Title	An index of syntactic development for Cantonese-Chinese preschool children
Author(s)	Wan, Ka-hei, Emily; 溫嘉曦
Citation	
Issued Date	2011
URL	http://hdl.handle.net/10722/192908
Rights	The author retains all proprietary rights, (such as patent rights) and the right to use in future works.

An index of syntactic development for Cantonese-Chinese preschool children

Wan Ka Hei, Emily

A dissertation submitted in partial fulfilment of the requirements for the Bachelor of Science (Speech and Hearing Sciences), The University of Hong Kong, June 30, 2011.

Abstract

This research study aimed to develop an index of syntactic development for Cantonese-speaking children. Language samples taken from 14 normal children aged from 4;1 to 5;0, 16 normal children aged from 5;1 to 6;5 and 15 SLI children aged from 5;1 to 6;4 were analyzed and credited according to the framework developed. Normal children aged from 4;1 to 5;0 performed poorer on the index than those aged from 5;1 to 6;5 with the same clinical status. Children with language difficulty performed poorer than their normal age peers on the index as well. The index was validated against MLU and the two indices moderately correlated with each other. A linear combination of age, D and the index was entered into discriminant analysis, yielding a classification accuracy of 86.7% of all the children. The index was found to be a potentially useful clinical marker of SLI yet replication is needed to confirm the findings. Further modification of the index was discussed. The age and language growth sensitivity of MLU was discussed as well.

Introduction

Language sample analysis has been widely adopted to document language growth and assess preschoolers' language ability. Though there are lots of psychometric tests available to achieve these purposes, their ecological validity are always being compared against language samples. Ecological validity refers to how well a measure resembles everyday conversation settings (Costanza-Smith, 2010). Language sample analysis is being considered as more ecologically valid than standardized tests. In language sample analysis, children can decide topic that they want to elaborate and use language in a naturalistic context highly similar to daily conversation (Dunn, Flax, Sliwinski & Aram, 1996). They have to assemble their own sentences using their syntactic knowledge, choose appropriate vocabulary and use language properly to interact with communicators with regard to their pragmatic knowledge (Costanza-Smith, 2010). Getting back to psychometric measures, they are administrated in restricted and controlled environments that hardly resemble daily communication (Crais, 1995). Moreover, children's productions are elicited from sentence completions or unnatural sentence creations (Costanza-Smith, 2010). Psychometric tests can only assess what children know about the syntax and the semantics of a language, while no information about how they integrate form, content and use of language in daily conversation can be obtained (Lahey & Bloom, 1994).

Several indices used of quantifying language development have been established for English-speaking preschoolers using language sample analysis. Among these indices, MLU (Brown, 1973) and D (Malvern & Richards, 1997) are two widely investigated norm-referenced grammatical indices. Owing to their global features, they have been adopted in language sample analysis of Cantonese speaking preschoolers as well. MLU is a calculation of mean morpheme numbers per utterance (Brown, 1973). Studies have shown that MLU correlated with age in normal English-speaking preschoolers (Klee, Gavin & Stokes, 2007;

Blake, Quartaro & Onorati, 1993; Miller and Chapman, 1981), English-speaking preschoolers with language difficulties (Klee, Schaffer, May, Membrino & Mougey, 1989), and normal Cantonese-speaking preschoolers (Klee, Stokes, Wong, Fletcher & Gavin, 2004). As for another quantitative language measure, D, it is a relatively new calculation of lexical diversity. Unlike previous calculation of lexical diversity such as TTR, D was not affected by sample size (Richards & Malvern, 1998). It correlated with age as well in the studies of normal English-speaking preschoolers (Klee et al., 2007) and normal Cantonese-speaking preschoolers (Klee et al., 2004). These two indices were said to be developmentally sensitive.

Apart from being sensitive to age, MLU and D could differentiate SLI children from their normal peers. MLU exhibited group difference between Cantonese speaking children with and without language difficulty (Wong, Klee, Stokes, Fletcher & Leonard, 2010). The same pattern held in D as well in English study (Owen and Leonard, 2002) and Cantonese studies (Wong et al., 2010). The two measures, MLU and D, in combination with age could be a diagnostic marker of SLI both for English-speaking preschoolers (Klee et al., 2007) and Cantonese-speaking preschoolers (Klee et al., 2004). However, a replication study of Klee et al. (2004) did not yield supporting results (Wong et al., 2010).

Grammatical development quantification of Cantonese-speaking preschoolers is limited to MLU and D only. However, in studies of English-speaking preschoolers, criterion-referenced procedures such as LARSP (Language assessment, remediation and screening procedure) (Crystal, Fletcher & Garman, 1976, 1981), Assigning Structural Stage Procedure (Miller, 1981) and Index of Productive Syntax (IPSyn; Scarborough, 1990) are available. LARSP is a qualitative language profile of preschoolers. It analyzed clause, phrase, word and expansion in language samples. In the study by Blake et al. (1993), the phrasal and clausal stages in LARSP were adopted to develop an index. Frequency counts of the targeted phrases and clauses were

computed for each child using language samples. As for Assigning Structural Stage Procedure and IPSyn, both of them were criterion-referenced quantitative grammatical index for English-speaking children. They looked for noun phrase expansions, verb phrase expansions, questions, negation and complex sentence structures in language samples. Quantification procedures were similar to that in Blake et al. (1993). The shared features between the syntactic complexity measure of Blake et al. (1993) and IPSyn was that, they correlated with age and MLU in normal children. The same relationship was found in language delayed children using IPSyn as well (Scarborough, Rescorla, Tager-Flusberg, Fowler, & Sudhalter, 1991). These indices' sensitivity to age did not decrease with age growth while MLU did so after utterance length of 4.5 (Blake et al., 1993) and 3 (Scarborough et al., 1991). So it was concluded that quantitative grammatical index offered more sensitive age differentiation than MLU (Blake et al., 1993; Scarborough et al., 1991).

Qualitative language profile was developed for Cantonese-speaking preschoolers in the study by Fletcher, Leung, Stokes and Weizman (2000). In their study, types and tokens of closed and open class vocabulary, syntax and sentence structures used by two- to five-year-old children were documented from language sample analyses. But this profile was not used for quantifying grammatical development in Cantonese population. Given that MLU and D, the only two quantifying language indices in Cantonese, may not be a sensitive language growth index after utterance length exceeds 3, a development of a syntactic development index will be necessary. It is hoped that the index can document language growth quantitatively and assess language ability of preschoolers as well.

Another call for development of a Cantonese syntactic complexity index stems from a lack of studies in Cantonese-speaking preschoolers isolated form and structure expression development. Only a few studies have investigated into this issue, such as aspect markers

(Leung, 1995), serial verb construction (Cheng, 2006), connectives (Ma, 2006) and classifiers (Tse, Li & Leung, 2007). The study done by Fletcher et al. (2000) was a comprehensive one but it did not identify the qualitative grammatical development. These findings are relatively scarce to define and quantify preschoolers' syntactic complexity using isolated grammatical forms and structures, compared with studies in English. Hence the value of conversational sample in assessing language-impaired children is limited to the use of MLU and D only

If a syntactic complexity index is to be developed for Cantonese preschoolers, its validity should be tested against another language measure. MLU and D are the only norm-referenced indices derived from language sample analysis in Cantonese. However, MLU has been validated against other similar quantitative grammatical summaries (e.g. syntactic complexity measure, Blake et al., 1993; IPSyn, Scarborough, 1990). So MLU is a better option than D. It was found to be highly correlated with quantitative syntactic measures of English-speaking preschoolers, such as IPSyn in the study of Scarborough (1990) (r=.92) and syntactic complexity measure in the study by Blake et al. (1993) (r=.88). Though these reported correlations in English cannot be totally generalized to Cantonese as its grammatical morphemes are not obligatory (To, Stokes, Cheung & T'sou, 2010), they support the use of MLU in validation.

Establishing a valid syntactic complexity index from language sample analysis can contribute to the identification of SLI preschoolers. In Hong Kong, the only available expressive language assessment for preschoolers is The Reynell Developmental Language Scales-Cantonese Version (Hong Kong Society for Child Health and Development, 1987). The ecological validity of psychometric tests in examining preschoolers' language ability has been compared against language sample analysis previously. To be in short, psychometric tests elicit children's language productions in either vocabulary or sentence form in isolation.

Children need not make up their own sentences (Costanza-Smith, 2010). Testing environments hardly resemble real communication situations (Crais, 1995). Language difficulty encountered in daily communication may not be truly reflected and language problems may be left undetected (Miller, 2005). Moreover, standardized tests do not examine error productions in conversation. In the study by Dunn et al. (1996), SLI children produced more errors in spontaneous speech than normal age peers. As for language sample analysis, children's use of language in natural context is assessed. Communication problems caused by morphosyntactic deficit of SLI children can be truly and fully revealed. Qualitative difference in morphosyntactic system between SLI children and normal age peers can be addressed. For instance, SLI children showed less facility in utilizing aspect markers in conversation (Stokes & Fletcher, 2000). Experimental studies also showed that SLI children were less proficient than their age peers in constructing passive sentences (Leonard, Wong, Deevy, Stokes and Fletcher, 2006) and who-object questions (Wong, Leonard, Fletcher & Stokes, 2004). Error pattern produced during conversation can also be examined in language sample analysis.

This project aimed to address the following research questions:

1. Was the index of grammatical development developed in this study sensitive to developmental changes in Cantonese-speaking preschoolers?

2. How strong was the correlation between the index of grammatical development developed in this study and MLU in Cantonese-speaking preschoolers?

3. Did children in the SLI group receive lower score on the index of syntactic development than the normal age-matched peers?

4. Did children in the SLI group produce more syntactic errors than the group of normal children of the same age?

5. How well could the index of syntactic developmental, age, MLU and D in combination discriminate children with SLI from their normal age peers?

Method

Forty five language samples from the language archives in the study of Wong et al. (2010) were analyzed for this study. The forty five samples were divided into three groups. Fourteen normally developing children aged from 4;1 to 5;0 were classified into the TD 4-5 group, sixteen normally developing children aged from 5;1 to 6;5 were in the TD 5-6 group and fifteen SLI children aged from 5;1 to 6;4 were in the SLI 5-6 group. Among these samples, the MLU and D of the language samples from the TD 4-5 group were reported in Wong et al. (2010).

According to Wong et al. (2010), the typically-developing children did not have any history or parental concerns of language, speech & hearing problem. All the children in the TD 4-5 group and the TD 5-6 group scored above -0.5 SD and -0.3 SD respectively on the Receptive subtests of the Cantonese version (Hong Kong Society for Child Health and Development, 1987) of the Reynell Developmental Language Scales – Revised (RDLS-R, Reynell & Huntley, 1985). All the children in the SLI 5-6 scored below -1.2 SD on the mean of RDLS-R, with thirteen children scoring below -1.25 SD. All the children passed Columbia Mental Maturity Scale. They have also passed a pure-tone audiometry screening tested at 0.5, 1.0, 2.0 and 4.0 kHz with loudness at 25-30 dB HL. A history of seizure, neurological or psychological problems should not be present in all the children.

To collect a language sample, each child was involved in a 15- to 20-min conversation with one or two speech therapist research assistants trained in language sampling previously. Then a team of eight students majoring in Speech and Hearing Sciences, Psychology and Chinese Linguistics transcribed the samples after training. For further details of language sampling and transcription, please refer to Wong et al. (2010).

Four subtests of syntactic structure were identified: noun phrase elaboration (NP), verb phrase elaboration (VP), complex sentence structure (SS), question (Q) and connectives & sentence adverbs (CON/SA). These categories were adopted and validated in syntactic development studies of English-speaking preschoolers (Scarborough, 1990; Miller, 1981). Items in each category are as follows: A noun phrase comprises a head noun with an adjective, a relative clause and a classifier as modifiers (Zhang, 2007; Matthews & Yip, 1994); a verb phrase can be elaborated by an aspect marker, a verbal complement, a modal auxiliary, a manner predicate and a clausal complement; SS includes pivotal construction, prepositional phrase and serial verb construction; Q includes yes/no question and wh-question; CON/SA includes connectives and sentence adverbs. For details and examples of the items included in each category, please refer to Appendix A.

The chosen items should possess potential in illustrating developmental progress in preschoolers. For instance, two-year-old children used "result + *zo2*" (*laan6 zo2*) to describe a consequence while three-year-old children used resultant verb complement (verb + result + *zo2*) (*dit3 laan6 zo2*) and older children can produce it with greater diversity (Cheung, 1991). The types of classifier employed by preschoolers increased with age, from three to five years old (Tse et al., 2007). Children started to use more general classifier "*go3*" to more specific sortal classifiers (e.g. *tiu4, bun2*), and then advanced to mensural classifiers (e.g. *deoi3*, *hap6*) and measurement classifiers (e.g. *jat6, jyut6*). These items should also be potentially discriminative between SLI and typically-developing children. For instance, SLI children produced qualitatively and quantitatively less aspect markers than their age-matched counterparts during conversation (Fletcher, Leonard, Stokes & Wong, 2005). In some experimental studies, they were also less capable of producing correct shape and function classifiers and omitted classifiers more frequently than age-matched peers (Stokes & So,

1997). Construction of passive sentences posed more difficulty to SLI children than to normal age-matched peers as well (Leonard et al., 2006).

Refinement was made to eliminate items showing weak changes, i.e. those rarely produced or equally well-mastered by children in both age groups (Scarborough, 1990). For example, negation was removed from the index as it was found to be equally well-mastered by typically-developing children in both age groups. Novel structures identified occasionally will be added to the index for analysis. To reveal the syntactic development among the two normal age groups, items that emerge as later developmental form will receive more credits after refinement. For instance, why- and how-questions receive one more credit than what-, who- and where-questions, as the study by Fletcher et al. (2000) showed that the former question types emerged later than the latter types.

Targeted items were then identified from each utterance in the samples. Three kinds of productivity criteria (*contextual criterion, context criterion and phrasal criterion*) were adopted form IPSyn (Scarborough, 1990) for crediting a second example of the same items. If a form appears for its second time, it must meet the *contextual criterion*, i.e. be in a distinct adjacent context for crediting of a second point. The *lexical criterion* should be satisfied as well, i.e. to be expressed in a different form compared with the first appearance. If the item is a sentence/phrase, the second example should meet the *phrasal criterion*, i.e. half of its wordings should be different from the first example. A grammatical item can be credited for a maximum of three times if the above-mentioned criteria can be satisfied. Syntactic errors will be identified in each sample and being classified into addition, substitution, omission and wrong order of word. To address inter-rater reliability, 10% of the samples, i.e. 1 sample from the TD 4-5 group, 1 from the TD 5-6 group and 2 from the SLI 5-6 group, would be marked by another rater again. The percentage of agreement is 95.1%, revealing high accuracy of agreement between the two raters.

Results

In this study, all of the group comparisons were conducted using one-way ANOVA procedures followed by Tukey HSD post hoc comparison, conducted at $p<.05$ level. Pearson's r tests were used to quantify the strength of the correlations among the variables of interest. According to Cohen (1988), *r* ranging from 0.9 to 1 represents a very high correlation, 0.7 to 0.9 represents a high correlation, 0.5 to 0.7 indicates a moderate correlation, and 0.3 to 0.5 indicates a low correlation.

For our questions of interest, the reports focus on the comparisons between the TD 4-5 group and the TD 5-6 group, and between the TD 5-6 group and the SLI 5-6 group. Comparison between the TD 4-5 group and the SLI 5-6 group will be presented only when there were significant differences.

Results of The Standardized Tests

The descriptive statistics for CMMS, CRVT, RDLS-E and RDLS-R of each group are summarized in Table 1, accompanied with the results of Tukey HSD post-hoc comparisons in which significant difference was marked at $p< .05$.

Table 1. Sample size, age (mean and range) and scores of the standardized language test (Mean, SD and range) for each group

Group		**TD 4-5**	**TD 5-6**	**SLI 5-6**
N		14	16	15
Age(month)	M (SD)	55.71 $_{a, b}$	67.25 $_{a}$	67.13 $_{b}$
	Range	49 - 60	61 – 77	61 – 76
CMMS	M (SD)	108.93$_{c}$ (5.99)	113.06$_{d}$ (7.21)	100.07$_{c,d}$ (9.55)
	Range	98 – 120	101 – 126	83 – 111
CRVT	M (SD)	58.79$_{e}$ (4.30)	61.69$_{f}$ (2.27)	53.80 $_{e, f}$ (4.41)
(max=65)	Range	52 -64	56 – 65	45 – 63

RDLS-E	M (SD)	57.57 $_{g}$ (5.02)	62.69$_{g,h}$ (4.22)	54.73 $_{h}$ (5.95)
(max= 73)	Range	49 – 66	56 – 70	44 – 66
RDLS-R	M (SD)	55.64 $_{i, j, k}$ (3.23)	59.81 $_{j, i, k}$ (2.86)	48.47 $_{k, i, j}$ (4.47)
(max=67)	Range	50 – 62	56 – 65	41 – 54

Note: Means having the same superscript differ significantly at p<.05.

CMMS = Columbia Mental Maturity Scale presented in standard score; RDLS-E = Reynell Developmental Language Scales- Expressive raw score; RDLS-R = Reynell Developmental Language Scales- Receptive raw score; CRVT= Hong Kong Cantonese Receptive Vocabulary Test presented in raw score

According to the ANOVA results, there was a significant difference in age among the three groups, $F(42, 2) = 29.28$, $p<.0001$, and the follow up comparisons confirmed that the TD 5-6 group and the SLI 5-6 group did not differ in age, $p= .997$.

ANOVA results revealed significant group differences in all standardized tests ([CMMS]: F (2, 42) = 11.26, $p<.0001$; [CRVT]: F (2, 42) = 17.43, $p<.0001$; [RDLS-E]: F (2, 42) = 9.74, $p<.0001$; [RDLS-R]: F (2, 42) = 39.60, $p<.0001$). For CMMS, which is a non-verbal cognitive test, post-hoc comparison indicated that no significant difference was found between the two TD groups, $p=.321$, while the TD 5-6 group (mean= 113.01, SD= 7.21) scored significantly higher than SLI 5-6 group (mean= 100.01, SD= 9.55), $p< .0001$. The TD 4-5 group (mean= 108.93, SD= 5.99) has significantly higher score than the SLI 5-6 group as well, $p=.01$. The same pattern across the three groups was found in CRVT, i.e. the two TD groups do not differ significantly, $p= .099$; whereas the TD 6-5 group (mean= 61.69, SD= 2.27) scored significantly higher than the SLI 5-6 group (mean= 53.8, SD= 4.41), $p<.0001$. The TD 4-5 group (mean= 58.79, SD= 4.30) scored significantly higher than the SLI 5-6 group as well, $p= .003$.

With regard to the expressive subtest of RDLS, the TD 5-6 group (mean= 62.69, SD= 4.22) scored significantly higher than their normal younger peers (mean= 57.57, SD= 5.02), p= .024; and the SLI 5-6 group (mean= 54.73, SD= 5.95) respectively, p <.0001. For the receptive subtest, given that the scores were used as the inclusion criteria for the SLI group, as expected, the TD 5-6 group (mean= 59.81, SD= 2.86) had significantly higher score than the TD 4-5 group (mean= 55.64, SD= 3.23), p=.008; and the SLI 5-6 group respectively (mean= 48.47, SD= 4.47) , p< .0001. And the younger TD group scored significantly higher than the SLI 5-6 group as well, p< .0001.

Results of the Syntactic Complexity Measure, the Component Measures, MLU and D

The descriptive statistics of MLU, D, the index and its component measures (noun phrase expansion measure [NP], verb phrase expansion measure [VP], sentence structure measure [SS], connective and sentence adverb measure [CON/SA] and question measure [Q]) of each group are presented in Table 2.

Table 2. Mean, SD and range of MLU, D, the syntactic complexity index and the subtests of each group

Group		TD 4-5	TD 5-6	SLI 5-6
MLU	M (SD)	4.73 (0.96)	4.57 (0.60)	4.06 (0.68)
	Range	3.47 - 6.59	3.62 – 5.62	2.82 – 5.22
D	M (SD)	58.95 (13.08)	66.39_{a}(14.32)	48.76_{a} (8.48)
	Range	39.65 – 85.72	40.01 – 90.68	29.33 – 59.90
Index	M (SD)	$38.64_{b,c,d}$ (6.81)	$47.3_{c,b,d}$ (9.46)	$27.73_{d,b,c}$ (8.88)
(max= 96)	Range	25 – 51	30 – 62	15 – 45
NP	M (SD)	3.79_{e} (1.63)	3.88_{f} (1.41)	$2.27_{e,f}$ (1.71)
(max= 12)	Range	1 – 6	1 – 6	0 – 5
VP	M (SD)	17.5_{g} (3.46)	21.13_{h} (5.06)	$12.87_{g,h}$ (4.27)

(max= 42)	Range	10 – 23	11 – 29	7 – 23
SS	M (SD)	7.43 (3.72)	9.25_{i} (3.99)	5.2_{i} (3.19)
(max= 18)	Range	2 - 14	2 – 14	0 – 10
CON/SA	M (SD)	6.57 (3.63)	7.23_{j} (3.17)	3.73_{j} (3.45)
(max =12)	Range	0 - 12	4 – 12	0 – 12
Q	M (SD)	3.36_{k} (1.69)	$5.68_{k, l}$ (2.91)	3.67_{l} (1.91)
(max= 12)	Range	0 – 7	1 – 11	0 – 7

Note: Means having the same superscript differ significantly at $p < .05$.

NP= Noun phrase expansion; VP= Verb phrase expansion; SS= Sentence Structure;

CON/SA= Conjunction and sentence adverbs; Q= Question

In the following reports, one-way ANOVA results on the abovementioned language measures will be presented first, followed by the post-hoc comparisons.

Group Comparisons for MLU, D, the Index and the component measures

Differences among the three groups in MLU were approaching significance (F (2, 42)=3.18, p=.052).The mean MLU of the SLI group (mean=4.06, SD=0.86) was lower than the TD 4-5 group (mean= 4.23, SD= 0.96). Post-hoc analysis confirmed that this is the only comparison that was approaching significance, p=.055.

With regard to D, significant group difference was found, (F(2, 42)=8.051, p=.001). Post-hoc comparison revealed that the TD 5-6 group (mean=66.39, SD=14.32) had higher D than the TD 4-5 group (mean=58.95, SD= 13.08) but the difference was not statistically significant, p=.232. The TD 5-6 group has significantly higher D than the SLI 5-6 group (mean= 48.76, SD= 8.48), p=.001. This is the only significant group difference found in D. So the TD 4-5 group did not score significantly higher than SLI 5-6 group, p=.077.

As for the index, significant group differences were found (F(2, 42)=8.07, p=.008). The

post-hoc analyses revealed that, the TD 5-6 group (mean=47.3, SD=9.46) scored significantly higher than the TD 4-5 group (mean=38.64, SD=6.81) (p= .021). The TD 5-6 group has significantly higher score than the SLI 5-6 group as well (mean=27.73, SD=8.88) (p< .0001)

With regard to the subtests, significant group differences were found in all the indices ([NP]: F(2, 42)=4.92, p=.012; [VP]:F(2, 42)=13.98, p< .0001; [SS]: F(2, 42)=4.83, p=.013; [CON/SA]: F(2, 42)=4.54, p=.016; [Q]: F(2, 42)=4.79, p=.013). Results of post-hoc comparisons are as follows:

For NP, no significant difference was found between the two normal groups in NP (p= .987). However, TD 5-6 group (mean= 2.27, SD=1.71) scored significantly higher than SLI 5-6 group (mean= 3.88, SD=1.41) (p=.019). The TD 4-5 group (mean=3.79; SD=1.63) had significantly higher score than the SLI 5-6 group as well (p=.035). The same pattern was found across the three groups in VP. The two TD groups did not score significantly different from each other (p= .070) while TD 5-6 group (mean=21.13; SD=5.06) scored significantly higher than the SLI 5-6 group (mean= 12.87, SD= 4.27) (p<.0001). The TD 4-5 group (mean=17.5; SD=3.46) scored higher than the SLI 5-6 group as well (p=.017).

In SS, significant difference was only found between the TD 5-6 group (mean= 9.25, SD= 3.99) and their SLI age peers (mean=5.2, SD=3.19), p=.009. The same pattern of group difference was also found in SA/CON, i.e. TD 5-6 group (mean=7.23; SD=3.17) scored significantly higher than the SLI 5-6 group (mean=3.73; SD=3.45) (p=.017). This was the only significant difference found across the groups in this subtest. The last subtest to be reported is Q. The TD 5-6 group (mean=5.68; SD=2.91) scored significantly higher than their normal younger peers (mean=3.36; SD=1.69) (p=.850). The TD 5-6 group had significantly higher score than the SLI 5-6 group (mean=3.67; SD=1.91) as well (p=.045), whereas no significant difference was found between the TD 4-5 group and the SLI 5-6 group (p=.928).

The results to be reported next is the correlation between variables of interest

Relationship between Age, MLU and the index

The correlations among age and MLU, index and D were conducted on the 30 children in two normal groups only.

Relationship of between Age and other language measures (MLU, the index)

Age and MLU weakly, negatively and insignificantly correlated with each other, $r(30)= -.122$, $p=.522$, two-tailed. The negativity can be attributed to the fact that some children from the TD 5-6 group produced utterances that were on average shorter than those of the TD 4-5 group, resulting a slightly lower mean MLU in the TD 5-6 group compared with that in the TD 4-5 group ([TD 4-5]: mean=4.73; [TD 5-6]: mean= 4.57). Age and the index mildly correlated with each other, $r(30)=.345$, $p=.062$, two-tailed. The relationship was approaching significance.

Relationship between the index and MLU

The index and MLU moderately correlated with each other, $r(30)= .502$, $p= .005$, two-tailed. The correlation between them was studied again after removing the effect of age. The partial correlation was significant, $r(30)=.584$, $p=.001$, two-tailed, showing that the relationship between MLU and the index was significant even if the variance due to age was removed. The relationship is depicted in Figure 1 in Appendix B.

Since the index exhibited group difference between the three groups and D did so between the TD 5-6 group and the SLI 5-6 group, discriminant analysis was carried out as proposed to investigate if a linear combination of age, MLU, D and the index could distinguish the three groups of children from each other, and the degree of accuracy that these measures can correctly classify the children into the respective groups. However, since MLU did not show significant group difference, it was not entered into the analysis.

Discriminant Analysis

Age, the index and D were entered into linear discriminant analysis. Since the index and D correlated with age, they were entered into the analysis simultaneously. The Box's M was 12.01 with significant value of .551, indicating that the groups do not differ significantly from the covariance matrices. So the analysis could be carried on. The overall Wilks's Lambda was significant, $\Lambda = .194$, $\chi^2(6, N = 45) = 67.27$, $p < .0001$, showing that the three groups of children were being successfully differentiated by the three predicators variables (age, D and the index). The residual of Wilks's Lambda was significant as well, $\Lambda = .194$, $\chi^2(6, N = 45) = 67.27$, $p < .0001$, indicating that the predicators could differentiate among the groups after eliminating the effects of the first discriminant function. Since these tests showed statistical significance, interpretation of both functions was carried out.

The resulting discriminant function equation of the first function was (0.204 x age) + (-0.038 x index) + (-0.027 X D) -9.982 and that of the second function was (0.082 x age) + (0.092 x syntactic complexity score) + (0.018 x D) -9.725. The first function was more strongly correlated to age than to the language measures, as its correlations with age, the index and D were .951, -.320 and -.331 respectively. The second function had a stronger language component, with higher contribution from the index (.786) than from D (.215). The correlation of the second function with age was .380.

Then every child's group membership was predicted using the two functions. This leaded to correct classification of 39 out of 45 (86.7%). The classification accuracy in the TD 4-5 group, the TD 5-6 group and the SLI 5-6 group were 92.9%, 81.3% and 86.7% respectively. Six cases were misclassified in total. One child in the TD 4-5 group was misclassified into the SLI 5-6 group; two children out of three in the TD 5-6 group were misclassified into the SLI 5-6 group whereas the remaining one was misclassified into the TD 4-5 group; two children in the SLI 5-6 group were misclassified into the TD 5-6 group.

To validate the discriminant function, "leave-one-out classification" was done for reclassification of the same cases. A slightly different result was produced: 38 out of 45 (84.4%). One more case was misclassified. One more children from the TD 5-6 group was misclassified into the TD 4-5 group.

Given that seven cases were being misclassified after validation, the clinical usefulness of the composite measure was being evaluated. These evaluation procedures went beyond the early work. All the misclassified cases which belonged to the two normal groups originally were regarded as true negative cases, whereas those being misclassified in the SLI group originally were regarded as false negative cases. The composite measure's sensitivity was 83.3%; specificity was 86.7%; positive likelihood ratio was 5.2 whereas negative likelihood was 0.16.

Syntactic Errors Analysis

The final results to be reported are the syntactic error analysis. The number of tokens and the percentage of occurrence of each error type in the TD 5-6 group and the SLI 5-6 group were presented in Table 3.

Table 3. No. of tokens (% of occurrence) of each error type in the TD 5-6 group and the SLI 5-6 group (N=31)

	TD 5-6	**SLI 5-6**
Addition	4 (16%)	12 (21.1%)
Substitution	14 (56%)	23 (40.4%)
Omission	6 (24%)	17 (29.8%)
Error in word order	1 (4%)	5 (8.8%)
Total number of errors	**25**	**57**

Among all these error types, substitution was found to be dominating in both groups. Further analysis revealed that substitution of classifiers was dominant in both groups. So the content of these classifier substitutes was further analyzed and presented in Table 4.

Table 4. No. of tokens (% of occurrence) of each classifier substitute in the TD 5-6 group and the SLI 5-6 group (N=31).

Substitute content	**TD 5-6**	**SLI 5-6**
Default *go3*	**5 (71.4%)**	**7 (70%)**
Other classifier	**2 (28.6%)**	**3 (30%)**
- *zek3*	1	3
- *gaa3*	1	0

According to the percentage of occurrence, the TD 5-6 group and the SLI 5-6 group used the default *go3* and other classifiers as substitute with more or less the same frequency respectively.

Discussion

This study aimed to investigate the age-sensitivity of the index of grammatical development, its validation against other language measure (MLU), its potential to differentiate the SLI children from their normal peers, its accuracy in giving a diagnostic label to each child and the error pattern of the SLI children and their age peers. The results were interpreted so as to address these research questions accordingly.

Age Sensitivity of the Index of grammatical development

The primary aim of this study was to examine whether the index of grammatical development was sensitive to age or not. The goal was met, as the TD 5-6 group scored significantly higher than the TD 4-5 group in the measure. The results were supportive to the previous hypothesis that there was a quantitative difference between the two normal age groups in syntactic development. The TD 5-6 group scored higher than the TD 4-5 group in all the subtests but statistical significance was only found in the subtest Q (Question). The following discussion will investigate the contributing factors to the quantitative difference observed in both the measure as a whole and each discrete subtest.

Let us recall that the index developed in this study analyzed the children's ability in noun phrase expansion (NP), verb phrase expansion (VP), constructing specific sentence structure (SS) and question (Q), and employing connectives and sentence adverb (CON/SA). Items in these subtests include the utilization of closed class words, open class words and construction of specific sentence forms.

In the two subtests NP and VP, some of the items involved employment of closed class words, such as classifier, modal auxiliary and aspect marker to elaborate the meaning of noun or verb phrases respectively (Matthews & Yip, 1994). In the subtests Q and CON/SA, question words and connectives, which are close class words as well, were used to request information and connect sentences or ideas respectively. Previous study has shown that five-year-old children could use more varieties of close class than four-year-old children. For instance, in the study by Ma (2006), the five -year-old children utilized more varieties of connectives when compared with the four-year-old peers in spontaneous speech. And the children of the TD 5-6 group in this study did use more types of connectives than their younger normal peers. Examples of connectives used by the two groups were summarized in Table 5.

Table 5. Examples of connectives in English translation used by the TD 4-5 group and the TD 5-6 group

		TD 4-5	TD 5-6
Categories of Connectives	Coordinative/Additive	*tung4maai4*	*tung4maai4,* ***jyut6…jyut6***
	Causal	*jan1wai6, so2ji5*	*jan1wai6, so2ji5*
	Concessive	*bat1gwo3, daan6hai6,* ***seoi1jin4***	*bat1gwo3, daan6hai6,* ***kei3 sat6***
	Conditional	*jyu4gwo2*	*jyu4gwo2*

	Temporal		***sau2sin1, zung1jyu1***
Note: Connectives being underlined were the unique productions of the corresponding group			

The TD 5-6 group used temporal connectives while the TD 4-5 group did not. The TD 5-6 group could produce more types of connectives in the category of coordinative. The findings echoed the study by Ma (2006) that types of connective production increased by age.

Apart from connectives, the TD 5-6 group produced more types and tokes of simple and compound directional verbal complement than the TD 4-5 group in this study. The older group could use six and ten simple and compound directional verbal complements (DVC) respectively, such as "*dit3 lok6 lei4*", "*juk1 lei4 juk1 heoi3", "ling1 hoi1 nei1 gau6 je5", "pou5 hei2 nei5",* whereas the younger group could only use five and eight simple and compound DVCs respectively with smaller tokens. Recall that two examples have to be expressed in two different forms for crediting two points in the measure (lexical criterion), and that each items in the measure could be fully credited with three examples. Being able to use more different types of closed class, the TD 5-6 group could obtain a higher score in the measure.

Even if the diversity of certain closed class was comparable between the two groups, the TD 5-6 group could use them more frequently and skillfully than the TD 4-5 group. For instance, in the study by Fletcher et al. (2000), the 4-year-old children produced more or less the same number of types of question words compared with the 5-year-old group. Both groups could produce what-, where-, who-, why- and how-question. However, in this study, significant group difference was found between these two age groups in the subtest Q. Let us recall there are three categories in the subtest Q. Producing advanced questions such as why-, how- and when- question (Q-Wh-2) could receive one more credit than what-, where-, who-, which- (Q-Wh-1) and yes/no question (Q-Y/N). Further examinations showed that the TD 5-6

group as a whole raised questions more frequently than the TD 4-5 group. Moreover, the older group produced nearly double the number of Q-Wh-2 than the TD 4-5 group (TD 5-6: 17 times; TD 4-5: 10 times). So the significant group difference between them could be attributed to the higher frequency of all question words and advanced question words employment. Or the TD 4-5 group was less skillful and produced more errors in the questions than the older peers. Erroneous productions could not be credited.

Apart from using the more diversified types of closed class words, the increase in the types of nouns and verbs, which are open class words, has also attributed to the growth of the measure from the younger normal group to the older normal group. The study by Fletcher et al. (2000) found that the five-year-old children could use more types of nouns, transitive and intransitive verb than those in the four-year-old group. Recall that two examples should be in distinct surrounding contexts in order to be credited twice (content criterion). For instance, the two examples "*hoi1 zo2 dou6 mun4"* and *"hoi1 gan2 dou6 mun4"* could not be credited twice for the use of the aspect markers (underlined with dotted line) since both the aspect marker expand the meaning of the verb "*hoi1*", i.e. with the same contexts. Being able to use more different types of nouns and verbs, the five-year-old normal group can produce exemplars with higher chance to be credited in subtest NP and VP.

Validation of the index using MLU

Another purpose of this study is to validate the index of syntactic development using MLU. The correlation between them was moderate (r=.502) with statistical significance. The correlation strength is weaker than those between grammatical productivity indices in English and MLU, such as IPSyn in Scarborough (1990) (r=.938); and the syntactic complexity measure (r=.88) in Blake et al. (1993). The moderate correlation indicated that utterance length may not always be valid in estimating the index. For instance, a child in the TD 4-5 group with MLU of 5.6 scored 36 in the index of syntactic development, while another one in

the TD 5-6 group with similar MLU (=5.62) scored 62. This phenomenon echoed the findings of the previous studies that the sensitivity of MLU to grammatical development will decrease after a mean length of roughly 3.0 is reached (Scarborough, Rescorla, Tager-Flusberg, Fowler & Sudhalter, 1991; Scarborough, 1990; Rondal, Ghiotto, Bredart & Bachelet, 1987). In the studies by Scarborough et al. (1991) and Scarborough (1990), the correlations between MLU and IPSyn were always lower for higher MLU.

And since the MLU and age correlation is also lower in Cantonese sample (r=.44) (Klee et al., 2004) than in English samples (r=.66) (Klee et al., 2007), MLU might not be a good index of language growth, especially with the older age group. In the study by Klee et al. (2004), there were two instances where MLU declined as age increased (from 35.5 to 41.7 months, from 60.2 to 65.7 months). Such decline was found in this study as well, from the TD 4-5 group and the TD 5-6 group. This echoed the conclusion in the study by Scarborough et al. (1991) that MLU is a less sensitive language measure than IPSyn.

Therefore, using MLU to explain and predict grammatical complexity may not be valid and adequate enough for children with MLU longer than 3. However, it still can be a useful validation tool for language measure provided that children are at early language developmental stage.

Performance of the 5-year-old normal and SLI group on the measure

Given that the index could show grammatical development in the normal groups, the group difference found between the SLI 5-6 group and their age-matched peers is of interest. The results indicated that the index could differentiate the two groups from each other and all the subtests could accomplish the task as well with statistical significance. The results are not surprising. It has been mentioned previously that lots of studies in Cantonese showed that the morphosyntatic system of SLI children was not as well-developed as their normal age peers. Since it is more important on the clinical perspective to find out if an index can make accurate

diagnosis than if the groups performed differently on it (Wong et al., 2010), the discussion will proceed to the diagnostic accuracy of the linear combination of age, D and the index.

Diagnostic accuracy of the linear combination of age, D and the index

Recall that the discriminant analysis correctly classified 25 of out 30 cases in the two normal groups and 13 out of 15 in the SLI group. One less child in the normal group was correctly classified in the "leave-one-out" validation procedure. The sensitivity and specificity of the index were 83.3% and 86.7% respectively. Sensitivity and specificity levels of 90% and above are regarded as good and 80% as fair (Plante & Vance, 1994). Upon these two indices, it can be concluded that the diagnostic accuracy of the composite measure is fair.

To evaluate the clinical usefulness of the composite measure as a diagnostic or screening tool, the positive and negative likelihood ratio (LR+ and LR-) should be considered. The LR+ and LR- of the composite measure were 5.2 and 0.16 respectively. According to Dollaghan (2007), a clinically-useful instrument should possess a LR+ greater than 10 and a LR- smaller than 0.1. So the test cannot achieve such purpose.

Although the clinical usefulness of the composite measure could not be confirmed from LR+ and LR-, it is yet a potential clinical marker of SLI since its sensitivity and specificity levels were fair. Replication of the study in an independent sample will be necessary to confirm the findings.

Error pattern

The last research question to be addressed is the comparison of error tokens between the TD 5-6 group and the SLI 5-6 group. As a group, the SLI group produced more errors than the normal group during conversation. This confirms the result of a study by Dunn et al. (1996) that SLI children produced more syntactic errors in spontaneous speech than normal children. Classifier substitution error will be the principal discussion area in this session as it was found to be dominating in both groups. Both groups used a similar proportion of default *go3* (TD 5-

6: 71.4%; SLI 5-6: 70%) and other classifiers as substitute (TD 5-6: 28.6%; SLI 5-6: 30%) of all their errors. However, it should be noted that only one child in the SLI group used other classifiers as substitutes and that child used *zek3* in all three occasions of substitution. So it is quite possible that the SLI child was using *zek3* as his own default substitute. Therefore, it can be concluded that the SLI children used general classifiers as substitute more frequently than the normal group. The results did not support the findings of the study by Stokes et al. (1997), which concluded that language-impaired children as a group used other classifier substitutes more frequently than the normal age-matched peers. The discrepancy may be due to the different methods adopted by the two studies. The study by Stoke et al. (1997) elicited children's production of classifiers using questions or sentence completion, while errors of classifier use were noted during conversation in the present study. If children had to assemble their own sentences during conversation, they behaved differently in two studies.

Future Modification

Given that the index was able to illustrate developmental difference between two age groups while such difference could not be observed in most of the subtests, it should be modified to develop an index of syntactic development with higher sensitivity to grammatical complexity. The modification can be done on the selection and categorization of items, crediting system of the index and implementing the index on more participants.

In terms of selection, certain grammatical items may not be good indicators of syntactic development as both normal groups used them with equal abilities. These items were not eliminated in the pilot study as previous studies have shown that the older group either produced more types of that item or produced that item more frequently than the younger group. For instance, in the study by Fletcher et al. (2000), the five-year-old group produced similar number of types of aspect markers as the four-year-old group but the older group used them more frequently than the younger group. So this item was kept in the index. However,

developmental trend could not be observed in the aspect marker usage in this study. Similar to the study by Fletcher et al. (2000), both the TD 4-5 and the TD 5-6 groups could use the same number of types of aspect marker (*zo2, gan2, zyu6, gwo3* and *haa5*). But unlike the previous study, the two groups produce comparable tokens of the item (TD 4-5: 49 tokens; TD 5-6: 48 tokens). No subtle grammatical development can be shown by this item. It should be eliminated if the study is to be replicated on children of the same age groups.

With regard to the further modification of the crediting system, more items should have adopted a crediting hierarchy. For some of the items in the subtests, such as aspect marker, modal auxiliary and question words, more credits were given to the exemplars which were known to emerge later in the developmental milestone. For instance, usage of question word "*dim2gaai2*" (why) received 1 more credits than question word "*mat1je5*" (what). But such system was not applied throughout the index since the pilot study did not cover every sample employed in this study. Therefore, certain developmental trends could not be identified. After every language sample had been analyzed, it was found that the crediting hierarchy should have been adopted in connectives, classifiers and prepositional phrase as well. For instance, referring to Table 5, temporal connectives were only used by the 5-year-old children. In the study by Ma (2006), the five-year-old children used more temporal connectives than the four-year-old group and they first emerged when a child was 4-year-old. These findings justified that more credits should be given to temporal connectives, or less credits should be given to other common connectives equally used by these two groups, such as *jan1wai6* (because), *tung4maai4* (and). For prepositional phrase, it was found that the four-year-old group received credits from this item using *hai2-* and *tong4-* construction only. However, the five-year-old group was using more varieties of prepositional phrase such as *zoeng1-*, *dang2-* and *caang1-* construction. Again, more credits should be given to the late emerging items as well.

Clinical implication

Though the composite measure is not a clinically-useful instrument according to the value of the LR+ and LR-, the fact that the diagnostic accuracy of the measure is fair gives insight into the development of assessment procedures of the SLI children. Spontaneous speech samples may supplement the diagnosis of SLI children and compensate the drawbacks of psychometric tests. There have been lots of controversies regarding the reliance on psychometric tests. These tests may tend to emphasize a single aspect of language (American Speech-Language-Hearing Association, 1989). For instance, the most complex language form being assessed in the RDLS-R is multiple-step command embedded with locative and colour concepts. No narrative element is included. For the RDLS-E, the most complex production required is at sentence level which focuses on the semantic content of sentences and their form. Any cohesiveness in daily dialogue and other uses of the language (e.g. request) are not assessed.

To ascertain the value of language sample analysis in assessing preschoolers' language ability, the above-mentioned improvements of the index should be take into account and replication of the study can help. The listing of structures that makes up the index provides a framework for examining language development for descriptive purpose over time. Once can examine if a child has made progress in using a more complex form of an item, or using a form more productively, during conversations or elicited productions.

Acknowledgement

I would like to thank my adviser, Dr. Anita Wong for her precious time and invaluable advice on this research project. This project will be impossible to be finished without her guidance and encouragement. My heartfelt appreciations go to Miss Feiona Man Ka Yee as well for her contribution of time to be the inter-rater in this study. Finally, I would like to thank my fellow classmates for their words of emotional support over the course of my study.

References:

American Speech-Language-Hearing Association. (1989). Issues in determining eligibility for language intervention. *Asha, 31,* 113-118.

Brown, R. (1973). *A first language: the early stages.* Cambridge, MA: Harvard University Press.

Blake, J., Quartaro, G. & Onorati, S. (1993). Evaluating quantitative measures of grammatical complexity in spontaneous speech samples. *Journal of Child Language, 20*, 139-152.

Cheung, S. L. (1991). The notion of "Result" in Cantonese Children. *Papers and Reports on Child Language Development, 30,* 17-24.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences.* (2nd ed.). Hillsdale, N.J.: L. Erlbaum Associates.

Costanza-Smith, A. (2010). The clinical utility of language samples. *Perspectives on language learning and education, 17*(1), 9-15.

Crais, E. R. (1995). Expanding the repertoire of tools and techniques for assessing the communication skills of infants and toddlers. *American Journal of Speech-Language Pathology, 4*, 47-59.

Dollaghan, C. A. (2007). *The handbook for evidence-based practice in communication disorders*. Baltimore, MD: Brookes.

Dunn, M., Flax, J., & Sliwinki, M. (1996). The use of spontaneous language measures as criteria for in identifying children with Specific Language Impairment: an attempt to reconcile clinical and research incongruence. *Journal of Speech and Hearing Research, 39,* 643 – 654.

Fletcher, P., Leonard, L. B., Stokes, S. F., & Wong, A. M.-Y. (2005). The expression of aspect marker in Cantonese-Speaking children with Specific Language Impairment. *Journal of Speech, Hearing and Language Research, 48,* 621-634

Fletcher, P., Leung, C. S., Stokes, S., & Weizman, Z. (2000). *Cantonese pre-school language development: A guide.* Hong Kong: University of Hong Kong, Department of Speech and Hearing Science.

Hong Kong Society for Child Health and Development (1987). *Manual of the Reynell Developmental Language Scales, Cantonese (Hong Kong) version.* Hong Kong-SAR-China; Author.

Klee, T., Gavin, W. & Stokes, S. (2007). Utterance length and lexical diversity in American- and British-English speaking children: What is the evidence for a clinical marker of SLI. In R. Paul (Ed.), *Language disorders from a developmental perspective: Essays in Honor of Robin S. Chapman.* Mahwah, N. J.: London: Lawrence Erlbaum Associates, Inc.

Klee, T., Scahffer, M., May, S., Membrino, I., & Mougey, K. (1989). A comparison of the age-MLU relation in normal and specifically language-impaired preschool children. *Journal of Speech and Hearing Disorders, 54,* 226-233.

Klee, T., Stokes, S. F., Wong, A. M.-Y., Fletcher, P. & Gavin, W. J. (2004). Utterance Length and Lexical Diversity in Cantonese-Speaking Children With and Without Specific Language Impairment. *Journal of Speech, Language and Hearing Research, 47,* 1396-1410.

Lahey, M., & Bloom, L. (1994). Variability and language learning disabilities. In K. G. Butler (Ed.), *Language learning and disabilities in school-age children & adolescents: Some principles and applications.* Needlham Heights, MA: Allyn & Bacon.

Leonard, B. L., Wong, A. M.-Y., Deevy, P., Stokes, S., & Fletcher, P. (2005). The production of passives by children with SLI acquiring English or Cantonese. *Applied Linguistics, 27,* 267-299.

Leung, C. S. (1995). *The development of aspect markers in a Cantonese-speaking child between the ages of 21 and 45 months.* Unpublished doctoral dissertation, University of Hawaii, Manoa.

Matthews, S. & Yip, V. (1994). *Cantonese: A Comprehensive Grammar.* London: Routledge.

Miller, J. F. (1981). *Assessing language production in children: Experimental procedures.* Baltimore: University Park Press.

Miller, J. F. & Chapman, R. S. (1981). The relation between age and mean length of utterance in morphemes. *Journal of Speech and Hearing Research, 24,* 154-161.

Miller, J. F. (2005). *Language Sample Analysis: The Wisconsin Guide.* Madison, Wisconsin: Wisconsin Dept. of Public Instruction.

Owen, A. J., & Leonard, L. B. (2002). Lexical diversity in the spontaneous speech of children with specific language impairment: Application of D. *Journal of Speech, Language and Hearing Research, 45,* 927-937.

Paul, R. (2007). *Language Disorder From Infancy Through Adolescence: Assessment and Intervention.* (3rd ed.). St. Louis, Missouri: Mosby Elsevier.

Plante, E. & Vance, R. (1994). Selection of preschool language tests: A data-based approach. *Language, Speech, and Hearing Services in Schools, 25,* 15-24.

Richards, B. J., & Malvern, D. D. (1997). *The New Bulmershe Papers: Quantifying lexical diversity in the study of language development.* Reading, England: University of Reading.

Richards, B. J., & Malvern, D. D. (1998). A new research tool: mathematical modeling in the measurement of vocabulary diversity (Award reference no. R000221995). Final Report to the Economic and Social Research Council, Swindon, UK.

Rondal, J. A., Ghiotto, M., Bredart, S., & Bachelet, J. (1987). Age-relation, reliability and grammatical validity of measures of utterance length. *Journal of Child Language, 14*, 433-446.

Scarborough, H. S. (1990). Index of Productive Syntax. *Applied Psycholinguistics, 11*, 1-22.

Scarborough, H. S., Rescorla, L., Tager-Flusberg, H., Fowler, A. & Sudhalter, V. (1991). The relation of utterance length to grammatical complexity in normal and language-disordered group. *Applied Psycholinguistics, 12*, 23-45.

Stokes, S. F. & Fletcher, P. (2000). Aspectual forms in Cantonese children with specific language impairment. *Linguistics, 41*(2), 381-405.

Stokes, S. F. & So, L. K.-H. (1997). Classifier use by language-disordered and age-matched Cantonese-speaking children. *Asia Pacific Journal of Speech, Language and Hearing, 2*, 83-101.

To, C. K.-S., Stokes, S. F., Cheung, H.-T., & T'sou, B. (2010). Narrative assessment for Cantonese-Speaking Children. *Journal of Speech, Language and Hearing Research, 53*, 648-669.

Tse, S.-K., Hui, L. & Leung, S.-O. (2007). The acquisition of Cantonese classifiers by preschool children in Hong Kong. *Journal of Child Language, 34*, 495-517.

Wong, A. M.-Y., Klee, T., Stokes, S. F., Fletcher, P. & Leonard, L. B. (2010). Differentiating preschool children with and without SLI using MLU and lexical diversity (D). *Journal of Speech, Language, and Hearing Research, 53*, 794-799.

Wong, A. M.-Y., Leonard, L. B., Fletcher, P., & Stokes, S. F. (2004). Questions without movement: A study of Cantonese-speaking children with and without specific language impairment. *Journal of Speech, Language, and Hearing Research, 47*, 1440-1453.

Zhang, H. N. (2007). *A Grammar of Cantonese as Spoken in Hong Kong.* Hong Kong: The Chinese University of Hong Kong.

Appendix A

Examples and definitions of the syntactic structures examined in this study

The definitions and some of the examples of the items were from To et al., (2010), Zhang, (2007), Leung, (1995), Matthews & Yip, (1994).

A. Noun Phrase Expansion (NP)

Definition: The meaning of a noun phrase can be modified and expanded by demonstrative, articles, adjectival modifier (including possessives and quantifiers) and relative clause. In this study, only structures exist in Cantonese linguistics and start emerging in the developmental stages of participants will be credited. Their definition and examples are as follows:

(NP-CL) Classifier *(loeng4 ci4)*

Definition: Each noun has a specific assigned classifier. There are mainly two types of classifiers: **measural** classifiers and **sortal** classifiers. The former specifies quantity of the head noun while the latter denotes the intrinsic quantity (e.g. shape) of the object.

Examples: *yat1 baan1 yan4* (measural classifiers)

One CL people (a group of people)

sam1 tiu4 jy2 (sortal classifiers)

Three CL fish (Three fish)

(NP-A) Adjectival modifier

Definition: They are adjectives which proceed and modify meaning of a noun phrase.

Example: *daai6 fei1 kei1*

Large airplane (large airplane)

(NP-RC) Relative clause

Definition: Relative clause is **predicative clause** elaborating a noun phrase. It always serves to restrict the identity of a noun phrase. It is ahead of the noun phrase and can be connected to it with linking particle *ge3* (LP) or *go2 go3.*

Example: *ngo5dei6 soeng5 tong4 go2zan6si4 waan2 caai2 daan1ce1 gaa3*

We have lesson when play ride bike PRT

(We rode on the bike when we were having lesson)

B. Verb Phrase Expansion (VP)

Definition: According to Matthews and Yip (1994), modal auxiliary, predicate adverb, manner predicate (*gum2*-construction) can serve the same function as well. Clausal complement can modify the meaning of perceptual and mental verbs (To et al, 2010).

Verbal Complement *(wai1bou2)*

Definition: The former is an action or description which is described by the latter complement. The complement can be a verb or an adjective. A verb particle may be present.

Different types of verbal complement

(VP-RVC) Resultant complement

Definition: The second predicate indicates the result of the first predicate.

Example: *sai3 lo2 zing2 laan6 sai3 di1 syu1*

Younger brother make broken all PRT book

(Younger brother has broken all the books)

(VP-DVC) Directional complement

Definition: Simple and compound direction complement can be assembled using these items: *lei4* (come) , *heoi2* (go) *soeng5* (up), *lok6* (down), *hoi1* (away), *mai4* (close), *ceot1* (out), *jap6* (in), *gwo3* (over), *hei2* (up), *dou6* (here), *faan1*(back)

Example: Simple direction complement

han4 *seong5* *lok6 lau2*

walk up six floor

(walk up to the sixth floor)

Compound direction complement

cung1 *lok3 heoi6* *wing6 ci4*

rush down come pool

(Rush into the swimming pool)

(VP-OVC)Other verbal complement

I. Phase complement

Definition: The second predicate indicate the special status of the first predicate. The second predicates are always *zoek6, dou2, hei2, gwo3*

Example: *keoi5* *fan3* *zoek3* *zo2*

He/she sleep finish ASP

(He/she is asleep)

II. Intensifying complement

Definition: An adverb is added before a stative verb to intensify the status of that adjective/verb

Example: *nau* *sei2* nei5

angry die (as intensifier) you

(Very angry with you)

III. Potential complement

Definition: It describes the possibility of an event or action. An infix is inserted into the complement to express the possibility. The adverb '*dak1*' indicates positivity while '*m4*'

indicates negativity.

Example: *faan3 m4 zoek3*

Sleep NEG finish (Cannot be asleep)

sik6 dak1 baau2

eat ADV full ((Somebody is) full)

IV. Descriptive complement

Definition: Its position is less restricted that it does not necessarily follow the predicate immediately.

Example: *keoi5 ceong3 dak1 hou2 hou2 teng1 aa3*

He/she sing ADV good good listen PRT

(He/she sings very well.)

V. Complement of Extent

Definition: It expresses the extent of the action in the predicate using the special structure "predicate + *dou3*"

Example: *siu3 dou3 lok1 dei2*

laugh V-PRT trundle floor

(Somebody laughs to the extent that he/she trundle along the floor)

(VP-AS) Aspect Markers

Definition: Aspect marker indicates the duration of time for which an event lasts while does not specify the exact time that the event happens. Aspect markers in category VP-AS-1 appear before 3 years and 3 months old (Leung, 1995).Their emergence is earlier than those in category VP-AS-2.

Different types of aspect marker:

(VP-AS-1)

I. Perfective marker *zo2*

II. Imperfective marker *gan2, zyu6*

III. Experiential marker *gwo3*

(VP-AS-2)

IV. Habitual *hoi1*

V. Inchoative *hei2 (seong6) lai4*

VI. Delimitative *haa5*

Example: *tai2 gan2 din6si6*

Watch ASP TV

((Somebody is) watching TV)

(VP-VA) Verbal Adverbs

Definition: Verbal adverbs are used to elaborate and describe ***verbs***, with respect to their ***extent, manner, tone, time and effective area*** etc. So their linguistic functions are very limited.

Example: *m4 tung1* *yat1 zaai1* *fan3* *me1*

Is-it-possible together sleep PRT

(Is it possible to sleep together?)

(VP-MA) Modal Auxiliary

Definition: Modal verbs are used to express the extent of ***possibility, ability, wishes, obligation*** and other related concepts. They appear in front of a verb phrase.

(VP-MA-1)

I. *wui3* (will/would)

II. *ho2 ji5* (can, may)

III. *yiu3* (want)

IV. *seong2* (wish to)

(VP-MA-2)

I. *zeon2* (allow)

II. *jing1 goi1* (should , ought to)

III. *m4 sai2* (no need)

IV. *ling4 jyun4* (rather)

V. *sik1* (know)

VI. *bei2* (let)

VII. *zung1ji3* (love, like).

Example: *ngo5* *soeng2* *fan3 kaau3*

I want sleep

(I want to sleep)

m *zeon2* *nei5* *waan2*

not allow you play

((I) Don't let you play)

(VP-CC) Clausal Complement

Definition: They are ***predicative or verbal clausal complement*** appearing ***after perceptive or mental verbs.*** Clauses following verbs such as *hoi1 ci2* (begin), *zeon2 bei6* (get ready), *gai3 zuk6* (continue) are also regarded as clausal complement.

Example: *ngo5* *zeoi3* *geng1* *sik6 laak6 jyu2*

I most scared eat spicy fish

(I am most scared of spicy fish)

(VP-MP) Manner predicate

Definition: It ***appears before a verb*** to indicate or explain the manner of the action. It can be indicated and preceded by adverbial *gum2* or *gum2 jyn2.*

Example: *koei2 ho2 daai6 lek6 gum2 daa2 ngo5*

He/she very big force thus hit me

(He/she hits me hard.)

ngo5 dei6 jat1 cai4 co5 aa1

We together sit PRT

(We sit together.)

C. *Sentence Structure (SS)*

(SS-SV) Serial Verb Construction *(lin4wai6)*

Definition: The subject is connected to two or more verbs without any intervening conjunction in between them.

Example: *ze4 ze1 caak6 a4 sai2 min6*

Elder sister brush teeth wash face

(Elder sister brushes teeth and washes face.)

(SS-PC) Pivotal construction *(gim1jyu5)*

Definition: The pivotal noun phrase (the underlined NP in the examples) serves as the object of the former verb and the subject of the latter verb simultaneously.

Example: *dong3 zek3 sau2 hai6 jan4 aa3*

pretend CL hand is person PRT

(Pretend that the hand is a person.)

(SS-PP)Prepositional phrase

Definition: Prepositional phrase usually appears before the verb(dotted lined) (there are exceptions such as *hai2*-construction). The phrase comprises a preposition and a noun phrase. (The verb is dotted underlined, the preposition is doubly underlined and the noun phrase is singly underlined in the examples)

I. *bei2*-passive construction

dai4 dai2 *bei2* *ze4 ze1* *cou4 seng2 zo2*

younger brother by elder sister noise wake ASP

(The younger brother is waken by the noise made by the elder sister)

II. *zeong1*-construction

zeong1 *syu4 zai2* *fong3 lok6 heoi3 zyu2*

place potato put down go cook

(To put the potato into something and cook)

III. *deoi3*-construction

zek6 gau2 *deoi3 zyu6* *go3 caak6* *fai6*

CL dog face ASP CL thief bark

(The dog bark face-to-face to the thief.)

IV. *lin4*-construction

lin4 *sai3 lo2* *dou1 gok3 dak1 m4 gau3 baau2*

Even younger brother also feel NEG enough full

(Even the younger brother is not full)

V. *hai2*-construction

zek3 gau2 *hai2* *toi2 dai2 dou6 fan3 gaau3*

CL dog at table bottom sleep

(A dog sleeps under the table.)

VI. *tong4*-construction

tong4 *keoi5 dei2* *waan4*

with they play

(Play with them.)

D. Question

(Q-YN) Yes/No question

Definition: There are seven types of yes/ no questions. They are particle questions, A-not-A questions, copula questions, *mei6* questions, existential questions, tag question and echo question. Only those used in the samples will be listed. Their definitions and examples are as follows:

I. Particle question

Definition: It is the simplest question form by addition of a particle to a declarative sentence.

Example: *mat1 nei5 m4 zi1 gaa3 me1?*

What you NEG know PRT PRT

(Don't you know?)

II. A-not-A questions

Definition: The construction of this question involves a negative morpheme *m4* in between a repetition of the verb or adjective

Example: *nei5 zung1-m4-zung- ji3 sik6 tong aa3?*

You like-not-like eat sweet PRT

(Do you like sweet?)

III. Copula questions

Definition: Copula question is signified by copula verb *hai6 mai6* which precedes the

man verb. *hai6 mai6* is a kind of explicit A-not-A

Example: *nei5 hai6-mai6 yiu3 aa3?*

You be-not-be want PRT

(Do you want it?)

IV. *mei6*-questions

Definition: Addition of the negative morpheme *mei6* to a declarative sentence can form a question about whether an event has occurred or not.

Example: *nei3 sik6 zo2 faan6 mei6 aa3?*

You eat ASP food not-yet PRT

(Have you had your meal yet?)

V. Existential questions

Definition: Existential question uses the words *yau5 mou5* to ask for existence of an object.

Example: *le1 dou6 yau5 mou5 ce1 ce1 aa3?*

PRT here have-not-have car car PRT

(Is there any car here?)

Wh-question

Definition: The wh-word appears in similar order as in an English question if it is a subject question. However, the wh-word acts as an object and appears after the verb like a direct object in an object question.

Subject question

Bin1go3 wan2 ngo5 a3?

Who seek me PRT

(Who is looking for me?)

Object question

nei2 wan2 bin1go3 a3

you seek who PRT

(Who are you looking for?)

(Q-Wh-1) Question word Type 1

Definition: The wh-question words in this category emerge from 24 to 30 months old.

I. 'Who' question – *bin1go3*

Example: *Bin1go3 daa2 din6waa2 lei4 a4?*

who call-phone come PRT

(Who was it that called?)

II. 'What' question – *mat1ye5*

Example: *Sik6 di1 mat1ye5 ho2 a?*

Eat SL what good PRT

(What shall we eat?)

III. 'Which' question – *bin1* + CL + noun

Example: *Nei5 gok3dak1 bin1 zek6 gau2 zui3 leng3 a4?*

You feel which CL dog most nice PRT

(Which dog you think is the nicest?)

IV. 'Where' question – *bin1dou6/ bin1syu3*

Example: *ngo5 dei6 heoi3 bin1dou6 waan2 a4?*

We go where play PRT

(Where shall we go for dim sum?)

(Q-Wh-2) Question word Type 2

Definition: The wh-question words in this category emerge from 30 to 48 months old.

I. 'When' and time question

Example: *nei2 gei2si4 faan1 lei4 a4?*

You when return come PRT

(When will you be back?)

II. 'Why' and reason question

Example: *dim2gaai2 wui2 gam2 ge?*

How come would thus PRT

(Why is it like this?)

III. 'How' question

Example: *nei2 dim2 yeung2 jing2 saa1loet2 gaa3?*

You how way make salad PRT

(How do you make salad?)

E. Sentence adverbs and connectives

Sentence adverbs (SA)

Definition: Sentence adverbs are distinct from verbal adverbs in the way that the former ***modify the meaning of the whole sentence*** instead of the verb phrase or the predicate only. Their position are less restricted than verbal adverb in the way that they can be ***positioned (a) between the subject/topic and the verb; (b) at the beginning of the sentence; (c) at the end of the sentence, by way of afterthought.***

Example: *ngo5 waak6 ze2 m4 lei4 dak1*

I perhaps no come able

(I may not come)

Sentence adverbs can be classified according to their functions. The functions include ***time*** (present), such as *yi1 gaa1* (now); time (future), such as *haa6 ci3* (next time); ***frequency,*** such as *do1 sou2* (by that time), *gan1 zyu6* (next); ***quantity***, such as *zing6 hai6* (only); ***presupposition,*** such as *dim2 zi1* (out of expectation)

Connectives/Conjunctions (CON)

Definition: There are two major categories of conjunction. The first one is used to **connect two words or even larger linguistics units**, such as phrases or clauses. They are used to reveal the relations between the two units. The second one is used to **connect idea other than the main clause of the speaker**. It can be used to connect idea of the previous speaker. There are different types of connectives classified into ***coordinative, clausal, concessive, hypothetical and temporal.***

Example: *maa4 maa1 tung1 ze4 ze1 dou1 heoi3 gaai1*

Mother and sister also go street

(Mother and sister also go out.)

Daan6 hai6 yiu3 maai5 faan1 lei4 lo1

But need buy back come PRT

(But (we) have to buy it back.) (A concessive connectives)

An index of syntactic development – Charting Sheet

Items (credit/s)	**Description**	**Exemplar1**	**Exemplar2**	**Exemplar 3**	
Noun phrase elaboration					**/12**
NP-CL(1)	Classifier				
NP-A(1)	Adjectival modifier				
NP-RC(2)	Relative Clause				
Verb phrase elaboration					**/42**
	Verbal complement				
VP-RVC(1)	Resultant Verb Complement				
VP-DVC(1)	Directional Verb Complement				
VP-OVC(1)	Other verbal complement				
	Aspect Marker				
VP-AS-1(1)	Perfective marker: *zo2* Imperfective marker: *gan2, zyu6* Experiential marker: *gwo3*				
VP-AS-2(2)	Habitual: *hoi1* Inchoative: *hei2 (seong6) lai4* Delimitative: *haa5*				
VP-VA(1)	Verbal Adverb				

	Modal Auxiliary				
VP-MA-1(1)	*wui3, ho2 ji5, yiu3, seong2*				
VP-MA-2(2)	*zeon2, jing1 goi, m4 sai2, ling4 jyun4, sik1, bei2, zung1ji3*				
VP-CC(2)	Clausal Complement				
VP-MP(2)	Manner predicate				
Sentence structure					**/18**
SS-PC(2)	Pivotal construction				
SS-SC(2)	Serial verb construction				
SS-PP(2)	Prepositional phrase				
Questions					**/12**
Q-YN(1)	Yes/No question				
	Wh-question				
Q-Wh-1(1)	What, who, where, which				
Q-Wh-2(2)	How, why, when				
Sentence Adverb and connectives					**/12**
SA(2)	Sentence adverbs				
CON(2)	Connectives				

Appendix B

Figure 1. Scatter plot of the index of grammatical development and MLU, for typically developing Cantonese speaking children (N=30)