



Title	Validation of a screening tool for identifying three-year- old Cantonese-speaking children with language impairment
Other Contributor(s)	University of Hong Kong.
Author(s)	Fong, Yik-yan
Citation	
Issued Date	2007
URL	http://hdl.handle.net/10722/55512
Rights	Creative Commons: Attribution 3.0 Hong Kong License

**Validation of a Screening Tool for Identifying Three-Year-Old Cantonese-Speaking
Children with Language Impairment**

Fong Yik Yan

A dissertation submitted in partial fulfillment of the requirements for the Bachelor of Science
(Speech and Hearing Sciences), The University of Hong Kong, June 30, 2007.

Abstract

The accuracy of subscales A and B of the Developmental Language Screening Scale (DLSS) in identifying 36 to 47 month-old Cantonese-Speaking children with language impairment was investigated. In Stage I, the DLSS was distributed to the caregivers of 149 children studying in K1 and 86 (57.72%) caregivers completed and returned the questionnaire. All 12 screened-positive children and 8 randomly selected screened-negative children received a clinical language assessment during Stage II. Results of screening and clinical assessment were compared in a double-blinded fashion. The DLSS was found to have poor accuracy, as revealed by the sensitivity value of 75% and specificity value of 50%. Since the DLSS has a high tendency to overlook children with language impairment and over-refer those without language impairment, recommendations were made for further research to develop a valid and locally applicable screening tool for identifying preschool children with language impairment.

Introduction

Definition

Language impairment is defined by the American Speech-Language-Hearing Association (ASHA) as an impairment in “comprehension and/or use of a spoken, written, and/or other symbol system” (1993, p.40), which may involve deficits in the form, content, and/or function of language during communication. Over the years, terms such as language disorder, language delay, and language deviance have been used to label children with language impairment (Paul, 2001). Based on the definition by Law (1992), language delay is used to describe children whose language development follows a typical pattern, but proceeds at a slower rate than their aged peers. Language disorder, on the other hand, is used to describe children whose language development is disturbed in some important, qualitative way (Law, 1992). In this paper, the term “language impairment” (LI) is used as a first level description to refer to both language delay and disorder.

The Developmental Language Screening Scale

It has come to general consensus that early intervention has a positive impact on children’s language development, behaviour, and later academic performance (Glascoe, 2005). However, the search for children at risk of language difficulties has been hampered by the lack of effective screening measures, such that as many as two thirds of preschool children with developmental disabilities such as language impairment remain undetected by the time they reach school age, by which all opportunities for early intervention have been dissipated (Glascoe, 2005).

Valid and locally applicable language screening measures were lacking in Hong Kong before the development of the Developmental Language Screening Scale (DLSS) (Lee, Luk, Yu, & Bacon-Shone, 1985). Utilizing a parent interview format, the DLSS was originally developed to provide an initial screening of developmental language problems applicable to the Hong Kong population (Lee et al., 1985). The DLSS is divided into six subscales (A-F), investigating verbal comprehension, verbal expression, non-verbal comprehension, non-verbal expression, interest in communication, and abnormalities of speech respectively. In a large-scale normative study (Lee et al., 1990), which served as part of an epidemiological survey on the prevalence of behavioral disorder and language impairment in the Hong Kong preschool population, scores of the A, B, and A+B subscales of the DLSS were normed on a stratified sample of 855 typically developing preschool children (424 females and 431 males) aged between 36 and 48 months. The concurrent validity of the DLSS was also investigated in the same study by Lee et al. (1990). The subscale A receptive language and subscale B expressive language quotients of the DLSS were examined with the receptive and expressive language quotients of the Cantonese version of the Reynell Developmental Language Scale (RDLS-C) (Hong Kong Society for Child Health and Development, 1987) in a correlation analysis. Data from 234 of the participants yielded moderate correlation of 0.45 and 0.40 ($p < 0.001$) for receptive and expressive language respectively.

In a follow-up study by Wong et al. (1992), 226 children (102 females and 124 males) were selected from the sample in the study by Lee et al. (1990). The prevalence of language impairment in the sample of 36 to 48 month-old children was 3.4%, with more males

diagnosed as having language delay than females (Wong et al., 1992). Although correlations of the DLSS and the RDLSS-C scores were reported, information on the accuracy of the DLSS as a screening measure for each of the participants to which it was administered was lacking. The DLSS has not yet been validated against an objective gold standard.

Previous studies of the DLSS utilized a discrepancy-based criterion for the diagnosis of language impairment (Lee et al., 1990; Wong et al., 1992), defining language impairment as “language age on the Verbal Comprehension Scale and the Verbal Expression Scale is equal or less than two-thirds of the chronological age” (Wong et al., 1992, p.252). Based on such a definition, children’s performances are compared to those of children at an age in which the raw score is typical, rather than comparing it to those of children at the same age (Paul, 2001). Thus, utilization of age-equivalent scores introduces the potential problem of not being able to identify the presence of a significant difference in the performance of a particular child, since no measure of normal variation is included (Paul, 2001). With such considerations, the present study utilizes the adaptive dysfunction criterion, shown by significant adaptive consequences, and the standardized criterion, evidenced by performance significantly below the norm on standardized tests, for diagnosis of language impairment (American Psychiatric Association, 1994; DSM-IV; ICD-10; World Health Organization, 1992).

Parent-Report Measures

The use of parent-report measures is being increasingly recognized as a valid and reliable method for identifying children with language impairment (e.g. Feldman et al., 2005; Klee et al., 1998; Klee, Pearce, & Carson, 2000; Stott, Merricks, Bolton, & Goodyer, 2002).

One of the most important advantages of parent reports of children's speech and language ability is that they can serve as a comprehensive and representative source of information on children's abilities to communicate in a wide variety of naturalistic situations outside the clinical setting (Dale, 1996; Feldman et al., 2005; Thal, O'Hanlon, Clemmons, & Fralin, 1999). The extensive experience parents have with their children enables a thorough evaluation of children's use of a range of language structures in different contexts and with different conversation partners so that influence of performance factors, such as frequency of words used, would be less prominent (Dale, 1996). Furthermore, through the use of parent-report measures, parents can constructively involve themselves as collaborators in their children's assessment or intervention, thereby fostering increased parent motivation and enabling provision of higher-quality information to professionals (Dale, 1996; Thal et al., 1999). As compared with other language screening methods such as testing and direct observation, parent-report measures are less costly to administer (Thal et al., 1999), since they require less time and involvement of trained professionals in its administration (Dale, 1996). The representativeness and cost-effectiveness of parent-report measures enable them to be suitable for both clinical and research applications (Dale, 1996; Thal et al., 1999). Clinically, parent-report measures can be used as a screening tool to allow professionals to identify children at risk of language impairment and provide subsequent in-depth evaluation (Dale, 1996; Thal et al., 1999).

Despite their advantages, there are also doubts in the use of parent-report measures in clinical practice (Dale, 1996). First, parents are typically perceived as lacking professional

training in language development, such that they may lack sensitivity to subtle aspects of language development, particularly those appearing later in the developing language stage, such as the development of logical reasoning, which would be observed by professionals (Dale, 1996; Thal et al., 1999). In addition, parents are also perceived to be unable to assess their impression of their children's abilities critically (Dale, 1996; Thal et al., 1999). A natural pride in their own children may lead to overestimation of their performance, whereas frustration in the case of poor performance may result in underestimation (Dale, 1996). Nevertheless, parent-report measures still hold promises for further development despite concerns with their validity in clinical practice. Diamond and Squires (1993) has postulated that other professionally administered screening methods, such as testing and direct observation, have an equal possibility to underestimate children's language abilities, especially in screening children with disabilities. Direct screening of children's language abilities, which typically take place in unfamiliar circumstances, may only reflect the forms that children are likely to use (Dale, 1996). Specific skills that are representative of children's true language ability may not be observed during the brief encounter.

The DLSS as a Parent-Report Questionnaire

Parent-report measures can take the form of parent interviews or parent-report questionnaires. Compared with parent interviews, the use of parent-report questionnaires is even more cost and time-efficient, as it requires minimal professional involvement in its administration. As parent-report questionnaires have found to be valid in screening preschool children with language impairment in a number of previous studies, the present

study aimed to be a pilot study to investigate the accuracy and concurrent validity of the DLSS as a parent-report questionnaire. The present study aimed to evaluate the accuracy of the DLSS as a parent-report questionnaire in identifying children with language impairment by presenting the sensitivity, specificity, positive likelihood ratio, and negative likelihood ratio to enable evidence-based practice for future use of the screening tool. The concurrent validity of the DLSS as a parent-report questionnaire was also examined by providing correlations between scores on the DLSS and those on standardized tests. Since the use of parent-report questionnaires is more cost and time-efficient in screening preschool children as compared to professionally administered screening procedures (Klee et al., 1998), validation of the DLSS as a parent-report questionnaire offers an attractive alternative to parent interview as used in previous studies (Lee et al., 1990; Wong et al., 1992) for early identification of children at risk of language impairment. This study aimed to examine: (1) The accuracy of the DLSS in identifying 36 to 47 month-old children with language impairment, and (2) The concurrent validity of the DLSS in identifying 36 to 47 month-old children with language impairment.

The present study was conducted in two stages. In Stage I, a screening was conducted to identify children at risk of language impairment using the DLSS. In Stage II, a double-blind follow-up clinical assessment was administered on a sample of the children who participated in Stage I using two standardized language tests as reference standards.

Method

Stage I Screening

Procedures.

All children in the K1 classrooms in three kindergartens were invited to participate in the study. Two of the kindergartens were situated in public housing estates in the Tseung Kwan O and Tsing Yi district. Children of these two kindergartens typically came from families in the lower extreme of social class. The remaining kindergarten was located in the town centre of Sai Kung. Children of this kindergarten came from families in both extremes of social class. A letter explaining the purpose of the study, a parent consent form, a case history form (Appendix A), and a copy of subscales A and B of the DLSS were distributed to the caregivers via the kindergartens. The case history form was used to obtain background information such as previous speech and language intervention, parents' concerns, parent education, parent occupation, and primary language spoken at home. Since the DLSS was originally designed to be completed as a parent interview, examples were added to several items of the original questionnaire to enhance parent understanding. Only subscales A and B were used for direct indication of children's language performance (Lee et al., 1990). Those willing to participate in the study were asked to return the parent consent form with the completed questionnaire and case history form to the kindergarten within one week after they received the questionnaire.

Participants.

There were 149 K1 children in the three kindergartens. This sample size was

anticipated to yield approximately 15 children with positive screens using a 15% projected positive yield (Klee et al., 2000). As the present study aimed to investigate the accuracy of the DLSS in screening 36 to 47 month-old children with language impairment, children who were younger or older than the age range specified, and children who had previously received speech and language assessment or intervention were excluded. Since the standardized language tests to be used in Stage II were originally designed for use with a monolingual population, children who were bilingual (indicated by a rating of six or less out of ten in the proportion of time spent using Cantonese reported in the case history form) were also excluded. The completed forms were collected and scored by a colleague of the principal investigator to avoid bias in administration of the reference tests.

Stage II Clinical Diagnosis

Participants.

Participants who scored less than two standard deviations in subscales A, B, or both A and B in the DLSS were considered as screened positive cases, while the remaining participants were considered as screened negative cases (Wong et al., 1992). A total of 20 children were invited to participate in Stage II. These included all the screened positive cases and a randomly selected sample of screened negative cases.

Procedures.

The follow-up clinical assessment was conducted within two weeks after the questionnaires were collected. The assessment was conducted by three final fourth year undergraduate Speech and Hearing Sciences students at the University of Hong Kong. To

ensure proper administration of the reference tests, the supervisor of this study, who is a practicing speech therapist, observed and gave feedback to the test administrators' running of the two reference tests on a three-year-nine-month old child in a pilot session, after which the test administrators' interpretation of test results were also discussed.

At this stage of the study, both the test administrators (including the investigator) and the caregivers were blind to the screening results to avoid bias and subjectivity, which may influence the study's results. Each assessment session lasted approximately one hour and fifteen minutes, in which participants were assessed using two reference tests: the Cantonese version of the Reynell Developmental Language Scale (RDLS-C) – Receptive Scale (Hong Kong Society for Child Health and Development, 1987) and a delayed sentence imitation task using stimuli from the Cantonese Adaptation of the Test for Reception of Grammar (CTROG) (Mok, 1995).

Measures.

Reynell Developmental Language Scale (RDLS-C) – Receptive Scale: The Receptive Scale of the RDLS-C was administered to provide follow-up assessment on verbal comprehension. This assessment tool is a commercially available and widely used test among speech therapists in Hong Kong.

Delayed Sentence Imitation: The Expressive Scale of the RDLS-C was replaced by a delayed sentence imitation task using stimuli from the Cantonese Adaptation of the Test for Reception of Grammar (CTROG) (Mok, 1995), which was a product of a master's thesis.

As criticized by Edwards, Garman, Hughes, Letts, & Sinka (1999), the English version of the

Expressive Scale of the RDLS (Reynell & Huntley, 1985), on which the Cantonese version was based, lacked informative results and was prone to possible subjectivity during administration. Use of the sentence imitation task has the advantage of being able to assess production of a large scope of grammatical structures within a short period of time, which would otherwise be impossible in language sample analysis (Mok, 1995). A 0-3 scoring method was used for assessment purposes (Mok, 1995). The sentence imitation task has demonstrated good construct validity, as well as high test-retest and inter-rater reliabilities (Mok, 1995). The validity of using sentence imitation to assess Cantonese-speaking children's expressive language ability was also demonstrated in a recent study by Stokes, Wong, Fletcher, and Leonard (2006).

A cut-off point of one standard deviation below mean was used to define caseness for both measures owing to the small sample size. Children scoring less than one standard deviation below mean in either or both measures were diagnosed to have language impairment. After the completion of Stage II, the caregivers of Stage I participants were notified of the screening results in the form of a reply slip, while those of Stage II participants were consulted of the assessment results via telephone conferencing. To ensure proper test interpretation, the principal investigator and her supervisor discussed all assessment findings and recommendations of these children before the consultation. For participants who were diagnosed to have language impairment, resources for treatment options were provided to the caregivers. The children's performance were only distributed to their corresponding kindergartens after consent was granted from the caregivers.

Inter-rater Reliability of Reference Test Scores.

Inter-rater reliability was evaluated by comparing the scores rated by two test administrators on two randomly selected participants. Since the RDLS-C involved classification of responses into mutually exclusive categories (i.e. correct and incorrect responses), percentage agreement (calculated by dividing the number of instances the two raters agree by the number of instances of opportunities to agree) was used to evaluate inter-rater reliability. As the delayed sentence imitation task involved use of a 0-3 scoring method, the Pearson Product-Moment Correlation Coefficient r , calculated using the Statistica computer software, was used for the purpose. Percentage agreement of the test administrators' ratings on the RDLS-C was 100%, whereas agreement on the delayed sentence imitation task yielded an r of 0.87. The strong positive correlations indicated good inter-rater reliability.

Results

Outcomes from Stage I

Of the 149 questionnaires distributed in Stage I, 86 were completed and returned (response rate = 57.72%). Based on the exclusion criteria mentioned previously, 17 participants were excluded from the sample of 86 participants agreeing to participate in the study. Among the 17 participants excluded, six were older than 47 months, four were bilingual, and seven had previously received speech and language assessment or intervention. Using the criteria previously defined, 12 out of the 69 participants included were screened positive cases.

Outcomes from Stage II

All 12 of the screened positive cases and eight children who were randomly selected from the 57 screened negative cases participated in Stage II. Of the sample of children assessed, seven out of the twelve screened positive cases (58%) and two out of the eight screened negative cases (25%) were boys. The mean age of the participants on the day of the follow-up clinical assessment was 41.9 months.

Using a cut-off point of one standard deviation below mean in either or both of the reference tests, eight of the 20 participants (four boys and four girls) were diagnosed to have language impairment. Six of the 12 screened positive cases were true positives, while six of the eight screened negative cases were true negatives. Six of the 12 screened positive cases were false positives, while two of the eight screened negative cases were false negatives. The relationship between results of the DLSS and those of reference tests can be expressed as follows:

	Tested Positive	Tested Negative
Screened Positive	True positives (a)	False positives (b)
Screened Negative	False negatives (c)	True negatives (d)

Discriminant Analysis

A discriminant analysis was conducted to investigate the discriminant accuracy of the DLSS to evaluate the extent to which the results of the DLSS captured the participants' true language ability, as reflected from their performance in the reference tests. Metrics of discriminant accuracy, such as sensitivity, specificity, positive likelihood ratio (LR+), and

negative likelihood ratio (LR-), of the DLSS were calculated. Sensitivity refers to the extent to which participants who are true language-impaired cases are identified as being so in the screening measure ($a / (a + c)$). Specificity refers to the extent to which typically developing children are identified as being so during screening ($d / (b + d)$). LR+ is an index of the degree of confidence that a screened positive case truly has language impairment ($\text{sensitivity} / (1 - \text{specificity})$), whereas LR- is an index of the degree of confidence that a screened negative case truly does not have language impairment ($(1 - \text{sensitivity}) / \text{specificity}$). As suggested by Plante and Vance (1994), a benchmark of at least 80% should be reached for sensitivity and specificity values for a diagnostic measure to be considered to have acceptable discriminant accuracy. Dollaghan (2004) suggested that an accurate diagnostic measure should have high LR+ and low LR-. A LR+ of 10 or more and a LR- of 0.2 or less indicates a high likelihood of true positive and true negative results respectively, while a LR+ of 4 or less and a LR- of 0.4 or more indicates intermediate results only, in which additional testing is required to confirm the diagnosis (Dollaghan, 2004). Results of the discriminant analysis revealed a sensitivity of 75%, specificity of 50%, LR+ of 1.5, and LR- of 0.5.

Participants seen in Stage II of the study were divided into subgroups according to their outcomes in Stage I on the basis of their DLSS scores (i.e. screened positive and screened negative). Descriptive statistics were obtained for results of subscales A and B of the DLSS and those of the reference tests (Table 1). Differences in performance between the subgroups were assessed using the Mann-Whitney U test with statistical significance set at $p < 0.05$. Table 1 shows the mean scores and standard deviations of the subgroups'

performance in the measures. As expected, the mean scores of the screened positive and screened negative subgroups differed significantly on subscale A ($U = 1.00, p < 0.05$), subscale B ($U = 7.00, p < 0.05$), and subscales A and B ($U = 1.50, p < 0.05$) of the DLSS. A lack of statistically significant differences was however evident between the performances of the two subgroups on the receptive language ($U = 24.50, p > 0.05$) and expressive language ($U = 24.00, p > 0.05$) reference tests. Such findings nevertheless confirmed what was expected considering the unacceptable discriminant accuracy of the DLSS, as shown by the low sensitivity, specificity, and LR+ values and high LR- value.

Table 1. Comparison of screening group performance in the DLSS and reference tests

Measure	Screened Positive		Screened Negative		U
	Cases (n = 12)		Cases (n = 8)		
	mean	SD	mean	SD	
DLSS (Subscale A)	9.58	1.62	13.50	0.76	1.00*
DLSS (Subscale B)	15.83	3.24	21.75	2.55	7.00*
DLSS (Subscales A + B)	25.42	3.92	35.25	3.11	1.50*
RDLS-C – Receptive Scale	38.25	9.74	46.88	7.49	24.50
Delayed Sentence Imitation	64.25	37.80	100.88	49.37	24.00

Note: DLSS = Developmental Language Screening Scale; CRDLS = Cantonese version of Reynell Developmental Language Scale

* $p < 0.05$, two-tailed.

Concurrent Validity of the DLSS

To assess the concurrent validity of subscales A and B, Pearson Product-Moment Correlation analyses were run to determine the degree of association between children's scores from the subscales and those from reference tests, with statistical significance set at $p < 0.05$. As shown in Table 2, positive correlations ($r = 0.27 - 0.34$) were found between scores in subscale A, scores in subscale B, and the total score in subscales A and B and that in the reference tests, indicating those participants achieving higher scores in the DLSS also tended to achieve higher scores in the reference tests. However, as expected given the abundance of falsely identified cases, the correlations were weak and lacked statistical significance.

Table 2. Correlations between scores on the DLSS and reference tests ($N = 20$, $df = 18$)

	RDLS-C – Receptive Scale Mean = 41.70 (SD = 9.97)	Delayed Sentence Imitation Mean = 84.79 (SD = 46.25)
DLSS (Subscale A) Mean = 11.15 (SD = 2.37)	0.27	
DLSS (Subscale B) Mean = 18.20 (SD = 4.16)		0.32
DLSS (Subscales A + B) Mean = 29.35 (SD = 6.07)	0.31	0.34

Note: DLSS = Developmental Language Screening Scale; RDLS-C = Cantonese version of Reynell Developmental Language Scale

Discussion

Discriminant Accuracy of the DLSS as a Parent-Report Questionnaire

The sensitivity (75%) and specificity (50%), in addition to the LR+ (1.5) and LR- (0.5) of subscales A and B of the DLSS reflects that the measure fails to enable accurate identification of children with language impairment, based on the benchmarks proposed by Plante and Vance (1994) and Dollaghan (2004). The low sensitivity indicates that the DLSS has a relatively high tendency to overlook children with language impairment. The findings of two false negative and six true positive cases suggest that as many as one quarter of children with language impairment (2/8) could have been overlooked by the DLSS. The low specificity, on the other hand, indicates that the DLSS has a relatively high tendency to over-refer children without language impairment. Counting the two children who were diagnosed to have language impairment prior to the study, the percentage of 36 to 47 month-old Cantonese-speaking children diagnosed to have language impairment in the sample was 6.99% (10/143). Fifteen percent (12/80) of the 36 to 47 month-old participants in Stage I of the study were screened positive on subscales A and B of the DLSS. If the percentage occurrence of 6.99% were used, 53% of the screened positive cases ($[(15-6.99)/15] \times 100\%$) would have been over-referred. The findings are consistent with the low LR+ and high LR- obtained, which suggest that a screened positive case does not necessarily imply presence of language impairment, while a screened negative case does not necessarily imply absence of language impairment.

The high rate of misidentifications indicates that subscales A and B of the DLSS fails

to meet the level of accuracy required to be used as a screening tool for identifying three-year-old Cantonese-speaking children with language impairment. As stated earlier, using parent-report questionnaires has the advantage of being more economical than conventional screening methods such as parent interviews and is thus a more practical method for mass screening (Thal et al., 1999). However, if the DLSS were to be used for screening at the present stage, providing follow-up services for all positive screens would be uneconomical due to the large number of over-referred cases. Furthermore, false identifications may create unnecessary anxiety for both the child and the family concerned. Considering the economical and social consequences of misidentification, the DLSS fails to be an effective tool for screening three-year-old Cantonese-speaking children with language impairment.

A number of factors may have caused the abundance of falsely identified cases. First, several items in subscales A and B of the DLSS require caregivers' report on behaviours that should have emerged two to three years ago (e.g. whether the child babbles or makes noise meaningfully). As suggested by Dale (1996), caregivers' report on children's past language abilities is typically less accurate than that on children's current language abilities, since having caregivers report on past behaviours places greater demands on their memory and may thus be vulnerable to inaccurate recall. There are also items requiring caregivers' report on cumulative behaviours (e.g. whether people who know the child well can understand his/her speech with some guessing, and whether people who know the child well can understand his/her speech without the need to guess). Inclusion of items that are

cumulative and concerned past behaviours in the questionnaire may have contributed to an increase in the number of false positive cases, as detailed analysis of the completed questionnaires revealed that caregivers tended only to fill in the child's most current ability. Items that are cumulative and concern past behaviours were typically missed out despite examples were added to enhance understanding when the DLSS was adapted for use as a parent-report questionnaire in this study. In addition, as observed in the consultation session after the study in which assessment results were discussed with the caregivers, the caregiver of one of the false positive cases has very negative perceptions on her child's language ability, which may be attributed to high expectations towards her child. From results of the study, the caregiver's negative perceptions has led to low scores in the DLSS despite the child's good performance in the reference tests, which has given rise to an outlier in the data set obtained. This example illustrates that biased perceptions may limit caregivers' expectations such that they may underestimate their child's ability when filling in the questionnaire. Furthermore, caregivers may lack knowledge in the terminology used (e.g. contextual cues, nouns) such that items that should have been relevant to the child may not have been responded to adequately.

Concurrent Validity of the DLSS

Results of the current study indicate that the DLSS lacked concurrent validity in identifying three-year-old Cantonese-speaking children with language impairment, as evidenced by the positive but statistically insignificant correlations between scores from the DLSS and those from the reference tests. The weak correlations may be due to the

abundance of falsely identified cases, which may be attributed to the factors discussed earlier (please refer to the discussion on the discriminant accuracy of the DLSS). Given that there were quite a few children with poor performance in screening turned out to perform well in the reference tests, or vice versa, a strong positive linear relationship could not be expected between scores of the DLSS and that in the reference tests. Given the small sample size, the association between results of the screening and reference tests may also be subjected to great variability, such that a discernible trend cannot be observed, thereby resulting in the lack of statistical significance in the screening's validity. In their study of the DLSS, Lee et al. (1990) identified significant correlation between scores in subscales A and B and the language quotients in comprehension ($r = 0.45$) and expression ($r = 0.40$) based on scores in the RDLS-C. However, as the language quotients were derived on the basis of age-equivalent scores using a discrepancy-based criterion, comparisons could not be made between the correlations obtained in the two studies.

Feasibility of Using the DLSS as a Parent-Report Questionnaire

Of the 149 questionnaires distributed, 57.72% were completed and returned. Such a response rate was comparable to those reported in previous studies involving the use of mailed-in parent-report questionnaires for screening preschool children with language impairment, in which response rate ranged from 53% (Klee et al., 1998) to 75% (Stott et al., 2002). A response rate of 57.72% indicates that just over half of the eligible children were screened in the study. Such a response rate could be influenced by a number of factors. First, caregivers may be unwilling to spend time to complete the questionnaire. As stated in

the cover letter, about 15 minutes was required to complete the questionnaire. Second, caregivers may be doubtful about providing personal information to the investigator. Though the cover letter has stressed that all information obtained would be kept confidential, some caregivers may still feel skeptical about volunteering information, such as their family history of speech and language impairment, parents' educational level, and parents' occupation, as required by the case history form to a stranger. Additionally, the caregivers' noncompliance may have been due to denial of their child's problem. Some of the caregivers who suspected their child as having problems in language, or those whose child had been diagnosed as having language impairment but denied the problem may fear that returning the questionnaire would lead to confirmation of their child's problem. No information was available, however, on the nonrespondents to confirm any of the above hypotheses.

Criteria for Defining Language Impairment

Counting the two children who were previously diagnosed to have language impairment prior to the study, the percentage of 36 to 47 month-old Cantonese-speaking children diagnosed to have language impairment in the sample was 6.99% (10/143) using a cut-off point of one standard deviation below mean in either or both of the reference tests. This figure is twice as large as the prevalent rate of 3.4% identified by Wong et al. (1992). In their study, Wong et al. (1992) utilized a cut-off point of "language age on the Verbal Comprehension Scale and the Verbal Expression Scale is equal or less than two-thirds of the chronological age" (p.252) for the diagnosis of language impairment. Based on such criteria,

a 42 month-old child would only be diagnosed to have language impairment if a raw score of 25 or less was obtained in the Receptive Scale and a raw score of 30 or less was obtained in the Expressive Scale of the RDLS-C, which is equivalent to 3 SDs (0.14th percentile) and 2.6 SDs (0.47th percentile) below mean respectively. The larger percentage of diagnosed language impairment cases in the present study may therefore be attributed to the more conservative criteria used. As stated by Tomblin, Records, and Zhang (1996), a score of 1 SD below the mean resembles the point where a majority of practicing clinicians makes a clinical decision to diagnose children as having language impairment. If the discrepancy-based criterion as defined by Wong et al. (1992) were used, only children with profound impairments in language would be eligible for language intervention services. Children scoring between -1 SD (16th percentile) and -2.6/-3 SD in Receptive and Expressive Scales of the RDLS-C, who also present significant risks for problems in language, would not be diagnosed and would therefore not have the opportunity to receive the clinical follow-up they need using such criterion for diagnosis.

Due to limitations in resources, the present study employed a school-based sampling method, in which participants were randomly selected from three pre-selected mainstream kindergartens. In the study by Wong et al. (1992), a stratified proportional sampling method was employed, in which participants were stratified according to the geographical region and size of the schools. In the study by Klee et al. (2000), a district-based sampling method was employed. Due to the sampling method used, the performance of participants in the present study may not be equally distributed, as factors such as their family's social economical

status were not controlled. Children with severe language impairments, who should be more easily identified, may also have already been allocated necessary services in special education settings. Such differences in sampling methods explain why the percentage of children diagnosed with language impairment in the present study is less than one-half of the 15% prevalence rate put forth by Klee et al. (2000), despite the same cut-off point for diagnosis was employed.

Research Implications

The developing language stage, which typically occurs in children between two to five years of age, has been described as “the most explosive stage of language development” (Paul, 2001, p.290). In Hong Kong, children will only receive regular screenings of their language and communication abilities under the Developmental Surveillance Scheme of the Integrated Child Health and Development Programme at 12 months, 18 months, and 48 months, with additional visits provided for those who require closer monitoring. Under such arrangements, children with potential risks of language impairment in the developing language stage are not likely to be identified until nearly the end of the stage. As evidenced from the present study, only two of the ten children diagnosed with language impairment received a diagnosis prior to the study, indicating the majority of children (80%) have not been identified during their previous visits in the Developmental Surveillance Scheme. Even if they were to be identified during their visit at 48 months, it is most likely that opportunities for early intervention will have been dissipated by the time they receive a diagnosis. In addition, the screening protocol used in the Developmental Surveillance

Scheme, or the DLSS used in the original parent interview format, have not yet been validated against an objective gold standard. Therefore, there is an immediate need for the development of a locally applicable screening tool that is validated against an objective gold standard for identifying preschool children with language impairment.

Future research can encompass validating the DLSS its the original parent interview format against an objective gold standard. The test administrators' clinical impression can also be incorporated in clinical diagnosis in addition to results of the reference tests to form a "gold standard", as stated by Klee et al. (1998), to reduce the effect of performance factors in hope to minimize the amount of falsely identified cases.

Alternatives to further studies of the DLSS can involve validating the screening protocol currently used in the Developmental Surveillance Scheme against an objective gold standard. Since children receive screening at 18 months of age under arrangements of the Developmental Surveillance Scheme, the predictive validity of the screening protocol should also be investigated. Since a certain proportion of children with poor language performance at 18 months may be late talkers, the accuracy of screening for language impairment at this age may be questioned. Particular emphasis should therefore be placed on the screening protocol's ability to differentiate children who are late talkers to those who are truly language impaired.

Given the questionable accuracy of screening for language impairment at 18 months of age, another direction for future research would be to develop a new parent-report screening measure to serve as a cost-effective supplement to the Developmental Surveillance Scheme to

enable mass screening when children enters preschool to identify children at risk of language impairment. Identifying children with language impairment early in the developing language stage ensures opportunities for early intervention. As postulated by Dale (1996), parent-report measures tend to be more accurate if a recognition format was utilized, with items limited to current behaviours to reduce the demands placed on respondents' memory. The response rate of parent-report questionnaires can be increased by highlighting the importance and benefits of early identification of children with language impairment to caregivers when introducing the screening, increasing the amount of written instructions on how to complete the questionnaire, as well as using simpler everyday language to enhance understanding. As possible bias may be present in caregivers' perception of the child, future work can also compare the accuracy and validity of a screening tool when only caregivers' report was considered to that when both the caregivers' and teacher's report was considered. In addition, the effect of caregiver's perception can also be investigated to evaluate the effect of overly high or low expectations on children's performance on the validity and accuracy of a parent-report measure.

References

- American Psychiatric Association. (1994). *Diagnostic and statistical manual of mental disorders* (4th ed.). Washington, DC: American Psychiatric Association.
- American Speech-Language-Hearing Association. (1993). Guidelines for caseload size and speech-language service delivery in schools. *American Speech-Language-Hearing Association, 35*(suppl.10), 33-39.
- Dale, P.S. (1996). Parent report assessment of language and communication. In K.N. Cole, P.S. Dale, & D.J. Thal (Eds.), *Assessment of Communication and Language* (pp. 161-182). Baltimore: Paul H Brookes.
- Diamond, K.E. & Squires, J. (1993). The role of parental report in the screening and assessment of young children. *Journal of Early Intervention, 17*, 107-115.
- Dollaghan, C.A. (2004). Evidence-based practice in communication disorders: what do we know, and when do we know it? *Journal of Communication Disorders, 37*, 391-400.
- Edwards, S., Garman, M., Hughes, A., Letts, C., Sinka, I. (1999). Assessing the comprehension and production of language in young children: an account of the Reynell Developmental Language Scales III. *International Journal of Language and Communication Disorders, 34*(2), 151-171.
- Feldman, H.M., Dale, P.S., Campbell, T.E., Colborn, D.K., Kurs-Lasky, M., Rockette, H.E., et al. (2005). Concurrent and predictive validity of parent reports of child language at ages 2 and 3 years. *Child Development, 76*(4), 856-868.
- Glascoe, F.P. (2005). Screening for developmental and behavioural problems. *Mental*

Retardation and Developmental Disabilities Research Reviews, 11, 173-179.

Hong Kong Society for Child Health and Development. (1987). *Reynell Developmental Language Scales, Cantonese (Hong Kong) Version*.

Klee, T., Carson, D.K., Gavin, W.J., Hall, L., Kent, A., & Reece, S. (1998). Concurrent and predictive validity of an early language screening program. *Journal of Speech, Language, and Hearing Research*, 41, 627-641.

Klee, T., Pearce, K. & Carson, D.K. (2000). Improving the positive predictive value of screening for developmental language disorder. *Journal of Speech, Language, and Hearing Research*, 43, 821-833.

Law, J. (1992). What is language impairment? In J. Law (Ed), *Early Identification of Language Impairment in Children* (1st ed.) (pp.20-40). London: Chapman & Hall.

Lee, P.W.H., Luk, E.S.L., Bacon-Shone, J., Lau, J., Wong, V., & Lo, L. (1990). The Developmental Language Screening Scale: A validation and normative study. *The Hong Kong Journal of Paediatrics*, 1, 7-22.

Lee, P.W.H., Luk, E.S.L., Yu, K.K., & Bacon-Shone, J.H. (1985). A developmental language screening scale for use in Hong Kong. *The Hong Kong Journal of Paediatrics*, 2, 152-165.

Mok, C. K-F. (1995). *Comprehension and expression of grammatical structures in Cantonese-speaking preschoolers*. Hong Kong: University of Hong Kong.

Paul, R. (2001). *Language disorders from infancy through adolescence: assessment and intervention* (2nd ed.). St. Louis: Mosby.

Plante, E. & Vance, R. (1994). Selection of preschool language tests: a data-based approach.

Language, Speech, and Hearing Services in Schools, 25, 15-24.

Reynell, J. & Huntley, M. (1985). *Reynell Developmental Language Scales*, revision of manual. Windsor: NFER-Nelson.

Stokes, S.F., Wong, A. M-Y., Fletcher, P., & Leonard, L.B. (2006). Nonword repetition and sentence repetition as clinical markers of specific language impairment: the case of Cantonese. *Journal of Speech, Language, and Hearing Research, 49*, 219-236.

Stott, C.M., Merricks, M.J., Bolton, P.F., & Goodyer, I.M. (2002). Screening for speech and language disorder: the reliability, validity and accuracy of the General Language Screen. *International Journal of Language and Communication Disorders, 37*(2), 133-151.

Thal, D.J., O'Hanlon, L., Clemmons, M., & Fralin, L. (1999). Validity of a parent report measure of vocabulary and syntax for preschool children with language impairment. *Journal of Speech, Language, and Hearing Research, 42*, 482-496.

Tomblin, J.B., Records, N.L., & Zhang, X. (1996). A system for the diagnosis of specific language impairment in kindergarten children. *Journal of Speech and Hearing Research, 39*, 1284-1294.

Wong, V., Lee, P.W.H., Lieh-Mak, F., Yeung, C.Y., Leung, P.W.L., Luk, S.L., et al. (1992). Language screening in preschool Chinese children. *European Journal of Disorders of Communication, 27*, 247-264.

World Health Organization. (1992). *International code of diseases* (10th ed.). New York:

World Health Organization.

Appendix A
Case History Form

Background Information 基本資料

Child Name 學童姓名：_____ Date 填寫日期：_____

Sex 學童性別：_____ Date of Birth 出生日期：_____

Age 年齡：_____ years 歲 _____ months 個月

Respondent 填寫人姓名：_____

Relationship with Child 與學童之關係：_____ Tel 聯絡電話：_____

Please check (✓) all appropriate items 請在適當的空格內填上“✓”號

Child Background 學童背景

1. Has your child ever received a speech and language assessment?
貴子弟曾否接受言語評估？
No 沒有 Yes 有 (Results 結果：_____)
2. Has your child ever received speech and language therapy?
貴子弟曾否接受言語治療？
No 沒有 Yes 有 (Reason 原因：_____)
3. Has your child ever been diagnosed to have any other problems (e.g. ear infection)?
貴子弟曾否被診斷患有其他病患 (例：中耳炎)？
No 沒有 Yes 有 (Please specify 請註明：_____)
4. What are your concerns towards your child's speech and language development?
您對貴子弟的言語發展有何顧慮？

Family Background 家庭背景

1. What language(s) do you use at home?
在家中使用的語言 (可多於一種)？

How much time will you use to speak to your child in Cantonese?

您會花多少時間與貴子弟以廣東話溝通？

never 從不					half 一半					always 經常
0 <input type="checkbox"/>	1 <input type="checkbox"/>	2 <input type="checkbox"/>	3 <input type="checkbox"/>	4 <input type="checkbox"/>	5 <input type="checkbox"/>	6 <input type="checkbox"/>	7 <input type="checkbox"/>	8 <input type="checkbox"/>	9 <input type="checkbox"/>	10 <input type="checkbox"/>

2. How many hours do you spend with your child each week?
 您每週約花多少時間與貴子弟相處？ _____ hours 小時
3. How many siblings does your child have?
 貴子弟共有多少兄弟姐妹？
 elder brother 兄 _____ elder sister 姐 _____
 younger brother 弟 _____ younger sister 妹 _____
4. Who usually take care of your child?
 貴子弟通常由誰照顧？ _____
5. Has any of your family member been diagnosed to have a speech and language problem?
 家中有沒有成員被診斷患上言語障礙？
 No 沒有 Yes 有 (Please specify 請註明： _____)

Parent Background 家長背景

1. Educational level 教育程度
- Father 父： Primary 小學 Secondary 中學
 Tertiary 大專/大學 Master or 碩 above 士或以上
- Mother 母： Primary 小學 Secondary 中學
 Tertiary 大專/大學 Master or 碩 above 士或以上
2. Occupation 職業 Father 父： _____ Mother 母： _____

Appendix B
Developmental Language Screening Scale

Please check (✓) **all** items that your child **is/was** able to perform.

若小朋友現在能夠做到或曾經做到某項目，請在該項目旁的空格內填上“✓”號。

A. VERBAL COMPREHENSION 言語理解

I. General 綜合表現

1. Responds when name is called
叫佢個名佢識得有反應
2. Responds when spoken to even when name is not mentioned
同佢講嘢，就算無叫佢個名，佢都識得有反應
3. Understands simple phrases used in familiar contexts
E.g. Every time dad comes home, would go fetch slippers when asked.
只係明白好簡單嘅句字，而且要同一個場合講先至明
例如：每次爸爸返屋企，叫佢「擺拖鞋」佢會照做
4. Knows meaning of some words without contextual cues
E.g. Can fetch a cup when asked without having to point at it.
知道一 D 簡單的字句 D 意思而不需要其他提示
例如：叫佢「擺個杯嚟」，你無需以手指指著個杯給予提示，幫佢理解
5. Follows instructions involving two named objects
E.g. Give me a bowl and a cup
可以叫佢擺兩件物件
例如：擺個碗同個杯俾我
6. Can be sent out of room to fetch two objects
E.g. Go to the bedroom and fetch the pillow and the blanket
可以叫佢出房以外擺兩件或以上的物件
例如：去睡房度擺個枕頭同張被嚟
7. Follows a sequence of three consecutive commands
E.g. Get the chopsticks, distribute them, then sit down
可以明白同埋遵照指示連續做到三樣嘢
例如：擺 D 筷子出嚟，派俾人，跟住坐好
8. Understands instructions involving decisions, i.e., if...then instructions
E.g. Put on a coat if you are cold, close the window if it is open
可以執行一 D 需要自己先行做決定嘅指令

例如：如果凍就著多件衫、如果隻窗開咗就柵番埋佢

II. Prepositions 介詞理解

1. Understands the meaning of the following words:

佢明白乜嘢叫:	喺裡面	in	<input type="checkbox"/>
	喺下面	under	<input type="checkbox"/>
	喺上面	on	<input type="checkbox"/>
	喺後面	behind	<input type="checkbox"/>
	之前	before	<input type="checkbox"/>
	之後	after	<input type="checkbox"/>

B. ABILITY TO USE SPEECH 言語表達

I. Grammar 文法運用

- Makes any sounds at all
可以發出任何聲音
- Babbles, or coos, without meaning
可以無意識咁發出「牙牙」聲或者其他聲音
- Babbles, or makes noise meaningfully
E.g. Makes noise when hungry
想要嘢嗰時會「牙牙」聲或者發出其他聲音
例如：肚餓時發出「啲啲牙牙」的聲音
- One definite word
識講一個字
- Gives names of two or more people or things when asked
問佢嗰陣，可以講出兩個或以上人或物件嘅名
- Spontaneously names two or more objects or people in the surrounding
可以自發地講出兩個或以上人或物件嘅名
- Speaks two-word phrases
E.g. kick ball, eat apple
識講由兩個字組合嘅句子
例如：踢 波、食 蘋果
- Speaks longer than two-word phrases with nouns and verbs, missing out linking words
E.g. I go to school, mum and I go shopping

可以講多過兩個字嘅句子

例如：我 返學、媽媽 同 我 去街

9. Appropriate use of words other than nouns and verbs (e.g. adjectives)

E.g. Adjectives (e.g. big, red), locatives (e.g. in front of, beside), adverbs of time (e.g. now, then)

除了名詞 (如：媽媽、狗仔) 同動詞 (如：食、跑) 之外，仲識講其他嘅字

例如：形容物件 (如：大、紅色)，形容位置 (如：前面、側邊)，形容時間 (如：而家、舊時)

10. Spontaneous use of sentences with linking words about present

E.g. Says “There are a lot of toys” when going into a toy shop

可以自發地講一 D 關於現在嘅句子

例如：入玩具店時話「好多玩具呀」

11. Forms sentences using *because*

E.g. Says “Because I fell” when asked why he/she has a plaster

識得用「因為」這字眼

例如：被問及點解要貼膠布時，回答「因為跌親」

12. Forms sentences with *but*

E.g. Says “But I want to stay longer” when asked to leave the park

識得用「但係」這字眼

例如：到公園玩耍要離開時，說「但係我想玩耐 D」

13. Talks about present, past, and future with complex grammatical constructions

E.g. I drank milk when I was small, I want to be a doctor when I grow up

可以講關於現在，過去及將來嘅嘢

例如：我以前飲奶、我第二時要做醫生

II. Asking Questions 問題運用

1. Asks questions using some simple overlearnt phrase(s).....

E.g. Is this it? Where’s doll?

可以用簡單的句子問問題

例如：喺咪咁？公仔呢？

2. Uses the following:

可以用以下字詞問問題: 邊度 Where?.....

乜嘢 What?.....

邊個	Who?.....	<input type="checkbox"/>
點解	Why.....	<input type="checkbox"/>
幾時	When?.....	<input type="checkbox"/>
點樣	How?.....	<input type="checkbox"/>

III. Intelligibility 咬字發音

1. Some speech, but mostly unintelligible
識講一 D 說話，但係其他人聽唔明白佢想講乜嘢
2. Only a few words (3+ at least) intelligible
佢講 D 說話其他人只係聽得明三個字左右
3. People who know the child well can understand but with some guessing ...
認識孩子嘅人，可以大致上猜到佢喺度講乜嘢
4. People who know the child well can understand without the need to guess.
認識孩子嘅人，唔需要猜亦聽得明佢講乜嘢
5. Strangers can understand with some guessing
陌生人可以大致上猜到佢講乜嘢
6. Strangers can understand without the need to guess
陌生人唔需要猜都明白佢講乜嘢

Acknowledgements

I would like to gratefully acknowledge Dr. Anita Wong for her suggestions and advice throughout this dissertation project. Many thanks are also due to Miss Iris Lam, Miss Melanie Law, and Miss Jessica Fan for their assistance in data collection and analysis.

Lastly, I would like to thank the principals, staff, parents, and children of the following kindergartens for their help and cooperation during data collection:

Sai Kung Lok-Yuk Kindergarten

Islamic Pok Oi Kindergarten

Yan Oi Tong Dan Yang Wing Man Kindergarten