



Title	Assessing expressive vocabulary in three-year-old children using two elicitation methods
Author(s)	Chan, Foon; 陳歡
Citation	
Issued Date	2009
URL	http://hdl.handle.net/10722/173658
Rights	Creative Commons: Attribution 3.0 Hong Kong License

Assessing Expressive Vocabulary in Three-year-old Children
Using Two Elicitation Methods

Chan Foon

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, 2009.

ABSTRACT

The study aimed to examine the reliability of Lam's (2008) expressive vocabulary test for 3-year-old Cantonese-speaking children in Hong Kong. The effect of age (3;02 – 3;07 vs. 3;08 – 4;01), elicitation method (dynamic video clips vs. static picture cards) and word type (noun vs. verb) on the children's performance were also investigated. A total of 49 participants were tested on Lam's (2008) final 57 word list and the original 18 verbs that were excluded in the study's last phase. It was found that children were able to name more verbs when actions were presented using video clips and their performance for nouns was no different whether video clips or pictures were used. Further analysis of data collected using video clips revealed that nine verbs met the required item statistics for inclusion in the expressive vocabulary test. These findings suggest how dynamic presentation of stimuli, actions in particular, allows us to have a more complete understanding of three-year-old children's vocabulary knowledge. Implications for clinical use were also discussed.

INTRODUCTION

Motivated by a need for an expressive vocabulary test for Cantonese speaking preschool children, Lam (2008) identified a set of words that are appropriate for use with three-year-old children in Hong Kong. The words were chosen after 3 phases. In the first phase, 269 words were identified after a review of Cantonese language sample archives, the local Chinese curriculum, and several expressive vocabulary tests developed for English-speaking children. Following the first phase, 20 preschool teachers were invited to rate on the difficulty level of each of these 269 words. This phase ended with a smaller list of 108 words. In the third and final phase, the 108 words were presented as picture cards and tested on 35 children. On the basis of the their difficulty and discrimination index scores, Lam (2008) recommended for inclusion in the test a final list of 57 words, including 40 nouns, 5 verbs and 12 descriptors (Appendix A).

One wonders why there is such a significant difference in the proportion of nouns and verbs on the final 57 word list. In a cross-linguistic analysis of mothers' report of young children's vocabulary, Bornstein et al. (2004) reported that children aged 20 months speaking languages as diverse as Spanish, Dutch, French, Hebrew, Italian, Korean and American English produced more nouns than verbs. They argued that these evidences pointed to a universal "noun bias". Do Chinese-speaking children show this noun bias in their early vocabulary development and produce fewer verbs than nouns?

Is "noun bias" universal and applicable to Cantonese-speaking children?

Tse, Chan and Li (2005) examined the vocabulary of 492 young Cantonese-speaking children, aged 36 – 60 months, when they engaged in toy play with another peer of the same age for 30 minutes. Tse et al. (2005) found that the proportion of verb tokens was 23.26% for 3-year-olds and 22.91% for 4-year-olds as compared to the proportion of noun tokens being 13.72% for 3-year-olds and 12.11% for 4-year-olds. This suggested that Cantonese-speaking

young children produced almost two times more verbs than nouns in their utterances.

In fact, Barrett (1999) also noted that the universal “noun bias” may not be truly universal. A cross-linguistic study done by Tardif (1999) provided direct evidence for this argument. Tardif (1999) compared the proportion of nouns and verbs in 24 English-speaking and 24 Mandarin-speaking children aged 5 to 20 months. The Mandarin-speaking children produced fewer nouns but more verbs when they were engaged in free play with toys and the same pattern was observed in parent’s reports of children’s vocabulary development. Those children, however, were reported to use more nouns when engaged in book reading activities (Tardif, 1999). Results from Tardif (1999) suggested that although Mandarin-speaking children might use more verbs than nouns in their early vocabulary development, it was plausible that the context and the method of elicitation can affect the type of vocabulary being elicited and hence our understanding of children’s vocabulary development, which will be discussed later.

Inclusion of more verbs and Specific language impairment (SLI)

As mentioned, the universal “noun bias” may not be applicable to Cantonese-speaking children, at least to those who are three- and four-years of age. This calls for the inclusion of verbs in any expressive vocabulary tests for Cantonese-speaking children in order to capture a more representative picture of their expressive vocabulary development. Besides, there are other reasons why we need to reconsider the inclusion of more verbs in expressive vocabulary tests. Studies have shown that children with specific language impairments (SLI) showed reduced diversity of verbs and increased semantic verb errors (Loeb, Pye, Redmond & Richardson, 1996). They also performed significantly poorer in verbs than their age peers (Fletcher, 1999). Verbs are problematic for these children and can potentially be useful in discriminating children with SLI from their typical peers.

However, expressive vocabulary test batteries currently available for Cantonese-

speaking children focus primarily on the assessment of children's knowledge of nouns. For example, there are a total of 26 items in the three sections testing expressive vocabulary in the Reynell Developmental Language Scale-Cantonese version (Reynell & Huntley, 1987). Among these items, only two are verbs and six are descriptors, with nouns taking up almost 70% of the total items. In the Hong Kong Cantonese Oral Language Assessment Scale (HKCOLAS), there is a section specially allocated for assessing nouns while verbs are only briefly assessed in story re-telling. Hence, it is plausible that these tests are not sensitive enough in distinguishing children with specific deficits in verb production. This is further supported by Loeb et al. (1996)'s study that although all six but one of the children with SLI scored within normal limits on the noun-dominant Expressive One-Word Picture Vocabulary Test (Brownell, 2000), none labelled verbs as well as their age peers. Hence, when standardized tests, which are often "noun biased", were used, verb problems in children with specific language impairment might not be identified.

Thus, a comprehensive expressive vocabulary test should include more verbs to make it more representative of developmental patterns and sensitive for the clinical diagnosis of language impairment. To develop an expressive vocabulary test for Cantonese-speaking children, we need to consider including more verbs on the list originally developed by Lam (2008). But why is it that a large proportion of the verbs (18 out of 23) were excluded from the final 57 word list in the last phase? (Appendix B)

Identification of final word list in Lam's (2008) study

According to Lam (2008), 18 of the original set of 23 verbs identified from Phase 2 were not included in the final 57 word list. Six of these verbs had high item facility (IF) value (>0.85), indicating that they were too easy for children at this age since a large proportion of children scored correct for them. The remaining 12 verbs were excluded due to low item discrimination (ID) value (<0.25), suggesting both high and low scorers obtained similar

performance on these items, which may be so difficult that they failed to discriminate children at different levels. Given that all these 12 verbs were either rated as easy or moderately difficult by a majority of the preschool teachers at phase 2, these verbs might actually be appropriate for children at this age. But nearly half of the children failed in six of these items. Hence, the reason behind children's failure in naming the verbs should be investigated.

Exclusion of verbs

First, Lam (2008) pointed out that the pictures illustrating some of the excluded verbs such as “摺” (fold), “推” (push) and “開” (open) were ambiguous, and therefore leading the children to produce a variety of off-target responses. For example, for the picture illustrating “開” (open), five out of the twenty-four children said “kAm2” (close) while the others gave responses such as “攞” (take), “擺” (put), “玩” (play), “打” (hit) and “搵” (find). The range of different responses elicited could be a result of the very nature of verbs. As verbs describe movement, motion and state of mind (Loeb et al., 1996), the use of static picture cards may not be the most effective in their elicitation. This is exactly what Cuetos and Alija (2003) found. In their study, 3-year-old children were more likely to produce off-target responses when static picture cards were used for eliciting verbs than when they were used for eliciting nouns. Cuetos and Alija (2003) suggested that alternative means which could capture the most prominent dynamic qualities of verbs should be used to elicit verbs.

Alternative method of elicitation for verbs

Recall that Tardif's (1999) study suggested that the context and method of elicitation could affect the type of vocabulary items children produce and thus our understanding of children's vocabulary development. This raises a question regarding the relatively larger proportion of nouns than verbs in Lam (2008)'s final 57 word list. Were the children in Lam's (2008) study really better in noun production, hence the large proportion of nouns

included? Is it possible that their knowledge in verbs was underestimated due to the use of static picture cards?

This review suggests that there is a need to explore an alternative method for the elicitation of verbs in children between three to four years of age. One alternative from Loeb et al. (1996) was, instead of using static pictures, examiners were to manipulate small figures and toys to probe for verbs. This elicitation method is better than the static one as it captures the dynamic properties of verbs. However, it is difficult to ensure that examiner(s) manipulate the action scenes in an identical fashion every time an assessment is carried out. So to circumvent this problem, we propose to record the action scenes used to elicit for verbs and present these scenes as dynamic video clips. The primary objective of this study therefore is to compare children's performance in verb production in two elicitation methods: the use of static media (picture cards) vs. dynamic media (video clips). The second objective is to replicate Lam (2008)'s study and expand her relatively small data set of 35 three-year-old children with a larger and different sample of children. By doing so, we would substantiate Lam (2008)'s finding and at the same time contribute to the existing normative data pool. Also, as children at this age are at their peak of language development, variations might be present in the age range of one year. Thus, the children were divided into two subgroups to examine whether the word list is sensitive to developmental changes within the year.

In sum, this study addressed these research questions:

1. Is the expressive vocabulary test developed by Lam (2008) reliable and replicable?
2. Do children perform differently on verbs and nouns in an expressive vocabulary test when the words are elicited using static picture cards and when they are elicited using dynamic video clips?
3. Do young 3-year-olds (3;02 – 3;07) perform differently when compared to older 3-year-olds (3;08 – 4;01)?

4. Should more verbs be included in the final word list identified by Lam (2008)?
5. What are some of the common naming errors for 3-year-old children?

METHODOLOGY

Participants

A total of 49 preschoolers (26 boys and 23 girls), aged between 38 months and 49 months ($M = 43.49$; $SD = 3.48$), participated in this study. They were children from the K1 classroom in three kindergartens in three different parts of Hong Kong (Happy Valley, Choi Hung and Kwai Hing) representing diverse socio-demographic backgrounds. No children were excluded from the study despite the fact that some of them did not fit into the selection criteria after review of the parent questionnaire received.

The reasons of nil exclusion are as followed, 1) four of the children exceeded the age limit of 4-year-old at the time of testing, by just one month. They made up almost 10% of the total number of participants and it would be important to keep them to maintain this sample size; 2) one child was receiving speech and language therapy at the time of testing, however it was only made known after data collection was completed. Since this study aimed at collecting normative data and children with speech and language impairment should be included in the normative sample in proportion to the actual prevalence rate for accurate representation of children's range of performance; and 3) three children were available only for one but not both testing sessions, however, they were evenly distributed with one child missing the static media session and two children missing the dynamic media session. In spite of this, their data were used in calculation of the descriptive statistics to contribute to a larger sample size.

The children were then grouped according to their age. Twenty-five of them belonged to the younger group aged 38 months – 42 months ($M = 40.56$; $SD = 1.69$) and 24 of them belonged to the older group aged 43 months – 49 months ($M = 46.54$; $SD = 1.79$). Upon return of questionnaires, no children was reported to have sensory deficits, thus all of them were scheduled for further testing.

Materials and Instruments

The children were tested on words from a combination of two word lists. One of them is Lam's (2008) final 57 word list developed for assessing expressive vocabulary in 3-year-olds Cantonese-speaking children. The other is the 18 verbs that were excluded after Lam (2008)'s direct testing of the children in the last phase. Together, the children were tested on an expanded word list of a total of 75 words, with 40 nouns, 23 verbs and 12 descriptors. These 75 words were elicited using both static picture cards (static media) and dynamic video clips (dynamic media) from each child (Appendix C).

For easy referencing, the final word list identified by Lam (2008) will be referred to as the final 57 word list and the word list tested in this study will be referred to as the expanded 75 word list, with target words with static nature (nouns and descriptors) referred to as nouns and target words with dynamic nature (action verbs) referred to as verbs.

For the dynamic form of test stimuli, the video clips were either self recorded or downloaded from the internet. For the static form of test stimuli, picture cards originally created by Lam (2008) were used for noun items with all but two cards illustrating verbs being replaced with screen shots from the video clips. This was done to minimize differences between the stimuli used in the two elicitation methods, ensuring that any changes in children's performance would be more likely to be due to the difference in elicitation method but not the difference in stimuli. The only two verbs not replaced were due to low quality of video clips which resulted in unclear illustration in screen shots and thus Lam's (2008) original picture cards were used.

The instruments used in this study were as follow.

1. A window XP operated laptop with Microsoft PowerPoint 2003 was used to administer the dynamic version of the expanded 75 word list.
2. For the dynamic test stimuli, the video clips were either recorded with Panasonic DMC-

FS3 digital camera or downloaded from the internet (www.youtube.com). For the static test stimuli, softcopy of the picture cards originally created by Lam (2008) and the screen shots of the video clips were colour printed with a colour photocopier.

3. A JNC (USB-350) digital voice recorder was used to record the children's responses.

Procedures

The expanded 75 word list was individually presented to each child using two methods of elicitation, namely static media for picture cards and dynamic media for video clips to minimize between group variations (Zechmeister, Zechmeister & Shaughnessy, 2001). The order of two elicitation methods was randomly assigned to each of the participants. This is done to balance off possible carry over effect (Zechmeister et al., 2001). The two elicitations were completed in two 10- 15 sessions by the same investigator in a quiet room in their own kindergarten within 3 to 7 days to minimize possible learning during the period.

The children were asked specific questions for different word types. To elicit names of objects or object parts, what-questions were used, such as “呢啲咩嚟架?” (What is this?) used for objects, “呢度全部都係咩嚟架?” (What are all these?) used for categories, “呢度呢啲咩嚟架?” (What is this part called?) used for the name of a specific part of an object. For attributes, contrast questions were used, for example, for the item “短” (short), “呢條褲長，咁個條褲呢?” (This pair of trousers are long, how about that pair?). For action verbs, questions like “佢係度做緊乜嘢呀?” (What is he/she doing?) were asked. The children were prompted to give a response for every stimulus presented and their responses were recorded online. Details please refer to Appendix D.

Scoring criteria and acceptable responses

For scoring criteria, it was adapted from Lam (2008). First, only the child's first response was scored and the second response was only scored in the following two conditions,

1) the first response was irrelevant to the target or 2) the first response was unclear due to low volume of speech or misarticulation and the second response was given upon verbal request of “講多次” (say it again). Besides, responses in the form of a phrase or a sentence which nevertheless included the target word were scored correct (e.g., “廚師煮緊飯” (The chef is cooking) was scored correct for “廚師” (chef)). On the other hand, misarticulated responses were scored as correct if the vowel and tone were produced correctly. For example, /te4/ or /t^he4/ was scored correct for /se4/ (snake) and /tyn4/ or /sy4/ was scored correct for /syn4/ (ship). Besides, children’s phonetic repertoire was also taken into account during scoring to ensure that they were not underscored.

In Cantonese, there are many occasions in which an object or a verb can be expressed with different wordings. Hence, alternative responses for some of the stimuli were accepted as correct.

Table 1. Alternative acceptable response (Adapted from Lam (2008))

Target response	Alternative acceptable responses
沖涼	洗白白
夾 (/gap3/)	/gap6/ or /gEp6/
斟	倒
顏色	色
泳池	水池
扭	/lIng2/ or /lIng6/
食物	嘢食
推	/UN2/
風箏	紙鳶
生果	水果
攪	撈
水喉	水龍頭
沙灘	海灘
砌	/tap3/ or /tap6/

RESULTS

Reliability and replicability of Lam's (2008) study

Recall that this study was a replication of Lam's (2008) expressive vocabulary test. The expanded 75 word list used in this study came from Lam's final 57 word list and the additional 18 verbs excluded in an earlier phase. With the original data available from the author of Lam (2008), the performance of the two cohorts of children was compared. The children in Lam (2008) were 42 months of age ($SD = 2.96$, $range = 37 - 47$ months) and those in this study were 43 months of age ($SD = 3.45$, $range = 38 - 49$ months). There was no statistical significance in the age of the two cohorts of children ($t(81) = 2.365$, $p > 0.05$). On the final 57 word list, Lam's (2008) children ($M = 30$, $SD = 11.67$) scored slightly worse than the children in this study ($M = 34$, $SD = 12.14$), but the difference was not statistically significant ($t(81) = 1.367$, $p = 0.175$). Such consistent performance on the final 57 word list suggested that the list included some of the typical words 3-year-old children could produce.

To examine the extent to which the level of difficulty of the items was similar in the two studies, item facility (IF), an index of the degree of difficulty of an item involving calculation of the percentage of participants scored correct on an item (Bailey, 1998), was calculated. A Pearson's correlation revealed a strong, positive and significant correlation. ($r = .748$) (Guilford, 1956), indicating that majority of the items that were easy (or difficult) for the children in Lam (2008) were also easy (or difficult) for those in this study. According to Guilford (1956), Pearson's correlation value of less than 0.2 was considered negligible, 0.2 – 0.4 was considered weak, 0.4 – 0.7 was considered moderate, 0.7 – 0.9 was considered strong and more than 0.9 was considered very strong. Figure 1 illustrates the linear relationship of the IF values in the two studies.

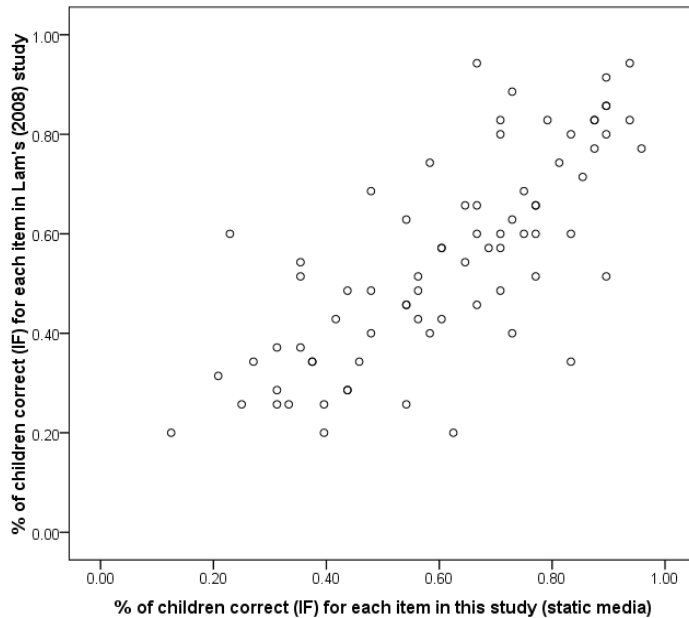


Figure 1. Linear relationship of the IF values in Lam's (2008) study and current study

Effect of age, elicitation method and word type on the children's performance in the vocabulary test

As there were 52 nouns and 23 verbs in the expanded 75 word list, each child's raw scores obtained for the two word types were converted to percentage scores before statistical analysis. A 2 (elicitation method) x 2 (word type) x 2 (age) mixed-model ANOVA revealed that the main effect for age was not significant ($F(1, 44) = 3.33, p > 0.05$). Thus, there was no significant difference in the performance of children from the younger age group ($M = 40.46, SD = 1.64, Range = 38 - 43$ months) compared to the older age group ($M = 46.41, SD = 1.82, Range = 44 - 49$ months). A significant main effect for elicitation method ($F(1, 44) = 19.37, p < 0.01, Partial\ Eta-squared = .31$) was obtained, suggesting that 31% of the obtained effect was attributable to the actual effect. The children performed significantly better when the stimuli were presented using dynamic ($M = 0.678, SD = 0.212$) than static media ($M = 0.619, SD = 0.205$). Besides, a significant main effect was also obtained for word types ($F(1,44) = 18.25, p < 0.01, Partial\ Eta-squared = .29$). The children performed significantly better in verbs ($M = 0.694, SD = 0.162$) than nouns ($M = 0.603, SD = 0.219$).

In addition to the main effects, a significant elicitation method and word type

interaction ($F(1,44) = 34.41, p < 0.01, \text{Partial Eta-squared} = .44$) was also obtained, revealing that the interaction effect had the largest proportion of the obtained effect attributable to the actual effect as compared to the main effects of elicitation method and word type. Examination of the cell means as plotted in Figure 2 indicated that although there was a large increase in mean percentage correct score for verbs when dynamic media ($M = 0.756, SD = 0.166$) was used than when static media ($M = 0.632, SD = 0.191$) was used, the mean percentage correct score for nouns did not differ significantly between the dynamic ($M = 0.600, SD = 0.225$) and the static media ($M = 0.607, SD = 0.219$).

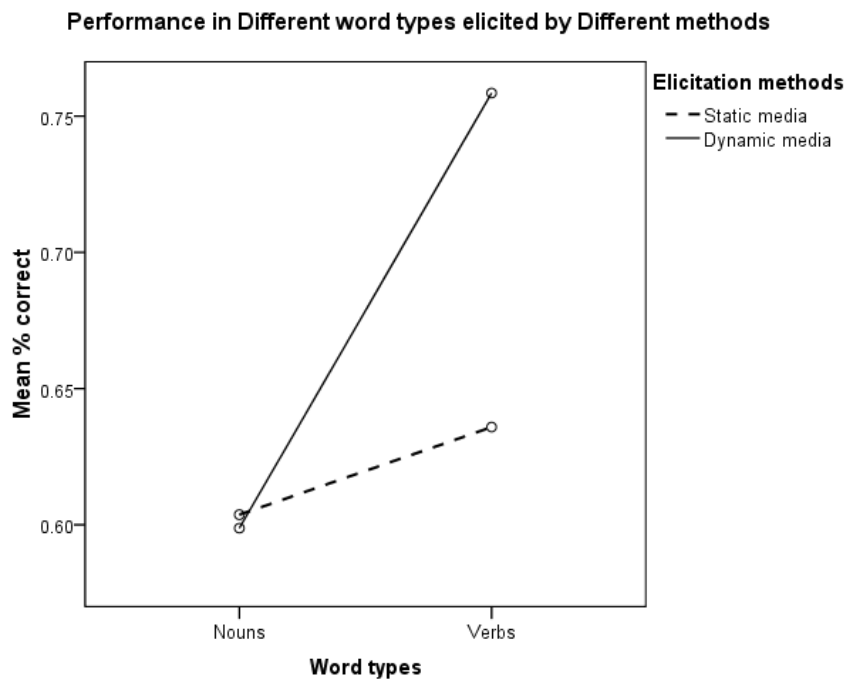


Figure 2. Mean percentage correct for different word types elicited by different methods.

Upon further analysis on the children's performance in the two word types, a very strong correlation ($r = .949$) was obtained in their performance in the 52 nouns across the two medias but only a moderate correlation ($r = .647$) was found in their performance in the 23 verbs (Guilford, 1956). These findings suggested that the children's naming of nouns was not affected by elicitation methods, as children who could name a noun in one media could successfully do so almost all the time in the other media as well. However, the situation was different in their naming of verbs. Some children who were not able to name the action

verb in the static media could successfully name it when it was presented in the dynamic media, resulting in a significantly higher group mean ($M = 0.756$). This difference in performance in the same action verbs across two medias results in a lower Pearson correlation. Figure 3 showed the distribution of the children's scores on the two word types elicited using two elicitation methods. Of note was an extreme score at the low end for verb production. This child's score was way apart from the rest of the group in verb production but the poor performance was not as obvious when it came to noun production as there were other children having similar performances.

Further comparison of the IF values obtained for both elicitation method revealed that 91% (21 out of 23) of the verbs had a higher IF value when elicited using the dynamic media with increments ranging from 0.019 to 0.409 ($M = 0.207$). These findings again suggested a general trend of better performance when verbs were elicited using dynamic media.

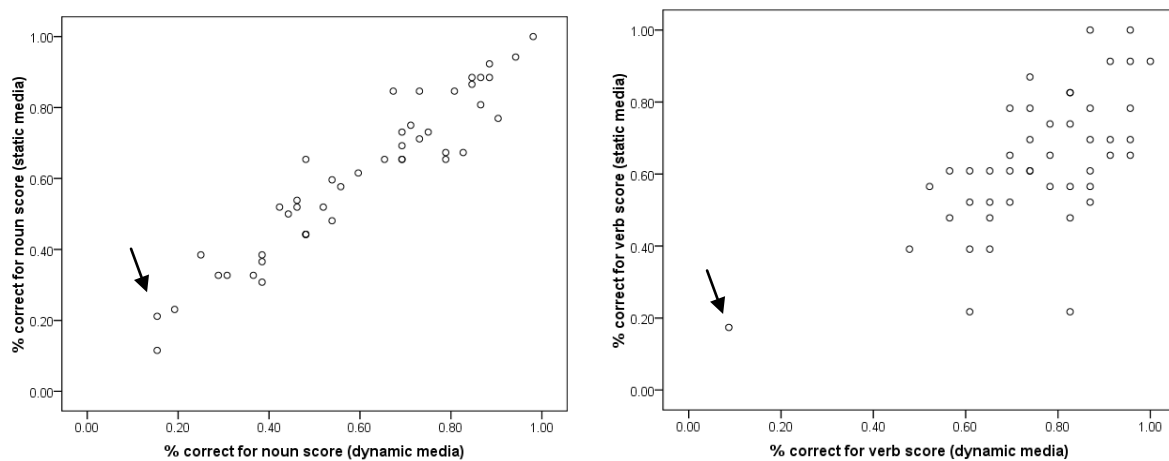


Figure 3. Children's performance in nouns (left) and verbs (right) in two elicitation methods.

Item analysis

Since the children's performance in verb production was significantly better when dynamic media was used, this prompted for a review of the final 57 word list suggested by Lam (2008). Similar to Lam's (2008) study, besides item facility (IF), item discrimination (ID), an index of the efficiency of an item in differentiating participants in the behaviour

being investigated (Anastasi, 1982), was also calculated (Appendix E).

Recall that In Lam's (2008) study, items with IF value falling between 0.15 and 0.85 and ID value larger than 0.25 were considered for inclusion in the final word list. This criterion is based on the consideration that an item with high IF value may be too easy since most participants answer it correctly while an item with low IF value may be so difficult that only a few participants get it correct (Oller, 1979). Such items are not suitable for use in a norm-referenced test, since there is too much variability of performance among the participants (Oller, 1979). On the other hand, items with low ID values suggest that both the high and low scorers perform similarly on those items and hence they are not discriminative of participants who performed better or worse. For items with high ID values, it suggests that there are significantly more high scorers than low scorers who score the items correct and hence they are able to discriminate between participants with good or poor performance.

In this study, similar to Lam's (2008) study, the statistic program SPSS 16.0 was used for the computation of ID values by using the point-biserial correlation technique. The IF and ID values of verbs elicited using both dynamic and static medias were calculated (Appendix F). Using Lam's (2008) criteria for inclusion, the acceptable range of both IF and ID was found for nine verbs elicited by the dynamic media and eleven verbs elicited by the static media. The discrepancy was reviewed and it was found that the verbs which met the selection criteria in one media but not another was due to either high IF value or low ID value, suggesting the children having better performance or the items being less discriminative.

Error analysis

A review of the error patterns indicated that the children made primarily substitution errors including 1) substitution of verbs for nouns; 2) substitution of attributes for nouns and 3) substitution of nouns for verbs. A significant amount of non-target verbs errors was also made.

Errors in nouns

Some of the non-target errors produced by the children in this study were substitution of verbs for nouns. Out of 49 children, three produced “剪剪” (cut) for “較剪” (scissors) and one produced “擦” (rub) for “擦膠” (rubber). In addition, “游水” (swim) was being elicited for “泳池” (swimming pool) and “沙灘” (beach) in 22 children, suggesting a tendency of children producing object function in replacement of its name. Others produced attribute of an object instead of its name, such as “圓轆轆” (round) being produced for “車轆” (car wheel). More common errors were listed in Appendix G.

Another interesting finding was the production of “tap” which can be expressed in Cantonese as either “水喉” (“water-pipe”) or “水龍頭” (“water-dragon-head”). Four children expressed “水頭” (“water-head”), which may be a combination of the two possible expressions, suggesting that children’s learning may be partial with different exposures or may be just an omission of the middle syllable of “水龍頭” (“water-dragon-head”).

Errors in verbs

Similar to errors in nouns, substitution of nouns for verbs was evident. One such error was the production of “擦膠” (rubber) for “擦” (rub) in eight children. Another common error was the production of non-target verbs a, such as “pAt1 “不” (scoop up) being elicited for “餵” (feed) when video of a child scooping with a spoon and putting the spoon into a man’s mouth was shown.

DISCUSSION

This study has achieved three outcomes. First, this study has successfully replicated Lam's (2008) study and again confirmed the usefulness of the chosen words for the assessment of expressive vocabulary in 3-year-old Cantonese-speaking children. Second, results from this study also suggested that the dynamic elicitation method enabled us to have a more accurate knowledge of 3-year-old children's vocabulary knowledge, especially for action verbs. However, there were shortcomings involving the use of video and careful consideration should be made during the choice of scenes. Third, nine verbs were identified in this study and proposed to be included in the revised final word list of the dynamic version of test for use in assessing the expressive vocabulary of 3-year-old Cantonese-speaking children in Hong Kong.

Lam's final word list

Since there was no significant difference found in the performance of Lam's (2008) final 57 word list between the children in Lam's (2008) study and the current study with a high correlation obtained, Lam's (2008) study was reliable and replicable with its clinical application in future substantiated.

Static media vs dynamic media

One of the aims of this study was to examine whether elicitation methods would affect 3-year-old children's performance in an expressive vocabulary test, especially for word type that is not static in nature. Since video clips are able to capture the dynamic quality of actions, and thus able to present the most prominent characteristics of verbs, results showed that 3-year-old children performed significantly better when verbs were elicited using dynamic media. From this, the relatively poorer performance of the children in Lam's (2008) study in verb production could be accounted for. Their poor performance may be due to the fact that static media was unable to capture the prominent characteristic of verb and leading

to confusion and non-target responses. As a result, the use of static media might have underestimated the actual knowledge of some children, which also explains the relatively small proportion of verb items (10%) in Lam's (2008) final 57 word list. In conclusion, static media, which cannot capture the dynamic quality of verb, may not be the most optimal elicitation method for verb items.

Therefore, when developing an expressive vocabulary test, not only should we consider the difficulty of the words we want to test the children, we should also consider the nature of the words and the method of elicitation. Thus, with alternative elicitation method, such as the use of dynamic media for eliciting verbs, children's vocabulary development can be better captured and underestimation prevented. Besides, as mentioned in the results section, significant changes were found only in the children's performance in verb production across the two medias. Since this is a repeated measure study, the differences obtained was unlikely a result of group differences, which further supported the effect of elicitation method on word type.

However, dynamic media is not all advantages without any shortcomings. Although computer administration of the expressive vocabulary test is likely to tap on children's actual knowledge, especially for action verbs, technical errors can sometimes cause problems. One such error is the production of non-target words for colour names, with 82% (28 out of 34) of the children produced "purple" correctly with static media but not dynamic media. Hence, colour names are very vulnerable to off-target responses as different computers have slightly different colour tones and they should be administered with careful considerations. However, this problem is relatively simple and can be easily resolved by a computer expert.

From the error analysis, we recognized that common errors were substitutions. This suggested tendency of some children to replace the target word with related features when they didn't know the word. It may also be possible that the errors were not true substitution

errors but simply reduplicated productions or confusions aroused during early learning. This is because in Cantonese, not only does verb sometimes constitute part of a noun, such as “剪” (cut) in “較剪” (scissors) and “擦” (rub) in “擦膠” (rubber), there was also a tendency of young Cantonese-speaking children to reduplicate their productions, such as producing “剪剪” (cut-cut) for “較剪” (scissors) and “擦擦” (rub-rub) for “擦膠”. On the other hand, for substitution of noun in verb, it may be the tendency of children naming the object seen in the picture or video when they didn’t know the action name.

As mentioned before, dynamic media is not all advantages without any shortcomings. One prominent error found in the elicitation with dynamic media which appeared only occasionally in the static media was the elicitation of non-target verbs. The example of “pAt1 不” (scoop up) being elicited for “餵” (feed) mentioned before was not completely incorrect however. This is because the action “pAt1 不” (scoop up) was indeed being presented in the earlier part of the video clip before “餵” (feed) even though it was not the target word. Hence, this non-target verb error was indicative that certain children extracted the earlier part of the scene (child scooping with a spoon) for production instead of the later part (putting the spoon into a man’s mouth).

This led us to think about the difficulty in verb learning. Verb learning is complicated and challenging as it involves identification of the part of the scene that represents the verb as compared to the relatively easier noun learning of names of object. Thus, errors in verbs may be due to children having identified the wrong scene during learning or extracted the wrong scene during production. Hence, when dynamic stimuli are used, one has to be careful in the selection of the scene to avoid any misinterpretation.

Although we can do our best to control the scene of the action for eliciting the verb, we are unable to control how any child extracts the scene. Some may argue that the scene

could be cut so precisely that only a single verb is illustrated, like showing only the “feeding” scene without the “scooping” scene when eliciting “餵” (feed). However, some verbs occur together with other verbs and separation is deemed impossible. One example would be the elicitation of “夾” (pick up something using a tool) using a scene which shows a person holding (揸住) a pair of chopsticks and picking up (夾) a piece of food. It is almost impossible to separate “holding” (揸住) and “picking up” (夾) as they happen simultaneously and one has to first “hold” the chopsticks before being able to use it to “pick up” the food. Thus, we can only do our best to minimize any confusion but it is impossible to control how children interpret the scene.

Even though verbs are difficult to elicit and much problem arises during their elicitation, we should not compromise and exclude the assessment of verbs in an expressive vocabulary test. This is because verbs are developmentally appropriate for children and they may be more discriminative than nouns in identifying children who may have language impairment. As revealed by the scatter plot showing the distribution of the children’s scores on the two word types under the two elicitation methods (Figure 3), there was one child whose score was way apart from the rest of the group in verb production but his poor performance in nouns was shared by several other children. This particular child was later identified to be the child who was reported to have delayed speech and language development and was receiving speech and language therapy service at the time of testing. This seemed to bring in light that children who have speech and language impairment may be more easily identified with their relatively weaker verb production as compared to noun production. These children may otherwise be masked if expressive vocabulary test involving primarily nouns were used. Thus, the inclusion of more verbs may allow Lam’s (2008) expressive vocabulary test to be more sensitive in discriminating weaker children who have speech and language impairment from children who are within the normal limits.

Inclusion of more verb items

Besides being developmentally appropriate and allowing the test to be more discriminative, inclusion of more verbs also serves to resemble the proportion of verbs expressed by 3-year-old children (Tse et al., 2005). From the results, 91% (21 out of 23) of the verbs obtained a higher IF value when elicited by the dynamic media than when elicited by the static media. Since the dynamic media was proposed for use in future, we shall consider the inclusion of more verbs in the dynamic version of test. With Lam's (2008) criteria for inclusion, there were a total of nine verbs with acceptable IF and ID values and these items were proposed for inclusion, resulting in a total of 61 words, constituting of 52 nouns and 9 verbs (Appendix H). The inclusion of more verbs was also in view of the discrepancy in the proportion of verbs found between the test and the actual production of 3-year-old children (Tse et al., 2005). With inclusion of more verbs, the proportion of verbs was then increased to 14.75% which has a greater resemblance as the speech sample obtained by Tse et al. (2005).

The items recommended for inclusion in the dynamic version of test, such as “擦” (rub), “扭” (twist), “攪” (stir) and “推” (push), were mainly verbs that depict prolonged dynamic quality or repetitive movements. This may account for their inclusion as the use of dynamic media could capture the repetitive nature of an action in addition to its dynamic quality. For verbs that were excluded in this study, they were generally too easy and at the same time, they failed to discriminate between the weaker children from the stronger ones. These excluded verbs seemed to include actions that children might perform or encounter daily, such as “梳” (comb), “沖涼” (bath), “斟” (pour), “吹” (blow) as in “吹泡泡” (blow bubbles), “游水” (swim), “跑” (run) and “剪” (cut). Hence, familiarity and exposure probably contributes to children's knowledge in expressive vocabulary.

To conclude, there are several implications for the inclusion of more verbs in a

standardized expressive vocabulary test for 3-year-old Cantonese-speaking children. First, the existence of a child having an extreme score in verb production but not in noun production suggested that verb production might be more discriminative than noun production in identifying children with language impairment. Secondly, as mentioned in the introduction, children with SLI has specific deficit in verbs and this would also advocate the inclusion of more verb items in any expressive vocabulary test.

Analyzing children's production error

Since 3-year-old children are at the peak of their language development, their vocabulary increases everyday. Errors should not be interpreted discretely as right or wrong but should be analyzed as they are revealing of the children's vocabulary knowledge and their language development. By analyzing the errors children made in their production, clinicians would gain a better understanding of their language development and the accompanying deficits. At the same time, knowledge of error patterns may reveal certain underlying language deficits and thus facilitate intervention planning and development of strategies to help the children during therapy.

CLINICAL IMPLICATION

Lam's (2008) 57 final word list was found to be reliable and replicable when tested on three-year-old Cantonese-speaking children in Hong Kong. The administration of the test and the scoring criteria is simple and extensive training is not required. Also, the duration of the test is relatively short (15 minutes) and therefore the test is practical for clinical usage, serving as a tool for clinicians to identify children with speech and language impairment. Nevertheless, as suggested by the findings of this study, the final word list used in the dynamic version of test was proposed to include a total of 61 words, consisting of 52 nouns and 9 verbs (Appendix H), so that a more comprehensive and accurate representation of children's language ability can be obtained.

As mentioned, verbs are especially problematic for children with specific language impairment (Fletcher, 1999). Hence, inclusion of more verbs has its significance in discriminating children with specific language impairment (SLI) at a young age of three. This implies early intervention possible which would facilitate the remedial of their speech and language development. However, we should also pay attention to parents' report and teachers' opinions of a child's speech and language development before making a diagnosis. This is due to the fact that children at such a young age may not be able to perform optimally in a structured clinical setting, such as during administration of standardized tests, resulting in poor performance and underestimation of their language ability.

From the result of this study, dynamic media (dynamic video clips) was more effective in eliciting verb items from 3-year-old Cantonese-speaking children than static media (picture cards). Thus, dynamic media of testing was proposed to be used in future clinical setting so as to achieve a more accurate understanding of children's expressive vocabulary development. Given the advancement of technology in this digital era with computers being readily available, it is possible for digital versions of assessment tools to be

administered readily. Clinicians can run an expressive vocabulary test using a computer in the clinics or run the test using mini laptops and PDAs (Personal Digital Assistant). With the advantage of portability provided by such devices, it would not be so much of a hassle even if clinicians have to carry the test around. Hence, the feasibility of running the digital version of expressive vocabulary test with dynamic stimuli is high and should be implemented.

However, caution is warranted in using the set of dynamic stimuli from this study. As mentioned, technical errors should be resolved before its application. Furthermore, some stimuli were believed to have been illustrated poorly, leading to the production of non-targets. One such example was “中間” (middle) as “車前面” (in front of the car) was always elicited instead. The stimulus picture showed three boys standing in front of a car and the question “Here are three boys in the picture, where does this boy (clinician points to the boy standing at the middle) stand?” was asked. As the boy in the picture concurrently stands in the middle and in front of the car, resulted in either position being interpreted and produced by the children. Thus, target stimuli should be reviewed and illustrated more clearly to obtain more reliable responses in future standardization of the test.

LIMITATIONS AND DIRECTIONS FOR FUTURE RESEARCH

Despite of the data collected in the current study, the sample pool remained limited with a total of fewer than 90 participants. Thus, extensive testing to increase the normative data pool is necessary in future research to verify the reliability and validity of Lam's (2008) study. In addition, although the age effect was not significant in this study, the result might not hold as there were only 24 participants in each age range and this sample size is relatively small. Hence, further investigation involving more participants in each age range or having a narrower age range should be carried out in future to confirm the finding of this study. While age effects within the 3-year-old were not significant, preliminary evidence from two 2-year old children confirmed that the word list was appropriate for three- but not two-year-old children. In collection of pilot data on children younger than three years of age, the final 57 word list was given to two 2-year-old children as part of this study and both of them scored correct for 20 items or less. There was huge difference in performance between them and the 3-year-old children ($M = 30$ for Lam (2008), $M = 34$ for the current study), with the 3-year-old obtaining over 50% more correct. However, the data obtained are limited and hence, future research may include both the 2-year-olds and 4-year-olds to further investigate the suitability of using this expressive vocabulary test on children of different age range.

REFERENCES

- Anastasi, A. (1982). *Psychological Testing*. New York: Macmillan.
- Bailey, K. M. (1998). *Learning about Language Assessment: Dilemmas, Decisions, and Directions*. Pacific Grove, [Calif.]: Heinle & Heinle Publishers.
- Bornstein, M. H. & Cote, L. R., with Maital, S., Painter, K., Park, S., Pascual, L., Pêcheux, M., Ruel, J., Venuti, P. and Vyt, A. (2004). Cross-Linguistic Analysis of Vocabulary in Young Children: Spanish, Dutch, French, Hebrew, Italian, Korean, and American English. *Child Development, 75(4), 1115 – 1139*
- Brownell, R. (ed.). (2000). *Expressive One-word Picture Vocabulary Test (EOWPVT) (3rd ed.)*. Novato, Calif.: Academic Therapy Publications.
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Lawrence Earlbaum Associates.
- Cuetos, F. & Alija, M. (2003). Normative data and naming times for action pictures. *Behavior Research Methods, Instruments, & Computers, 35(1), 168-177*.
- Dromt, E. (1999). Early lexical development. In Barrett, M. D. (eds). *The Development of Language (pp349 -372)*. Hove, East Sussex, UK : Psychology Press
- Fletcher, P. (1999). Specific Language Impairment. In Barrett, M. D. (eds). *The Development of Language (pp349 -372)*. Hove, East Sussex, UK : Psychology Press
- Guilford, J.P. (1956). *Fundamental Statistics in Psychology and Education*. New York: McGraw Hill.
- Lam, T. K. (2008). *Developing an Expressive Vocabulary Test for Three-year-old Children in Hong Kong. Unpublished BSc Honours Thesis*. Division of Speech and Hearing Sciences, The University of Hong Kong.

- Loeb, D.F., Pye, C., Redmond, S., & Richardson, L. Z. (1996). Eliciting Verbs from children with specific language impairment. *American Journal of Speech-Language Pathology*, 5, 17-30.
- Oller, J. W., Jr. (1979). *Language Tests at School*. London: Longman.
- Opper, S. (1996). *Hong Kong's Young Children: Their Early Development and Learning*. Hong Kong: Hong Kong University Press.
- Paul, R. (2007). *Language Disorders from Infancy through Adolescence: Assessment and Intervention* (3rd ed.). St. Louis, Mo: Mosby Elsevier.
- Reynell, J., & Huntley, M. (1987). *Reynell Developmental Language Scales—Cantonese* (Hong Kong version). Windsor, England: NFER-Nelson.
- Tarif, T., Gelman, S. A. & Xu, F. (1999). Putting the “Noun Bias” in Context: A Comparison of English and Mandarin. *Child Development*, 70 (3), 620–635
- Tse, S. K., Chan, C. and Li, H. (2005). Is the expressive vocabulary of young Cantonese speakers noun or verb dominated? *Early Child Development and Care*, 175(3), 215–227.
- Tse, S. K., Chan, K. K. & Kwong, S. M. (2006). *香港幼兒口語發展*. Hong Kong: Hong Kong University Press.
- Zechmeister, J. S., Zechmeister, E. B., & Shaughnessy, J. J. (2001). *Essentials of research methods in psychology*. New York, NY: McGraw-Hill.

Appendix A

Lam (2008)'s 57 final word list

Total: 40 nouns, 5 verbs (boxed), 12 descriptors

Item	IF value	ID value	Item	IF value	ID value
手指	0.20	0.366	蓋	0.26	0.246
瘦	0.20	0.460	海	0.27	0.351
中間	0.20	0.471	數目字	0.29	0.476
電掣	0.26	0.308	檸檬	0.29	0.383
直昇機	0.26	0.433	廚師	0.31	0.487
泳池	0.26	0.476	食物	0.34	0.393
膠布	0.43	0.448	彩虹	0.34	0.557
高	0.46	0.409	扭	0.37	0.44
間尺	0.46	0.324	玩具	0.40	0.542
粉紅色	0.46	0.320	顏色	0.40	0.499
紫色	0.49	0.667	風箏	0.41	0.413
蕃茄	0.49	0.320	水喉	0.43	0.608
啡色	0.49	0.593	動物	0.43	0.660
沙灘	0.49	0.538	脞	0.60	0.468
窗	0.51	0.250	夾	0.60	0.301
尾	0.51	0.334	長頸鹿	0.60	0.507
消防員	0.51	0.506	太陽	0.60	0.263
菠蘿	0.54	0.311	砌	0.60	0.305
電單車	0.54	0.311	凍	0.60	0.391
三角形	0.57	0.310	轆	0.63	0.359
葉	0.57	0.619	生果	0.63	0.502
寫	0.74	0.363	洗衣機	0.66	0.362
拖鞋	0.74	0.296	紅色	0.66	0.318
心形	0.77	0.564	短	0.66	0.582
滑梯	0.77	0.559	擦膠	0.66	0.344
蛇	0.8	0.373	禮物	0.69	0.448
鈕	0.80	0.515			
多	0.80	0.494			
較剪	0.83	0.292			
鏡	0.83	0.403			
切	0.83	0.269			

Appendix B

Verbs excluded from Lam (2008)'s 57 final word list

Total: 18

Item excluded due to
high IF value (<0.15 / >0.85) or low ID value (<0.25)

Item	IF value	ID value
吹	0.86	-0.032
跑	0.86	0.130
沖涼	0.89	0.276
飛	0.91	-0.058
游水	0.94	0.203
擦	0.94	0.546
餵	0.51	0.003
斟	0.57	0.166
推	0.57	0.195
錫	0.83	0.036
剪	0.83	0.164
熨	0.29	0.207
開	0.31	0.121
抹	0.34	0.168
摺	0.34	0.155
攪	0.37	0.028
煮	0.69	0.222
梳	0.71	0.021

Appendix C

The expanded 75 words list

Total: 52 nouns (nouns and descriptors) and 23 verbs (action verbs) (boxed)

1	脞	41	煮
2	蛇	42	擦膠
3	錫	43	泳池
4	洗衣機	44	扭
5	游水	45	食物
6	掣	46	熨
7	吹	47	鏡
8	直昇機	48	煲蓋
9	三角形	49	摺
10	沖涼	50	推
11	禮物	51	彩虹
12	紫色	52	瘦
13	菠蘿	53	心形
14	飛	54	切
15	紅色	55	中間
16	夾 (/gaap3/)	56	粉紅色
17	電單車	57	葉
18	高	58	餵
19	寫	59	剪
20	數字	60	動物
21	跑	61	啡色
22	玩具	62	鈕
23	窗	63	風箏
24	擦	64	開
25	間尺	65	多
26	長頸鹿	66	梳
27	尾	67	膠布
28	鞭	68	生果
29	短	69	攪
30	斟	70	海
31	顏色	71	滑梯
32	蕃茄	72	水喉
33	拖鞋	73	檸檬
34	凍	74	沙灘
35	廚師	75	砌
36	抹		
37	消防員		
38	太陽		
39	手指		
40	較剪		

Appendix D

Instructions used in the assessment (Adapted from Lam (2008))

A detailed description of the instruction used during the testing was as followed.

“姐姐部電腦有啲短片，咁你就要睇吓啲短片，跟住話俾姐姐知睇到啲咩㗎。(There are some video clips shown on my computer. You will have to look at those video clips and tell me what each video clip shows.)”. Three trials of demonstration with feedback given was carried out, such as “係㗎，你有話俾姐姐你睇到啲咩㗎，係 xx [target word] 㗎。(Well done! You have told me what the video clip showed, it is xx [target word].)” or “唔啱㗎，你睇完之後冇話俾姐姐聽你睇到啲咩㗎。我哋再試多次，話俾姐姐聽你睇到啲咩? (That’s not right! You did not tell me what you saw from the video clip. Let’s try again, tell me what the video clip shows.)”. To elicit the participants for correct production, investigator asked “呢個咩嚟㗎?” for nouns, “個 xx (person) 做緊咩啊?” or “個 xx (object) 點㗎?” for verbs and “呢 xx (object/person) 好 xx (the contrastive word for the target word) , 咁呢個呢?” for elicitation adjectives (e.g. 呢個人好肥, 咁呢個呢?) .

The participants were given verbal reinforcement such as “好好 (good)”, “好俾心機 (Good effort!)” to facilitate the smooth running of the task. In addition, the participants were allowed to take breaks at anytime if fatigue or boredom is indicated. However, feedback of right or wrong was no longer provided. At the end of the session, participants were rewarded some stickers for their cooperativeness.

Appendix E

Calculation of item facility (IF) and item discrimination (ID) (adapted from Lam (2008))

Item facility (IF)

It is an index of the degree of difficulty of an item, and it involves calculating the percentage of participants who are correct on the item (Bailey, 1998).

Possible values of IF range from 0 to 1.

Item discrimination (ID)

It is an index of the efficiency of an item for differentiating the participants in the behaviour being investigated (Anastasi, 1982).

The ID of an item can be computed using the formula,

$$ID = \frac{(\text{no. of "high scorers" who got the item correct}) - (\text{no. of "low scorers" who got the item correct})}{27.5\% \text{ of the total number of participants}}$$

where "high scorers" and "low scorers" referred to the 27.5% of the participants who achieved the highest scores and the lowest scores in the entire test respectively.

Appendix F

Item facility and item discrimination values for verbs in different medias

Results from current study (static media)			Results from current study (dynamic media)		
Item	IF value	ID value	Item	IF value	ID value
錫*	0.79*	0.317*	錫*	0.75*	0.485*
游水	0.94	0.376*	游水	0.98	0.365*
吹	0.90	0.389*	吹	0.94	0.381*
沖涼	0.73*	0.218	沖涼	0.87	0.079
飛	0.90*	0.524*	飛	0.92	0.515*
(夾)*	0.23*	0.512*	(夾)*	0.40*	0.485*
(寫)	0.58*	0.198	(寫)	0.66*	0.126
跑	0.90	0.413*	跑	0.87	0.315*
擦*	0.67*	0.415*	擦*	0.75*	0.468*
斟*	0.69*	0.507*	斟	0.79*	0.178
抹	0.83*	0.199	抹	0.94	0.216
煮*	0.48*	0.310*	煮	0.72*	0.223
(扭)*	0.35*	0.437*	(扭)*	0.51*	0.449*
熨*	0.31*	0.390*	熨*	0.51*	0.291*
摺	0.46*	0.108	摺	0.72*	0.206
推*	0.60*	0.403*	推*	0.75*	0.384*
(切)	0.88	0.342*	(切)	0.96	-0.065
餵*	0.35*	0.479*	餵*	0.51*	0.410*
剪	0.94	0.154	剪	0.96	0.332*
開	0.21*	0.191	開*	0.62*	0.310*
梳	0.85	0.321*	梳	0.89	0.543*
攪*	0.31*	0.377*	攪*	0.51*	0.477*
(砌)*	0.71*	0.320*	(砌)	0.92	0.295*

* Acceptable IF or ID

() Items included in Lam's (2008) final word list

Appendix G

Common errors made by the children in the current study

Target	Children's Response	No. of such responses
較剪	剪剪	3
鈕	扣扣	2
擦膠	擦	1
泳池 / 沙灘	游水 / 沙灘	22
廚師	煮	15
水龍頭	洗	9
擦	擦膠	8
餵	pAt1 “不” (scoop up)	5
夾	揸筷子 / 食	33

Appendix H

The final 61 words list

Total: 52 nouns (nouns and descriptors) and 9 verbs (action verbs) (boxed)

1	脞	36	食物
2	蛇	37	熨
3	錫	38	鏡
4	洗衣機	39	煲蓋
5	掣	40	推
6	直昇機	41	彩虹
7	三角形	42	瘦
8	禮物	43	心形
9	紫色	44	中間
10	菠蘿	45	粉紅色
11	紅色	46	葉
12	夾	47	餵
13	電單車	48	動物
14	高	49	啡色
15	數字	50	鈕
16	玩具	51	風箏
17	窗	52	開
18	擦	53	多
19	間尺	54	膠布
20	長頸鹿	55	生果
21	尾	56	攪
22	鞭	57	海
23	短	58	滑梯
24	顏色	59	水喉
25	蕃茄	60	檸檬
26	拖鞋	61	沙灘
27	凍		
28	廚師		
29	消防員		
30	太陽		
31	手指		
32	較剪		
33	擦膠		
34	泳池		
35	扭		

ACKNOWLEDGMENT

I would like to express my deep gratitude to my dissertation supervisor, Dr. Anita Wong, for her guidance, stimulating comments and support throughout the study. I would also like to express my thanks to Miss Lam Tze Kam, author of the 'Developing an expressive vocabulary test for three-year-old children in Hong Kong' (2008), in sharing her data with me. In addition, I would also like to take this opportunity to thank all the teaching staff and clinical supervisors for their guidance and support in my four years of study.

I gratefully acknowledge the principals, teachers, staffs, parents and children in the following schools for their assistance and cooperation during data collection:

1. St. Margaret Mary's Catholic Kindergarten
2. St. Stephen's Catholic Kindergarten
3. Kam Lam Catholic Kindergarten

Last but not least, I would like to thank my family members, classmates and friends for their encouragement and support.