU are in coordinate CR, whereas subordinate and mixed CR are generally less. We also found that each home-language group stands out on one CR, i.e. PESH has the most MCU with mixed CR, MESH has the most MCU with subordinate CR and PMSH has the most MCU with coordinate CR

With these observations, we shall attempt to answer our research questions in Section 1, i.e. (1) How complex is the syntax of Mandarin spoken by Chinese preschoolers in Singapore? (2) Does this complexity of syntax vary among Chinese preschoolers with different home-language exposures? (3) What syntactic structures need to be developed in future curricular attainment?

For question (1), our annotation result showed that the utterances types of our children are mostly simple utterances (i.e. SWU and SPU) rather than complex utterances (i.e. SCU and MCU). Further analysis on phrase types of SPU showed that our children can readily produce utterance with complex verb phrase but not complex noun phrases, and they generally lack other phrase types. As suggested by researchers like Tardif, the advancement in complex VP may be language-specific. Our analysis on MLU found that our children did not have very complex SCU as the average length of utterance is only about 5.2, which is equivalent to MLU of 3-4 years old of their China counterparts (MLU = 4-6). As for our analysis on clause types of SCU, we found that the utterances of our children are also relatively simple as their clauses are merely in descriptive form and active voice, which are simple clause types that are acquired at very young age (Shi, 1989). For clause relation, we found that the MCU produced by our children are mainly coordinate construction which is deemed to be simpler

than other multi-clause construction (like subordinate CR). These findings suggest that the syntax of Mandarin spoken by Chinese preschoolers sampled in our study seem to be rather simple at large.

For question (2), our observation on utterance types showed correlation with home-language exposure, i.e. PMSH children seem to have higher Mandarin competence by having the most SCU among the three home-language groups, whereas PESH children achieve lower Mandarin competence by having the most SWU among the three groups, and MESH children achieve moderate Mandarin competence by having the most SPU among the three groups. However, MESH children show an additional advancement by having the most MCU among the three groups. Other than showing additional advancement in MCU, MESH children also show advancement in MLU as they had the longest MLU for both SCU and MCU among the three groups. Furthermore, in our analysis of clause relation, MESH children also have the most subordinate CR which is deemed to be more difficult in Mandarin oral production (Shi, 1989; Li 1995; Jin, 2004). With these advancements in MCU, MLU and subordinate CR, MESH children, the more balanced bilingual children, seem more capable of uttering syntactically more complex utterances as compared to the less balanced bilingual children of PESH and PMSH. In summary, this study found that children from the three home-language groups show variation on SC in terms of utterance types, MLU and CR of complex utterances. In section 3, we hypothesized that PESH children will have the least SC, whereas PMSH children will have the most SC, while PMSH children will have the moderate SC. But in a slight contrary to this hypothesis, our result suggested that MESH children have the most SC, while PMSH children have moderate SC, whereas PESH children have the least SC.

For question (3), our result showed that children in all three home-language groups show competence in producing SWU and SPU for utterance types. They also advance in active voice clauses and descriptive form clauses in terms of clause types. Furthermore, they are rather developed in complex VPs and simple NPs in terms of phrase types as well. Converse to these advancements in SC, children of all three groups generally lack advancement in complex NPs, passive voice clauses and imperative & interrogative clause forms. Though the lack of such phrase types and clause types maybe a result of data collection method, but the relatively low percentages still posed the possibility for their incompetence. Hence, our findings still carry implications to curriculum developers for ensuring further development for the less advanced syntactic competence found across children of all home-language groups. Apart from these common areas for further reinforcement in the curriculum, the observed differences among children of different groups have to be taken into consideration by educators, so that specific assistance can be provided to their specific needs, such as developing MLU of PESH

children and reviving their chunky coordinate MCU into more fluent coordinate or even subordinate complex constructions.

Before concluding on this study, we believe that our study still has room for improvement as it is a preliminary sketch of SC in the linguistically vibrant Singapore. Firstly, our data under analysis are mainly collected via interview and picture elicitation; hence the lacking of some syntactic features in the children's language output may be methodologically driven. Secondly, we have to admit that the sample of 18 children is still minimal to make claims for all children of different home-language exposure; hence further research with larger sample will shed more light for the correlation between SC and home-language exposure. Thirdly, though our data have demographic information such as age, sex and types of kindergarten the child attended, we have not studied the relation of these demographics with SC, which shall also be interesting to research on. Lastly, when making our annotation, we noticed that there are some other specific features of child language that we have not taken into consideration for this scheme; hence it will be more informative if our scheme can be refined to reflect more syntactic features. Though our study did not cover the above areas of research, we believe that our findings (though preliminary) will provide understanding on SC in Mandarin of Singaporean preschool children.

Generally speaking, during our analysis of SC in our selected data, we noticed that utterances at clause level are not only low in frequency, but low in quality as well. Many of these utterances (especially those with multi-clauses) seem rather loosely or weakly conjoined and sometimes lack logical sequence in description (Example 6a, 7b & 8a). Other than these syntactic problems, some utterance also demonstrated the influence of English (Example 7b “上次我姐姐，……”) and some Chinese dialect syntax (Example 6b “这个男孩子不小心倒掉水。……”). These influences, especially influence from the English syntax, have been a “long-living” pain among Mandarin educators. And with the increase in children from PESH, such syntactic “problem” will soon become more prominent in the formal Mandarin education. However, such influence is almost unavoidable in a multilingual society like Singapore as Mandarin is frequently in contact with English (and also other languages). Though unavoidable, we believe that such syntactic problems can still be resolved, not by avoiding or barring English speech in the formal classroom, but by use of comparative pedagogy or bilingual approach to help children understand the differences between the two syntaxes.

Description of Transcription Symbols

| Symbol | Explanation |
|-------------|---------------------------------|
| ## | Background Noise [Class] |
| () | Ungotten Talk |
| [| Overlap Onset |
|] | Overlap Termination |
| (.) | Gap between Utterances |
| (..) | Extended Gap between Utterances |
| | Prolongation Onset |
| -. :- | Prolongation Termination |
| ((Stutter)) | Stuttering Utterance |
| ((| Comment Onset |
|)) | Comment Termination |

Acknowledgement:

This paper makes use of data from the research project “An Investigation of Chinese Singaporean Children's Oral Competence in Mandarin - To construct at aural oral corpus and to generate a word list” (CRP 11/05 LYB), funded by the Centre for Research in Pedagogy and Practice, National Institute of Education, Singapore (<http://www.crpp.nie.edu.sg>). The views expressed in this paper are the author's and do not necessarily represent the views of the Centre or the Institute.

Reference:

Cheah, C. M. (2003) *Jiaoxue yu ceshi* (Teaching and Testing). Singapore: Singapore Chinese Teachers' Union.

CLCPRC (Chinese Language Curriculum and Pedagogy Review Committee) (2004) *Report of the Chinese Language Curriculum and Pedagogy Review Committee*. Singapore. Chinese Language Curriculum and Pedagogy Review Committee.

- De Houwer, A (1995) Bilingual language acquisition. In P. Fletcher & B. MacWhinney & (Eds.), *Handbook of child language* (pp. 219-250). Oxford: Blackwell.
- Dixon L. Q. (2004) *Learning to Read in a Non-native Language: The relationship between English Oral-Language and Early Literacy Skills of Kindergarten Children in Singapore*, PHD Dissertation. Harvard Graduate School of Education.
- Foster-Cohen, S. H. (1999) *An introduction to child language development*. London ; New York : Addison Wesley Longman
- Genesee, F. (1993) Bilingual language development in preschool children. In D. Bishop and K. Mogford (eds.) *Language development in exceptional circumstances* (pp. 62-79). Hove, East Sussex : Lawrence Erlbaum Associates.
- Gleason J. B. (c2001) *The Development of Language (5th ed.)*. Boston : Allyn and Bacon.
- Goh, C.M. & Silver, R. E. (2004) *Language acquisition and development : a teacher's guide*. Singapore: Longman.
- Goh, Y. H (2004) *Huayuwen Zai Xinjiapo de Xianzhuang yu Qianjing* (The Current Practice and Prospect of Chinese Language Education in Singapore). Singapore: Chuangyiquan Chubanshe.
- Gopinathan, S, Pakir, A., Ho W. K. & Saravanan, V. (eds). (1998) *Language, Society and Education in Singapore: Issues and Trends* (2nd edition). Singapore: Times Academic Press.
- Guo, X. (2004) Haiwai huaren shehui hanyu (huayu) jiaoxue de ruogan wenti: yi xinjiapo weili (Some questions concerning the teaching of Chinese in overseas Chinese communities: With the status in Singapore as an example), *Shijie Hanyu Jiaoxue*. 72(3), 79-88.
- Gupta, A. F. (1997) When Mother-tongue Education is not Preferred, in *Journal of Multilingual and Multicultural Development*. 18(6), 496-506.
- Ingram, D. (1989) *First Language Acquisition: Method, Description, and Explanation*. Cambridge [England] ; New York : Cambridge University Press, 1989.
- Jin H. G. (2004) *Yuyan Huode Lilun Yanjiu* (Study of Language Acquisition) Beijing: Zhongguo Shehui Yanjiu Chubanshe.
- Li, Y. M. (1995) *Ertong yuyan fazhang*. (Children Language Development). Wuhan: Huazhong shifan daxue chubanshe.
- Lightbrown, P. & Spada, N. (1999) *How are Languages Learned?*. Oxford ; New York : Oxford University Press

- Liu, Y. B., & Goh, H. H. (2005, December). *An investigation of Chinese Singaporean children's linguistic competence in Mandarin: A research design*. Paper presented at the 11th International Conference on Processing Chinese and Other East-Asian Languages, Hong Kong.
- Ong, Y. P. (1988) *Xinjiapo Xuequan Ertong Huayu Kouyu Cihui* (The Oral Mandarin Vocabulary of Singapore Preschool Children), Singapore: Paper presented in Special IE-BvLF Project Seminar – Together we grow... how to develop your children's potential.
- Ortega, L. (2003). Syntactic complexity measures and their relationship to L2 proficiency: A research synthesis of college-level L2 writing. *Applied Linguistics*, 4, 492–518.
- Romaine, S. (1999) Bilingual language development. In Barrett, M. (eds) *The Development of Language* (pp.251-276). Hove, East Sussex, UK : Psychology Press.
- Scarborough, H. (1990) Index of Productive Syntax. *Applied Psycholinguistics*. 11, 1-22.
- Sagae, K., Lavie, A., and MacWhinney, B. (2005) Automatic measurement of syntactic development in child language. In *Proceedings of the 42nd Meeting of the Association for Computational Linguistics*. Ann Arbor, Michigan
- Shanxi Daxue (Shanxi University) (2004) *973 Dangdai Hanyu Wenben Yuliaoku Fenci Cixingbiaozhu Jiagong Guifan* (973 Segmentation and Annotation Guidelines for Contemporary Chinese Text). URL: <http://www.chineseldc.org/EN/doc/CLDC-LAC-2003-003/label.htm>
- Shao, J (2001) *Xiandai Hanyu Tonglun* (General Aspects of Modern Chinese). Shanghai: Shang hai Jiaoyu Chubanshe.
- Shi, H. (1990) *Zhongguo ertong qingshaonian yuyan fazhan yu jiaoyu (yi) – 3-6 sui ertong yuyan fazhan yu jiaoyu* (The language development and education of children and adolescents in China (1) - The language development and education of 3-6 years old children). In Zhu, Z. *Zhongguo ertong qingshaonian xinli fazhan yu jiaoyu* (The psychological development and education of children and adolescents in China) . Beijing: Zhongguo Zhuoyue Chubans Gongsi.
- Szmrecsányi, B. (2004) On operationalizing syntactic complexity. *Le poids des mots*. Proceedings of the 7th International Conference on Textual Data Statistical Analysis. Louvain-la-Neuve, March 10-12, 2004. G. Purnelle, C. Fairon and A. Dister. Louvain-la-Neuve, Presses universitaires de Louvain. II: 1032-1039.
- Tardif, T. (2005) But are they really verbs? Mandarin words for action. In K. Hirsh-Pasek & R.M. Golinkoff (Eds.), *Action meets word: How children learn verbs*. Oxford University Press.

- Yaruss, J. S. (1999) Utterance length, syntactic complexity, and childhood stuttering. In *Journal of Speech, Language, and Hearing Research*, 42, 329-344.
- Snow, C. E. (1995). Issues in the study of input: Fine-tuning, universality, individual and developmental differences, and necessary causes. In P. Fletcher & B. MacWhinney (Eds.), *The Handbook of child language* (pp. 180-193). Oxford: Blackwell.