

1986

A comparative study of three language sampling methods using developmental sentence scoring

Cheryl Diane Dong
Portland State University

Follow this and additional works at: https://pdxscholar.library.pdx.edu/open_access_etds



Part of the [Speech and Rhetorical Studies Commons](#), and the [Syntax Commons](#)

Let us know how access to this document benefits you.

Recommended Citation

Dong, Cheryl Diane, "A comparative study of three language sampling methods using developmental sentence scoring" (1986). *Dissertations and Theses*. Paper 3589.
<https://doi.org/10.15760/etd.5473>

This Thesis is brought to you for free and open access. It has been accepted for inclusion in Dissertations and Theses by an authorized administrator of PDXScholar. Please contact us if we can make this document more accessible: pdxscholar@pdx.edu.

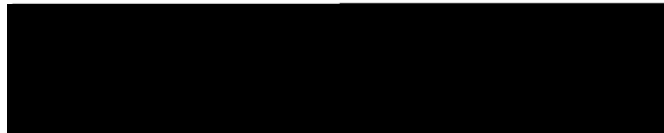
AN ABSTRACT OF THE THESIS OF Cheryl Diane Dong for the Master of Science
in Speech Communication with an emphasis in Speech-Language Pathology
presented June 3, 1986.

Title: A Comparative Study of Three Language Sampling
Methods using Developmental Sentence Scoring

APPROVED BY MEMBERS OF THE THESIS COMMITTEE:



Robert Casteel, Chairman



Mary Gordon

Oral language sampling has been used as a clinical and research
tool to describe a child's expressive language skills. Through the
years, many methods have been presented with which to score or analyze
utterances produced by a child during spontaneous conversation.

One method of analyzing a child's use of syntax, Developmental Sentence Scoring (DSS) by Lee and Canter (1971) was the focus of this study. The study which provided the normative data only revealed that toys, pictures, and stories were used to elicit the language samples. A review of the literature of oral language sampling revealed that different stimulus materials produce differing effects on the expressive output of children.

Therefore, the present study sought to determine the effect different stimulus material has on the language elicited from children. Its purpose was to determine whether a significant difference existed among language samples elicited three different ways when analyzed using DSS. Eighteen children between the ages of 3.6 and 5.6 years were chosen to participate in the study. All of the children had normal hearing, normal receptive vocabulary skills and no demonstrated or suspected physical or social delays. Three language samples, each elicited by either toys, pictures, or stories, were obtained from each child. For each sample, a corpus of 50 utterances was selected for analysis and analyzed according to the DSS procedure as described by Lee and Canter (1971).

The means and standard deviations of the DSS scores were calculated for the samples elicited by each stimulus material. A two-tailed t - test for related means was computed to determine if a statistically significant difference exists among the language samples which were elicited by toys, pictures, or stories.

Results of this study showed that storytelling consistently elicited the higher mean DSS score, followed by, in descending order, toys and pictures. Analysis of results indicated that using stories

($\bar{x} = 8.90$) to elicit language produced a significantly higher mean DSS score than when pictures ($\bar{x} = 6.87$) were used to elicit language samples ($t = 3.42, p < .01$). A difference trending toward significance was also obtained between samples elicited by toys ($\bar{x} = 8.01$) and those elicited by pictures ($\bar{x} = 6.87$). As each child served as his or her own control and the same examiner and procedures were used to obtain each sample, it is proposed that these differences were primarily related to the stimulus materials used to elicit the language samples.

A COMPARATIVE STUDY OF THREE LANGUAGE SAMPLING METHODS
USING DEVELOPMENTAL SENTENCE SCORING

by

CHERYL DIANE DONG

A thesis submitted in partial fulfillment of the
requirements for the degree of

MASTER OF SCIENCE IN SPEECH COMMUNICATION
with an emphasis in

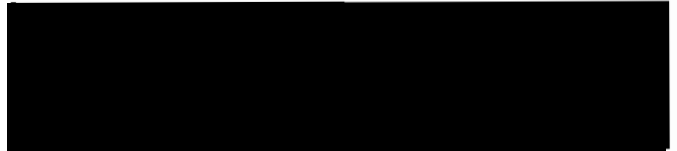
SPEECH-LANGUAGE PATHOLOGY

Portland State University

1986

TO THE OFFICE OF GRADUATE STUDIES AND RESEARCH:

The members of the Committee approve the thesis of Cheryl Diane Dong presented June 3, 1986.

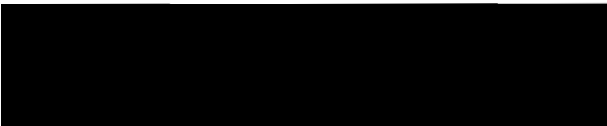


Robert Casteel, Chairman

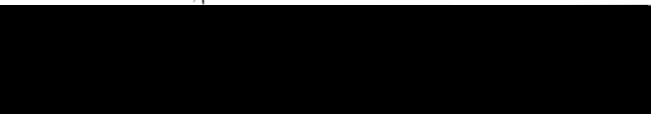


Mary Gordon

APPROVED:



Theodore G. Grove, Head
Department of Speech Communication



Bernard Ross, Dean
Graduate Studies and Research

DEDICATION

This manuscript is lovingly dedicated in memory of my mother,
Akiko Dong.

ACKNOWLEDGEMENTS

Leo Buscaglia writes

Nikos Kazantzakis suggests that ideal teachers are those who use themselves as bridges over which they invite their students to cross, then having facilitated their crossing, joyfully collapse, encouraging them to create bridges of their own.

To one of the "ideal teachers," Dr. Robert Casteel, I give my gratitude for all the time, energy and knowledge you shared with me. I feel very fortunate to have studied under you and as I go out "building bridges," I shall remember you and your teachings.

To Mary Gordon, I extend a special thank you for your time and knowledge. You were always willing to read, critique and assist me in the preparation of this document. Your assistance has been deeply appreciated.

Thank you also to Dr. Roger Jennings for his participation in the thesis orals. Your questions were thought-provoking and your suggestions were helpful and constructive.

To the staff of the Helen Gordon Child Development Center, the YMCA Child Development Center and the parents of the subjects, I extend a heartfelt thank you. Without your cooperation, this study could not have been successfully completed. A warm thank you to all the children who participated in this study. You were all delightful and I enjoyed every one of you!

Special thanks to:

Kathy "Ding" Zuehlsdorf who has been such a great friend to me and helped to make graduate school a fun place to be.

Cathy Pew who has provided much needed emotional support and friendship through the past two years.

Lisa Whitney who, as always, believed in me when I didn't believe in myself. You have been and always will be such an important part of my life.

Jill Gordon who had the "privilege" of living with a graduate student, entitling her to witness a wide variety of mood swings, paper and books scattered throughout the house, all night typing, and housework duties performed according to my homework schedule. Your stable influence, optimistic outlook and unfailing support made my life that much easier.

Lesta Maddock for being such a true friend. Thank you for being such an inspiration to me.

Lastly and most importantly, thank you to my family for providing me with the opportunity and emotional support with which to complete my professional training. Your love and encouragement has been deeply appreciated.

TABLE OF CONTENTS

| | PAGE |
|---|-------|
| DEDICATION | iii |
| ACKNOWLEDGEMENTS | iv |
| LIST OF TABLES | .viii |
| CHAPTER | |
| I INTRODUCTION AND STATEMENT OF PURPOSE | 1 |
| Introduction | 1 |
| Statement of Purpose | 4 |
| II REVIEW OF THE LITERATURE | 5 |
| Analysis of Oral Language Sampling | 5 |
| Mean Length of Response | 5 |
| Mean Length of Utterance | 6 |
| Length-Complexity Index | 6 |
| Type-Token Ratio | 6 |
| Empirical Data | 6 |
| Developmental Sentence Scoring | 7 |
| Uses of Developmental Sentence Scoring | 7 |
| Effect of Different Stimulus Materials on the Elicitation of Oral Language Samples | 10 |
| III METHODS | 18 |
| General Plan | 18 |
| Subjects | 18 |
| Sampling Method | 19 |
| Screening | 19 |
| Instrumentation | 18 |
| Examiner Reliability | 20 |
| Experimental Procedures | 21 |

| CHAPTER | PAGE |
|---|------|
| Language Sample Collection | 21 |
| Language Sample Transcription | 23 |
| Scoring | 24 |
| Analysis of Data | 25 |
| IV RESULTS AND DISCUSSION | 26 |
| Results | 26 |
| Discussion | 27 |
| Stimulus Material Effects on Elicited Language | 28 |
| Possible Variables Affecting Outcome of Study | 31 |
| Type of Language Elicited | 31 |
| Elicitation, Transcription, Analysis and Order Variables | 35 |
| Effects of Stimulus Materials on DSS Scores | 36 |
| Past Studies Using DSS | 39 |
| Use of DSS as a Clinical and Research Tool | 41 |
| V SUMMARY AND IMPLICATIONS | 45 |
| Summary | 45 |
| Implications | 46 |
| Clinical Implications | 46 |
| Research Implications | 49 |
| SELECTED BIBLIOGRAPHY | 51 |
| APPENDICES | 55 |

LIST OF TABLES

| TABLE | | PAGE |
|-------|--|------|
| I | A Comparison of DSS Means Obtained Using Three Different Stimulus Materials | 27 |
| II | DSS Scores, Percentiles, and Means by Session | 37 |

CHAPTER I

INTRODUCTION AND STATEMENT OF PURPOSE

Introduction

The most important single factor used to evaluate a child's growth and development is oral language (Longhurst and Schrandt, 1973). Therefore, it is of utmost importance that the speech-language clinician be able to describe a child's language adequately. An adequate description of language, as defined by Peterson and Marquardt (1981),

should define the speech and language skills observed, judge the communication ability . . . and make obvious a plan of action for remediation if the pattern presented warrants it.

To aid the speech-language clinician in describing a child's language skills, many formal, standardized tests are available which provide information such as vocabulary comprehension and grammar usage. In addition to the administration of formalized tests, the speech-language clinician may choose to assess a child's language skills by taking a language sample, defined by McLean and Snyder-McLean (1978) as a "verbatim, transcribed record of all utterances produced by a child within a given situation over a period of time." This record may then be analyzed or scored according to a variety of methods. McLean and Snyder-McLean stated that language sampling may

provide the clinician with a much more complete picture of a child's natural expressive language performance than that

obtainable from any standardized instrument.

Developmental Sentence Scoring (DSS), developed by Lee and Canter (1971), is one method of analyzing a tape-recorded sample of a child's language by making a detailed, readily quantified and scored evaluation of a child's use of standard English grammatical rules. The sample consists of 50 consecutive complete utterances. These utterances are written on a record sheet and given a weighted score for each of eight grammatical categories. Additionally, a sentence point is awarded if the utterance is grammatically complete and correct by adult standards. Total points for the 50 utterances are summed and divided by 50, resulting in a DSS score which can then be compared with normative data for children aged 2.0 years through 6.11 years (Lee, 1974a).

While DSS provides a detailed account of the method for transcribing, analyzing, and scoring the language sample, the guidelines are not as precise when describing the procedure to be used to elicit the language sample. Past researchers have found that there are many variables which may affect the elicitation of a language sample. Barrie-Blackley, Musselwhite, and Rogister (1978) listed the following variables which may affect the elicitation of a language sample: subject population, rapport-building techniques, examiner, situation or site, stimulus, instructions, consequences, and language task. Longhurst and Grubb (1974) stated that a major consideration of language sampling is the possibility that different collection procedures may produce measurable differences in the respondent's language. These differences may cause the speech-language clinician to under- or overestimate the child's knowledge of language structure.

The effect of different collection procedures used to elicit

oral language has been the focus of research studies. In two studies (Ahmed, 1973; Cowan, Weber, Hoddinott, and Klein, 1967), different pictures were found to affect the language sample elicited. Rogister (1975) found that telling a familiar story elicited more complex language and more verb usage than telling an unfamiliar story. It appears that the collection procedures used have an effect on the amount and complexity of the language produced.

DSS guidelines for selecting stimulus materials to be used in the elicitation of language state that stimulus materials used should be appropriate (Lee, 1974a). In the study by Koenigsknecht (1974) which provided the normative data for DSS, the stimulus materials consisted of miniature toys and figures, sets of pictures, and a familiar nursery story. These materials were presented in the above order and were used to elicit conversational speech from the children in the study. The last 50 utterances of the session were selected for scoring and analysis.

While DSS specified the above three tasks for eliciting oral language, it did not specify from which task the utterances for analysis should be selected. DSS normative data, therefore, according to Barrie-Blackley et al. (1978), "may be drawn from only one of the task segments . . . or a combination of tasks." This would appear to affect the reliability of DSS.

Spector (1981) referred to reliability as the "consistency of a measuring device." Reliability provides assurance that results from one assessment will be comparable to results from a second assessment. In order to achieve adequate reliability, standardization of procedures is needed. Standardization of administration, according to Davis

(1983), "maximizes the likelihood that the test is given in the same way each time so that . . . performance is assessed according to the same criteria each time." In order to achieve adequate standardization, the instructions for test administration must be explicit and comprehensive, and the examiner must follow the instructions exactly. It appears that a need exists to examine DSS administration procedures, particularly stimulus materials, in order to ensure adequate reliability.

Statement of Purpose

The purpose of this study was to compare Developmental Sentence Scores (Lee and Canter, 1971) of language samples from children which were elicited with three different types of stimulus materials.

The question this investigator sought to answer was the following: Does a significant difference exist among language samples elicited using three different stimulus materials (toys, pictures, storytelling) from children aged 3.6 years through 5.6 years when analyzed utilizing the Developmental Sentence Scoring (Lee and Canter, 1971) procedure?

CHAPTER II

REVIEW OF THE LITERATURE

Oral language sampling is a useful tool which the speech-language clinician may use to provide a detailed description of a child's expressive language performance. This review of the literature will present some of the various methods of analysis which are available to the speech-language clinician. As Developmental Sentence Scoring (DSS), developed by Lee and Canter (1971), was the focus of this study, the uses of DSS will be discussed. Studies showing the effect of different collection procedures upon the language sample elicited will be detailed in addition to the collection procedures used in DSS.

Analysis of Oral Language Sampling

The speech-language clinician may choose to assess a child's language skills by obtaining a language sample, defined by McLean and Snyder-McLean (1978) as a "verbatim, transcribed record of all utterances produced by a child within a given situation over a certain period of time." This record may then be analyzed or scored according to a variety of methods.

Mean Length of Response

In 1925, Nice suggested using the length of a child's response as a means of evaluating speech development. Mean Length of Response (MLR) is calculated by collecting 60 responses per child, discarding the

first 10 and analyzing the remaining 50 responses by totaling the number of words and dividing by 50. McCarthy (1954) developed normative data for children at six month age separations from 18 to 54 months. Templin (1957) also reported MLR norms for children from 3.0 to 8.0 years of age.

Mean Length of Utterance

Mean Length of Utterance (MLU) is calculated by collecting 100 responses per child, totaling the number of morphemes and dividing by 100. McLean and Snyder-McLean (1978) stated that MLU may be used as an indicator of the child's relative stage of linguistic development. Brown (1973) defined five stages of development of multi-word utterances; these stages were defined by the child's MLU.

Length-Complexity Index

Length-Complexity Index (LCI) was presented by Shriner and Sherman in 1967 and uses a numerical weighting system to analyze a child's response. This index results in a composite analysis of sentence length and complexity.

Type-Token Ratio

Johnson (1944) used Type-Token Ratio (TTR) to analyze oral language samples. This measure calculates vocabulary usage by counting each new word (type) and each additional use of that word (token). The number of tokens are then divided into the number of types, resulting in the TTR.

Empirical Data

Various types of empirical data have been used in the literature

to chart a child's progress in speech development. These include total number of words, mean of the five longest responses, number of one word responses, total number of words, proportion of total utterances which are sentence fragments, number of transformations and adverbial expansions, and different sentence relationships (Johnson, Darley, and Spriestersbach, 1963; Stalnaker and Creaghead, 1982).

Developmental Sentence Scoring

Lee and Canter (1971) described this method which specifies the developmental level of a child's syntax by analyzing a tape-recorded spontaneous speech sample. The corpus consists of 50 complete, different, consecutive, non-echoic utterances. Each of these utterances are assigned a weighted score for each of the following eight grammatical categories: indefinite pronouns and/or noun modifiers, personal pronouns, main verbs, secondary verbs, negatives, conjunctions, interrogative reversals, and wh-questions. If the utterance is grammatically correct according to adult standard English, the utterance is awarded one extra point. Total points for the 50 utterances are summed and divided by 50, resulting in a DSS score which can be compared with normative data for children from 2.0 years through 6.11 years (Koengis-knecht, 1974).

Uses of Developmental Sentence Scoring

DSS has been used in various studies as a means of differentiating the language behavior of certain groups of children. In a study conducted by Kramer, James, and Saxman (1979), DSS was used to compare language samples elicited by mothers at home to those elicited by

speech pathologists in the clinic. Results showed no difference between DSS scores obtained in these two settings.

In a study by Haynes and McCallion (1981), DSS was used to determine if significant differences existed among the expressive language of children with differing cognitive tempo's (reflectivity versus impulsivity). In this study, a reflective subject was one who responded slowly and made fewer errors while an impulsive subject was one who responded quickly and made a greater number of errors. Results showed that the two groups were similar in their spontaneous expressive language performance as measured by DSS.

In a study designed to examine the effect of various language elicitation techniques on collecting language samples from normal and language-disordered children, DSS was used to identify the two groups of children (Wren, 1985). Language-disordered children, as defined by this study, scored below the 20th percentile while the normal children scored between the 40th and 60th percentiles. Results of this study indicated that both quality and quantity are affected by the amount of structure in the various elicitation tasks. The authors concluded that a combination of techniques yielded a large representative sample while one single task could not be said to elicit a representative sample of the children's language.

DSS has also been used as a diagnostic measure with which to compare outcomes of other tests. Carrow (1974), when developing the Carrow Elicited Language Inventory (CELI), used DSS as a means of establishing validity of the CELI.

Werner and Kresheck (1981) compared scores from DSS, CELI, and the Structured Photographic Language Test (SPLT) (Werner and Kresheck,

1974) to determine if all tests yield similar scores and information. The subjects consisted of fifty-four 4.5- and 6-year-old normal-language children. The results showed that some children scored below normal on one measure and a small number of children scored below normal on two of the measures. The authors postulated the following reasons as to why these results were obtained: (1) The children in this study had expressive language behaviors differing from their peers. (2) The tests did not provide an accurate representation of the child's capabilities. (3) Factors within the test may have affected the child's performance. The authors concluded that no one measure is best in eliciting a language sample from all children.

DSS was also used as a diagnostic measure to compare both the Fluharty Preschool Speech and Language Screening Test (FPSLST) (Fluharty, 1978) and the Bankson Language Screening Test (BLST) (Bankson, 1977) in a study conducted by Blaxley, Clinker, and Warr-Leeper (1983). Results indicated that the BLST is generally accurate in identifying language-impaired children (as determined by DSS) while the FPSLST failed to identify a large proportion of these children. The authors noted that any conclusions drawn from this study may be limited as the only diagnostic measure used was DSS; they concluded, however, that the FPSLST may not be accurate in identifying language-impaired children.

Blaxley et al. (1983) stated that DSS has been "widely used as a standard clinical and research tool" both in the identification of language-disordered children and as a diagnostic measure to establish validity of other tests. Therefore, it appears that a need exists for DSS to have adequate reliability. Spector (1981) refers to reliability

as the "consistency of a measuring device." Adequate reliability increases the probability that results from one assessment will be comparable to results obtained from a second assessment. In order to achieve adequate reliability, standardization of procedures is needed. Davis (1983) stated that standardization of administration "maximizes the likelihood that the test is given in the same way each time so . . . performance is assessed according to the same criteria each time." In order to achieve adequate standardization, the instructions for test administration must be explicit and comprehensive, and the examiner must follow the instructions exactly.

Effect of Different Stimulus Materials on the
Elicitation of Oral Language Samples

As DSS was developed in an attempt to provide a standardized method of describing the developmental level of a child's syntax, it outlined specific procedures for recording, transcribing, selecting, segmenting, and scoring the corpus. These procedures were first reported by Lee and Canter in 1971 and again by Lee in 1974(a).

DSS guidelines are not as explicit when describing the stimulus materials or methods to be used when collecting the language sample, however. Lee (1974a) stated that "uniformity of interest level is more important than uniformity of stimulus procedures in this kind of clinical setting." In the past, however, studies have investigated the effect which different stimulus procedures have on the language elicited and found that stimulus procedures do have an effect upon the language produced.

In some of the investigations, the use of pictures to elicit

language has been the focus of the study. Cowan et al. (1967) investigated the effect of stimuli variation, 10 different activity pictures, upon the language elicited from four age groups of children (5, 7, 9, and 11 years). Results indicated that different pictures elicit sentences of different length as measured by MLR. The language task, however, was not standardized as different examiners were used. Each examiner was free to use instructions such as "Tell me what you see in the pictures" or "Tell me what is happening" or "Tell me what the people are doing." Additionally, encouragement was not systematically applied. It is difficult, therefore, to attribute the difference in MLR to the particular picture used or to the instruction used.

Ahmed (1973) investigated the effects single-object pictures and multi-object pictures have on language produced by educable and trainable mentally retarded children. Scores for LCI, total number of words, and TTR were significantly greater for multi-object pictures than single-object pictures.

Another variation of the picture task was examined in a study by Strandberg and Griffith (1969). In this study, children were given cameras and allowed to take pictures of 10 toys presented during the experiment and then allowed to take 10 pictures at their homes. When asked to tell about each picture, the results showed that the children's responses were longer and more complex when verbalizing about the pictures which were taken at their homes.

From these studies, it appears that pictures are not uniform in the language which they elicit; different pictures elicit differing lengths and complexity levels of language. Therefore, when using

pictures to elicit language, it appears important to describe the type of pictures which are being used i.e., single versus multi-object pictures, activity pictures, pictures with high personal relevance, etc.

The use of a picture task has also been compared with other methods of obtaining a language sample. A study by Longhurst and Grubb (1974) examined the effect of object elicitation and picture elicitation upon the language produced by mentally retarded subjects of three different age groups (10, 11, and 14 years). The results of the study showed that pictures elicit a larger score for total number of words, TTR, and MLU than do objects. LCI did not show a significant difference between these two conditions. With the population used in this study, it appears that a picture task elicited language better than an object task.

A conflicting result regarding the effectiveness of picture tasks in eliciting language was found by Wren (1985). In this study, the effect of the type of elicitation task used to collect language samples from two groups of 6-year-old children was examined. One group was language-disordered (as identified by DSS) and the other group was considered normal in language development (as identified by DSS). The type of tasks used were a) spontaneous interaction tasks which consisted of free play with puppets and props, b) elicited interaction tasks which included storytelling, explanation of a game, creation of a story from three pictures, and description of pictures, c) specific tasks which involved the use of a birthday party task to elicit specific structures and sentence types, and d) a sentence building task (when given a word, the child was asked to make up a sentence).

Results of the study indicated that both quantity and complexity of language produced by these children varied from task to task. The birthday party task yielded the most language while response to pictures yielded the least language. Play with puppets, picture description, and sentence building tasks yielded the least syntactically complex language while the description of a game and the birthday party tasks yielded the most complex language. The authors concluded that, for these 6-year-old children, no single task elicited a representative sample of their language but a large, representative sample could be obtained through a combination of tasks. It appears that the amount of structure in the various tasks affected the quality and quantity of the language produced.

Storytelling tasks have also been researched by investigators. Atkins and Cartwright (1982) investigated the effectiveness of three language elicitation procedures on the language produced by Head Start Children (ages 3, 4, and 5 years). The three procedures used were picture interpretation, storytelling, and response to imperative requests ("Tell me what you would do if you got lost in a big store."). The authors concluded from the results of this study that, for these preschool children, picture interpretation yielded the most desirable results, followed by the imperative task and the storytelling task in that order.

Stalnaker and Creaghead (1982) gathered language samples from twelve Head Start Children (ages 4.0 to 5.6 years) under the following conditions: retelling a story using toys, playing with toys, and answering questions while playing with toys. These samples were then examined and compared. Results showed that toys with questions

produced the larger number of total utterances while playing with toys produced the smaller number. The only significant difference found among the three conditions was MLU; retelling the story with toys produced the larger MLU and playing with toys while the investigator asked questions produced the smaller MLU. The researchers concluded that questioning children does not inhibit language and that asking them to retell a story may be a useful approach to use when gathering a language sample.

Register (1975) examined stimulus familiarity when using story retelling as a means of eliciting language. Three variations of the story retelling task included the following: the child was asked to tell a story she/he recently had heard with the use of pictures, child was asked to tell an unfamiliar story with the use of pictures, and child was asked to tell a familiar story. Results of the study indicated that the more complex language and verb usage were elicited through the retelling of the familiar story.

Results of these studies appear to show differences in the effectiveness of storytelling tasks, as compared to other methods, to elicit oral language. Factors which should be considered when using storytelling tasks include whether or not the use of toys or pictures are used as aids in telling the story and the familiarity of the story itself.

The results of the many studies appear to be non-conclusive as to the best method of eliciting language; however, it appears that different collection procedures may significantly affect the language which is elicited.

The hypothesis that the use of different elicitation procedures

might significantly affect the results when analyzed by DSS has also been investigated. Longhurst and File (1977) compared DSS scores from language samples of Head Start Children (ages 3.11 years through 5.0 years) obtained under the following conditions: single-object pictures toys, multi-object pictures, and adult-child conversation. Results of this study show significant differences among DSS scores obtained under the various conditions, with conversation yielding the highest rank of DSS score and percentile means, followed by, in descending order, toys, multi-object pictures, and single-object pictures. Haynes, Purcell and Haynes (1979) compared the following conditions and their effect upon the language elicited from children of two age groups (4 and 6 years): a conversation task, an unscreened picture task where both experimenter and child looked at the stimulus at the same time, and a screened picture task where the child was able to view the stimulus but the experimenter was not. For these two groups of children, both picture tasks produced a significantly greater MLU than the conversation task, with the screened picture task eliciting a significantly higher MLU than the unscreened picture task. When the samples were analyzed using DSS, however, the conversational condition showed significantly more complex language than either of the picture tasks. The authors stated that the result "strengthens the notion that a conversation technique is perhaps the most effective method of obtaining a language corpus." It appears that DSS scores of the children in these studies might have been affected by the type of stimulus materials used.

In the report of statistical information on Developmental Sentence Analysis (Koenigsnecht, 1974), a study was reported which

investigated the effect of stimulus variables on DSS scores. The stimuli used were picture and story materials. No significant differences were found between the two materials on overall DSS scores, however, the use of various stimulus materials did show an effect on usage with specific grammatical categories, such as indefinite pronouns, personal pronouns, secondary verbs, and interrogative-reversals. Subjects tended to use more and higher level personal pronouns in describing pictures and more and higher level indefinite pronouns or noun modifiers in storytelling. The subjects also tended to receive higher scores on secondary verbs and interrogative reversals when describing pictures.

In the study by Koenigsknecht (1974) which provided the normative data for DSS, the procedure of elicitation of language was as follows. The stimulus materials consisted of miniature toys (a doll family and plastic furniture, a transport truck with small cars inside, and a plastic barn with farm animals), pictures from a preprimer series (Robinson, Monroe and Artley, 1962), "The Three Bears" story, and pictures from "What's Its Name?" (Utley, 1950). These stimulus materials were presented in the above order and the last 50 utterances of the session were selected for scoring.

While the DSS normative study specified the above three tasks for eliciting oral language, it did not specify from which task the utterances for analysis should be selected. Barrie-Blackley et al. (1978) stated that DSS normative analysis, therefore, "may be drawn from only one of the task segments . . . or a combination of tasks." It would appear that a need exists to examine the effect which these three collection procedures (playing with toys, picture task, and

storytelling) have on DSS scores obtained from children. If the three procedures do make a significant difference upon DSS scores, this data may be used to evaluate the reliability of DSS normative data, as well as provide more information on assessing the expressive language of children.

CHAPTER III

METHODS

General Plan

Eighteen normally developing children, composed of six groups of three, were tested to determine the effect of different language elicitation procedures upon the language samples obtained. Each child produced a language sample under three different conditions and each sample was analyzed using Developmental Sentence Scoring (DSS) (Lee, and Canter, 1971). Each child served as her or his own control.

Three different stimulus materials were used to elicit language samples: (1) toys, (2) pictures, (3) stories. The subjects were randomly assigned to one of six groups to which the stimulus materials were presented in differing order to counterbalance an order effect.

This investigator, a speech-language pathology graduate student, collected and analyzed the language samples using procedures as outlined by Lee (1974a).

Three DSS scores from each child (one for each stimulus material) were compared to determine whether a significant difference exists among the three scores.

Subjects

The subjects in this study consisted of children chosen from the Portland, Oregon metropolitan area. Each subject met the following

criteria: (1) Chronological age between 3.6 years and 5.6 years.

(1) Normal hearing sensitivity in one ear as determined by a pure-tone audiometric screening at a level of 25 dB for the following frequencies (Hz): 500, 1000, 2000, and 4000. (3) Normal receptive vocabulary within two standard deviations from the mean score for age level as measured by the Peabody Picture Vocabulary Test - Revised (PPVT-R) by Dunn and Dunn (1981). (4) No demonstrated or suspected physical or social delays as reported by teacher or observed by this investigator.

Sampling Method

A single page in the form of a letter to parents was sent explaining the study and seeking parental consent (Appendix A).

Screening

Upon receiving parental consent for the child to participate in the study, the investigator conducted a hearing screening and administered the PPVT-R, Form M, (Dunn and Dunn, 1981) to the prospective subject. The screening procedure took place in a quiet room at the child's preschool. Information regarding physical and social development was obtained through teacher report and/or investigator observation. Children who met the previously established criteria were included in the study.

Instrumentation

A portable Beltone audiometer was used to conduct the audiometric screen of the subject's hearing. A Panasonic portable cassette tape-recorder, model number RQ-309DS, was used to record the language samples.

The Peabody Picture Vocabulary Test - Revised (PPVT-R), Form M (Dunn and Dunn, 1981), an instrument which provides an estimate of a child's receptive vocabulary, was used to establish normal receptive vocabulary age, consistent with chronological age.

Developmental Sentence Scoring (DSS) developed by Lee and Canter (1971) is one method of analyzing a child's spontaneous tape-recorded speech sample. DSS was used to make a scored evaluation of the child's use of standard English grammatical rules.

Examiner Reliability

The investigator collected and analyzed all of the language samples according to the procedures described by Lee (1974a). The investigator's training consisted of successful completion of SP 410B "Language Sampling," a course offered at Portland State University, Fall term, 1984. This course was taught by a professor who holds the Certificate of Clinical Competence in Speech-Language Pathology. In addition, the investigator viewed the film, "Developmental Sentence Scoring" (Lee, 1974b).

Interjudge reliability was determined between the investigator and a Speech-Language Pathologist who holds a Certificate of Clinical Competence from the American Speech-Language-Hearing Association after the collection and transcription of the language samples. After the data had been gathered, one sentence from each language sample was randomly selected by a third party for the two judges to score independently. The judges' scores for each sentence were compared and reliability was calculated to be 86 percent. When the two judges scored an utterance differently, the judges made a decision about the

way the utterance should be scored. The remainder of the analysis was based upon these decisions.

Intrajudge reliability was determined by the investigator re-scoring the above utterances one week after interjudge reliability had been calculated. After scoring these utterances independently, the investigator compared these scores to the scores determined previously and reliability was calculated to be 96 percent.

Experimental Procedures

Language Sample Collection

Each subject had her or his language sampled three separate times using a different stimulus material each time. Each sampling session lasted approximately twenty minutes and approximately one-half hour separated the sampling sessions. Each language sampling session involved the investigator and the child, with the investigator concluding each session after the child had produced approximately 60 different utterances. The investigator used a hand-held counter to track the number of utterances produced during the interview. Each language sample was tape-recorded, with the recording device set up prior to the child entering the room. A large piece of felt material was situated under the microphone to minimize extraneous noise.

As instructed by Lee (1974a), the investigator sought to elicit complete sentences and high-level grammatical forms. Lee advised that it may be necessary to ask simple naming or fill-in questions to get the conversation started, but that these should be discontinued as soon as possible and the investigator did so. The investigator introduced past tense, modal verbs, plural pronouns, and such forms into the

conversation to present the child with the opportunity to use such forms himself.

One language sample was elicited utilizing the following toys as the stimulus material: small barn with farm animals, a doll family and plastic furniture, and a transport truck with small cars inside. The investigator engaged in creative play with the child to obtain a spontaneous speech sample.

Another language sample utilized pictures as the stimulus material to elicit language. The pictures consisted of multi-object and activity pictures from Game Oriented Activities for Learning (GOAL) (Karnes, 1972). The investigator showed approximately thirty pictures to the child and instructed him or her to "Tell me what is happening in these pictures."

The final method used storytelling as the stimulus to elicit language. The investigator presented pictures from stories and asked the child to tell the story. The stories which were used included "Goldilocks and the Three Bears" by Kincaid (1981), "The Three Little Pigs" by Banta and Dempster (1972), "The Three Billy Goats Gruff" by O'Grady and Throop, and "Little Red Riding Hood" published by the Western Publishing Co., Inc. A decision was made to use more than one story when, in a pilot study, this investigator was unable to elicit 50 utterances from one story alone. As many stories as needed to elicit 60 utterances were presented to each subject.

Each child was randomly assigned to one of six groups. In order to mediate an order effect, language samples were elicited from the groups by the procedures in the following order:

Group I toys, pictures, stories

| | |
|-----------|-------------------------|
| Group II | pictures, toys, stories |
| Group III | stories, toys, pictures |
| Group IV | toys, stories, pictures |
| Group V | pictures, stories, toys |
| Group VI | stories, pictures, toys |

Language Sample Transcription

Following the collection of the language samples, the investigator transcribed the recordings into transcripts following the specific instructions for transcription of the sample, selection of the corpus, and segmentation of the utterances provided by Lee (1974a).

The language sample was transcribed by the investigator. The investigator played back the tape as often as necessary to obtain an accurate account. Unintelligible sentences were excluded from the sample, as well as echolalic sentences.

For analysis by DSS, the corpus contained 50 different, intelligible, nonecholalic, consecutive complete utterances. A sentence was defined as complete if it contained a noun and verb in subject-predicate relationship. Following DSS guidelines (Lee, 1974a), this included the following: some two word combinations ("Doggie bark, Baby crying"), some two word wh-questions ("What happen?, What fall?"), imperative sentences which consisted of single verbs with implied subject ("Look, wait"), negative imperatives ("Don't cry, No look!"), and single obligatory-do plus negative ("Don't"). A sentence did not necessarily have to be grammatically correct to be included in the corpus; the only requisite relationship was the basic subject-verb requirement.

As only complete sentences are used for DSS, all fragmentary, incomplete sentences were discarded. Interjections and nouns in direct

address were not scored or included in the transcription. Question markers were included as questions receive a DSS score. Imperative interjections, such as "look, lookit, see" and the sentence tags "you know, I think, I guess" were separated out and given sentence status. Lee (1974a) listed the following five rules for separating utterances:

- (1) Sentences which begin with an initial conjunction will be scored as complete but the initial conjunction will not receive a score.
- (2) When the conjunction "and" is used to join two independent clauses, only one "and" conjunction will be scored per sentence.
- (3) When "and" conjunction is used in a series, a compound subject or a compound predicate, the sentence will not be broken up.
- (4) Internal conjunctions other than "and" do not require a sentence to be broken up.
- (5) If a child has indiscriminately overused a conjunction, the clinician may choose to break up a lengthy sentence. If a child's language sample contains utterances which combine both a presentence structure and a complete sentence, the clinician will separate the sentence if it is an independent clause, deleting both the fragment and the conjunction. Only the independent clause will be scored. For example, the sentence "A rabbit and it hopped away," only "it hopped away" will be scored. The investigator of this study closely followed these guidelines when separating utterances.

Scoring

The language samples were scored as described by Lee (1974a) (Appendix B). This scoring gave weighted scores to an acquisitional order of pronouns, verbs, negatives, conjunctions, yes-no questions, and wh-questions. Additionally, a sentence point was given when the

utterance was syntactically, morphologically, and semantically correct according to adult standard English. If a structure was attempted, yet incomplete, an "attempt mark," a horizontal line, was inserted in place of the numerical score. A sentence point was not awarded if the utterance received any attempt marks. After scoring the 50 sentences individually, the scores were summed and divided by 50 to obtain each DSS score.

Analysis of Data

DSS scores were computed for each child's spontaneous language samples. Descriptive statistics were then applied to determine mean DSS scores for the language samples which were elicited using toys, pictures, and stories. To determine if a significant difference exists between the means of the language samples which were elicited in three different ways, a two-tailed t -test for related means was computed.

CHAPTER IV

RESULTS AND DISCUSSION

Results

Individual language samples were elicited from eighteen children between the ages of 3.6 years and 5.6 years. Each child produced three separate language samples using a different type of stimulus material each time. The language samples were then submitted to analysis using Developmental Sentence Scoring (Lee and Canter, 1971). The means of the DSS scores obtained under each of the experimental conditions were then compared to determine whether a significant difference exists among the scores.

A two-tailed t-test for related means was computed for the DSS mean scores. Table I shows the comparison between the mean DSS scores as well as the DSS means and standard deviations.

Examination of the data presented in Table I reveals that the means ranged from 6.87 to 8.90. DSS scores obtained using stories received the highest mean score of 8.90, followed by toys with a mean score of 8.01, and pictures with a mean score of 6.87.

Further examination of the data reveals that a statistically significant difference beyond the .01 level of confidence occurred between the DSS mean obtained under the storytelling condition and the DSS mean obtained under the picture condition. The storytelling

TABLE I

A COMPARISON OF THE DSS MEANS OBTAINED USING
THREE DIFFERENT STIMULUS MATERIALS

| Stimulus Material | Mean | S.D. | df | t-test |
|-------------------|------|------|----|--------|
| Toys | 8.01 | 1.40 | 17 | 2.10 |
| Pictures | 6.87 | 2.05 | | |
| Storytelling | 8.90 | 2.23 | 17 | 3.42* |
| Pictures | 6.87 | 2.05 | | |
| Storytelling | 8.90 | 2.23 | 17 | 1.64 |
| Toys | 8.01 | 1.40 | | |

* = $p < .01$

condition produced the higher DSS mean score with a mean of 8.90 while that obtained under the picture condition was 6.87. While no other comparisons of DSS means were statistically significant, there was a trend towards significance ($p < .10$) between the DSS mean of scores obtained using toys and that obtained using pictures. In that comparison, toys obtained the higher mean DSS score of 8.01 while the mean for pictures was 6.87.

Discussion

The analysis of results showed that storytelling consistently elicited the higher mean DSS score followed by, in descending order,

toys and pictures. Additionally, a significant difference was found between DSS scores of language samples elicited by storytelling and those elicited by pictures. The following discussion will examine a) the results of past studies regarding stimulus material effectiveness in comparison to the findings of the present study; b) possible variables affecting the outcome of the present study; c) implications regarding the effects of stimulus materials on DSS outcomes; d) implications regarding past studies using DSS; e) use of DSS as a clinical and research tool.

Stimulus Material Effects on Elicited Language

Lee (1974a) presented the hypothesis that storytelling might elicit higher scoring sentences than either the pictures or the toys. This hypothesis was confirmed by the outcome of this study. This finding would appear to be consistent with that presented by Stalnaker and Creaghead (1982) who found a storytelling task elicited a longer MLU when compared to a picture task. Subjects in that study ranged between 4.0 to 5.6 years. A conflicting result, however, was reported by Atkins and Cartwright (1982). In that study a picture task elicited the greater number of utterances with a longer MLU, followed by, in descending order, an imperative task (e.g., "Tell me what you would do if you got lost in a big store.") and a storytelling task. The subjects' ages ranged between 3 and 5 years.

One factor which might account for the difference in results found among these studies is the type of material used during storytelling. Barrie-Blackley et al. (1978) noted that the type of language task used to elicit a language sample may affect the

respondent's language. The language task may be described as the instructions, the stimulus and the situation. In the Stalnaker and Craghead study (1982), the storytelling task was conducted as follows: the child listened to the investigator tell a story which "appealed to Head Start Preschool children" while the investigator manipulated a set of toys to act out the story. After hearing the story, the child was instructed to tell the story, using the toys if desired. In the Atkins and Cartwright study (1982), the child heard a story while looking at pictures and then was instructed to retell the story without pictures. In the current study the child did not hear the story prior to telling it using pictures as a visual aid. The results of these three studies regarding the efficiency of storytelling as a method to elicit language may not be comparable due to the fact that different tasks and stories were presented. Storytelling as a task to elicit language is not homogeneous across research.

Another outcome of this study was that pictures consistently elicited language samples which received the lower mean score. Wren (1985) found that responses to pictures yielded the least language when compared to a variety of other elicitation methods. As presented in the review of the literature, however, other studies have found that picture tasks elicited a longer MLU than object or conversation tasks (Haynes et al., 1979; Longhurst and Grubb, 1974).

One possible reason for the difference found among these results is the research finding that different pictures elicit different types of language (Ahmed, 1973; Cowan et al., 1967; Strandberg and Griffith, 1969). In the study by Haynes et al. (1979), the pictures consisted of magazine pictures depicting activities generally "familiar" to

children. In the Longhurst and Grubb (1974) study, the pictures used were slides from the Peabody Language Development Kit (Level 2) (Dunn and Smith, 1967) and consisted of color story situation slides which used a variety of subjects and settings. In the current study, the pictures used were color cartoon type multi-object or activity pictures from the GOAL language program (Karnes, 1972). The pictures used in these three studies were considerably different. This may account for the difference in effectiveness of these picture tasks in eliciting language.

Another possible reason for the difference found among these results is the instructions used when presenting the picture tasks. In the study by Haynes et al. (1979), the children were instructed to tell the examiner as much as they could about each picture. In the Longhurst and Grubb (1974) study, the child was instructed to "Tell me all you can about this." In the current study, the instruction presented was "Tell me what is happening in this picture." It appears that the instructions used were quite different with the first two studies using a more open-ended instruction than the current study. It may be that the instruction presented in the current study was more limiting than the instructions presented in the other two studies. This difference in instructions may have affected the type of language elicited from these picture tasks.

One difficulty which exists when comparing the results of the current study to those found in the review of the literature is the fact that different types of analysis were used to compare the language samples. In many of the studies, MLU was used to analyze the samples (Atkins and Cartwright, 1982; Haynes et al., 1979; Longhurst and Grubb,

1974; Stalnaker and Creaghead, 1982). The correlation between DSS and MLU has not been clearly established. Koenigsnecht (1974) reported a study which examined the correlation of MLR (which measures total number of words as opposed to MLU which measures number of morphemes; both are measures of verbal output) and DSS and stated that results supported the view that

DSS measures the impact of not only the developmental level or maturity of syntactical usage but also the typical length of a child's utterances.

A study by Haynes et al. (1979), however, analyzed samples using both MLU and DSS. When analyzed using MLU, a picture task elicited language better than a conversation task. When analyzed using DSS, the conversation task elicited language better than the picture task. Therefore, it remains unclear what the correlation is between DSS and MLU. Results from Haynes et al. suggest that DSS and MLU measure different aspects of language. This makes it difficult to make an adequate comparison between the results of these studies with respect to the effectiveness of different stimulus materials.

Possible Variables Affecting Outcome of Study

Some possible reasons for the finding that storytelling in this study elicited a significantly higher DSS mean score than pictures include the type of language elicited using storytelling, the stimulus materials which were used, and the child's background regarding the stimulus materials.

Type of Language Elicited

Tilden-Browning (1985) suggested that when using storytelling,

children might be reproducing rote responses which they had heard (learned) at an earlier time. She hypothesized that retelling a familiar story might allow the children to use higher level grammatical forms than those used in spontaneous speech. This may account for the higher DSS scores obtained when storytelling was used to elicit the language sample. The greater standard deviation found in the storytelling task may also be a reflection of variance in rote learning of these stories. When examining the transcriptions of the children's samples during storytelling, the utterances do appear to be rote and stereotypical. For example, when telling Kincaid's (1981) version of "Goldilocks and the Three Bears," utterances were often repeated with one word change, such as "Papa bear said, 'Someone has been eating my porridge.'" followed by "Mama bear said, 'Someone has been eating my porridge.'" As these were not exact replications, they were included in the corpus selected for analysis. Due to the nature of storytelling, these utterances were appropriate, but it is uncertain whether repetitive utterances such as these would be produced during spontaneous speech. While storytelling might produce the higher mean DSS score as found with this study, it also may be that storytelling does not accurately represent a child's spontaneous expressive language.

The pictures which were used in this study may account for the reason that samples elicited by pictures received the lower mean DSS score. It may be that these pictures were not as interesting or motivating to the children as the toys or stories. When presenting the pictures, it was noted that the children tended to focus on one aspect of the picture and then dismiss it (perhaps due to the instructions

used). Attempts for elaboration were generally unsuccessful during the picture tasks while examiner attempts for elaboration during the toys or storytelling tasks were generally more successful.

While studies have shown that different pictures affect the elicitation of language (Ahmed, 1973; Cowan et al., 1967; Strandberg and Griffith, 1969), there does not appear to be any studies reporting how different stories may affect the language sample. Therefore, it may be possible that the different stories used in this investigation may have affected the language elicited. This study used the stories of "Goldilocks and the Three Bears" by Kincaid (1981), "The Three Pigs" by Banta and Dempster (1972), "The Three Billy Goats Gruff" by O'Grady and Throop, and "Little Red Riding Hood" published by Western Publishing Co., Inc. to elicit the language samples. While the story of "Goldilocks and the Three Bears" was consistently presented first, the following stories were not presented in a specific order. It is possible that different stories elicit differing levels of grammatical complexity and thus the high storytelling DSS mean may be a reflection of the grammatical complexity inherent in the individual stories used in this study. To determine if this might be the case, this investigator analyzed the first 50 utterances of both Kincaid's (1981) version of "Goldilocks and the Three Bears" and "Little Red Riding Hood" published by Western Publishing Co., Inc. DSS scores for these stories were 10.02 and 10.84, respectively. This information appears to indicate that the utterances used in these two stories are comparable when analyzed according to DSS. It would appear that the differing presentation of the stories probably did not significantly affect the outcome of this study.

Another factor which was not controlled in this study was the child's background regarding previous experiences with toys, pictures, and stories. Barrie-Blackley et al. (1978) stated that "variables relating to past experiences are important at all ages." She noted that a child who has heard bedtime stories every night or had opportunities to play with commercial toys might perform much better when presented with these items during a language sampling session than children who did not have these past experiences. Register (1975) examined familiarity of stories in her study and results suggested that more complex language was elicited during a familiar story task when compared to an unfamiliar story task or the retelling of a recently heard story. All children in the present study were able to relate the stories presented during the storytelling task, but it is unclear how much previous experience they had with these stories. Some might have heard the stories extensively while others may have heard them only once or twice. Also, some children in this study, as indicated to the investigator during the sampling session, owned the same toy barn and animal set that was used during the toy task. This suggests that some of these children had a great deal of previous experience with these exact toys while others may have had none. In addition, it is uncertain how much previous experience the children in this study had with viewing and describing pictures. Therefore, the language sample obtained may have been influenced by the subject's past experience with the stimuli used in this study. If a child had a high degree of familiarity with these materials, it might be that the sample obtained was of higher complexity language, not necessarily reflecting advanced expressive language but, instead, familiarity with the materials

involved.

Elicitation, Transcription, Analysis and Order Variables

Other less probable factors which may account for the current results were taken under consideration in the design of the present study. These factors include elicitation, transcription and analysis variables, as well as order effect.

As presented in the review of the literature, Barrie-Blackley et al. (1978) reported that subject population, examiner, situation or site, stimulus, instructions, consequences and language task are among the elicitation variables which may affect the respondent's language. Therefore, the examiner in this study adhered closely to Lee's (1974a) guidelines as to the elicitation of the oral language samples. As instructed, all samples were obtained in a quiet room at the child's preschool and the examiner attempted to elicit complete sentences and high-level grammatical forms from the subjects. The same stimulus materials were used with each child, although presented in differing orders. As each child served as his or her own control and as the investigator solely conducted all of the language sampling sessions, it does not seem probable that differences found in this study are attributable to differences in the subject population, site, stimulus, instructions or consequences used.

The investigator also transcribed and analyzed the samples solely, thus reducing the possibility that the difference found among the samples was due to the manner in which the samples were transcribed or analyzed.

Temporal reliability of four repeated applications in a two-week

period of the DSS language assessment techniques was reported by Koenigsknecht (1974). Results suggested that in the two-week period, DSS scores for 10 subjects increased progressively from the first sampling to the fourth sampling. These results were attributed to biases such as the children's becoming overly familiar with the stimulus materials used or commenting on the same things during repeated sessions. In the current study, examination of the raw data according to session (Table II) showed that the mean of DSS scores obtained in the first sampling session was 7.40, that obtained in the second session was 7.83 and the final session's mean score was 8.57. To counterbalance this order effect (the last DSS score higher than the previous two sessions), the presentation of materials was ordered so that one particular stimulus material was not consistently used as the stimulus for any particular session. The order of presentation of stimulus materials is shown in Appendix C. Therefore, while an order effect did occur, the randomized order of presentation makes it doubtful that results of this study are due to the order effect observed.

Effects of Stimulus Materials on DSS Scores

After examining the variables which might have affected the outcome of this study, the results suggest that DSS scores may be significantly affected by the stimulus materials used to elicit the language samples. Of three previous studies which examined variability of DSS scores according to stimulus materials (Haynes et al., 1979; Koenigsknecht, 1974; Longhurst and File, 1977) two appear to support these results while one does not. The study by Longhurst and File found that conversation elicited the higher DSS score followed by, in

TABLE II

DSS SCORES, PERCENTILES, AND MEANS
BY SESSION

| Subject | 1st Session | | 2nd Session | | 3rd Session | |
|----------|-------------|-------|-------------|-------|-------------|-----|
| | Score | % | Score | % | Score | % |
| 1 | 7.24 | 20 | 5.90 P | <10* | 7.24 | 20 |
| 2 | 7.22 | 15 | 5.66 P | <10** | 12.58 | >90 |
| 3 | 8.50 | 65 | 9.92 | 88 | 11.58 | >90 |
| 4 | 6.66 | 18 | 8.78 | 70 | 7.98 | 50 |
| 5 | 4.54 P | <10** | 5.92 S | <10* | 7.70 | 30 |
| 6 | 6.00 P | <10* | 13.90 | >90 | 8.42 | 50 |
| 7 | 5.74 P | <10* | 7.76 | 66 | 8.54 | 85 |
| 8 | 7.88 | 70 | 7.64 | 60 | 7.24 | 50 |
| 9 | 8.22 | >90 | 7.58 | 82 | 7.54 | 82 |
| 10 | 4.94 T | <10* | 6.28 | 35 | 6.02 | 25 |
| 11 | 7.60 | 42 | 11.88 | >90 | 13.30 | >90 |
| 12 | 6.92 | 26 | 8.18 | 55 | 9.58 | 82 |
| 13 | 6.46 | 28 | 5.08 P | <10* | 7.62 | 63 |
| 14 | 10.04 | >90 | 7.82 | 70 | 10.42 | >90 |
| 15 | 6.64 | 25 | 6.46 | 20 | 9.06 | 88 |
| 16 | 9.16 | 88 | 8.88 | 80 | 5.82 P | <10 |
| 17 | 9.02 | 73 | 5.28 P | <10* | 7.58 | 37 |
| 18 | 10.40 | >90 | 7.96 | 90 | 6.10 | 28 |
| Mean DSS | 7.40 | | 7.83 | | 8.57 | |

* = Below the 10th percentile line

** = At least 1 DSS point below the 10th percentile line

T = Sample elicited by toys

P = Sample elicited by pictures

S = Sample elicited by stories

descending order, toys, multi-object pictures and single-object pictures. The study by Haynes et al. found that higher DSS scores were obtained by samples elicited during conversation (child was asked questions about a variety of topics such as family, sports, activities, etc.) when compared to those elicited during picture tasks. These results appear to be consistent with the current finding that DSS scores appear to vary depending on the stimulus materials used to elicit the language samples. The study reported by Koenigsknecht, however, found no difference in overall DSS scores between samples elicited by pictures and those by stories. Two versions of each type of stimulus materials were counter-balanced over trials. There is conflicting evidence as to whether different stimulus materials significantly affect DSS scores.

In each of these studies, the sample size was small. Longhurst and File's (1977) sample consisted of twenty Head Start students. Haynes et al. (1979) used a sample of twenty-four preschool students. Koenigsknecht's (1974) sample consisted of ten preschool children and the current study's sample consisted of eighteen preschool children. It may be that further studies on a larger sample population, using the same type of data analysis, will need to be conducted in order to reach a conclusive finding regarding the effect of different stimulus materials on language samples. Across research in this area of investigation, however, a picture appears to be emerging that children respond in a heterogeneous fashion that does not favor one method over another. Perhaps each child should be approached as a unique respondent. This would suggest multiple assessments with some children in order to achieve an accurate, representative sample of expressive

language. Should these findings be confirmed by future research, many implications exist concerning the use of DSS as a research and clinical tool.

Past Studies Using DSS

If DSS scores are affected by stimulus materials, results from studies which have used DSS to differentiate the language behavior of selected groups of children should be viewed cautiously. Careful examination of past studies reveals that a variety of stimulus materials was used to elicit the language samples. Therefore, results of these studies may have been influenced by the stimulus materials used to elicit the language samples.

In a study which examined DSS scores of samples obtained in the home and in the clinic settings, Kramer et al. (1979) used different stimulus materials to obtain the samples. To obtain the clinic samples, a conversation task was used presenting toys and books where necessary. To obtain the home samples, a free play format was used. The results revealed that the home samples yielded longer MLU and higher DSS scores for most of the subjects. These results concerning the site of the language sample, however, are confounded by the stimulus materials used and vice versa. It may be that the results were due to the different stimulus materials or to the differing environment where the samples were obtained. With the design in this study, it is difficult to separate these two variables and their effects.

In a study by Wren (1985), the stimulus materials used to obtain the language samples to be analyzed using DSS were not specified. These samples were used to differentiate language-disordered children

from normally developing children. These two subject groups were then used in the study to examine the effectiveness of different stimulus materials. Results of the study indicated that in some cases, language-disordered children responded differently to a particular task than normal children. The author concluded that the amount of structure in the various tasks presented affected the quality and quantity of language produced by these two subject groups. While the study examined different stimulus materials, it did not appear to consider how stimulus materials affected the DSS scores which were used to separate the subjects into the two experimental groups. It may be that the differences between the two groups were not a reflection of language development but a reflection of the stimulus materials used to elicit the language samples.

DSS has also been used as a diagnostic tool with which to compare other tests. When establishing the validity of the CELI (1974), Carrow did not specify the stimulus materials used to obtain the language samples. In the study conducted by Werner and Kresheck (1981) which compared scores from DSS, CELI, and SPLT, it was reported that an open set of stimulus materials "motivating" to the child was used to obtain a sample of free speech. No further specifications were given. Also, the study (Blaxley et al., 1983) which compared the FPSLST and the BLST did not specify the stimulus materials. It is uncertain how the validity of these studies may have been affected by the stimulus materials used to obtain the language samples. When designing a study which uses DSS as a validity measure, future researchers should increase their awareness of how stimulus materials affect DSS scores, and of greater importance, how stimulus materials affect expressive

language.

Use of DSS as a Clinical and Research Tool

Future researchers should also be aware that stimulus materials presented during the language sampling session appear to affect the variability of responses obtained. The variability among the eighteen subjects involved in the current study may be examined when looking at the data presented in Table I. Storytelling elicited DSS scores with the greater standard deviation, toys elicited DSS scores with the smaller standard deviation and pictures fell between the two. The smaller standard deviation found among the samples elicited using toys may be indicative that toys elicited more uniform responses i.e., the subjects tended to produce utterances in much the same way. The storytelling task tended to elicit responses with greater variation among the subjects as indicated by the larger standard deviation. Therefore, storytelling appeared to elicit responses which varied from one subject to the next (perhaps due to their past experience with the stories presented as discussed previously). The standard deviation of samples obtained using pictures showed more variation than toys but less than storytelling. When deciding the type of stimulus material to use, the clinician or researcher should consider the purpose for which the samples are to be obtained. If the intent is to elicit a differentiation of responses between the subjects (such as separating language-disordered from normally developing children), results suggest that storytelling might provide more adequate information than pictures or toys. In this study, poor scores on pictures almost never reflected the child's potential.

Clinically, DSS has been used as a diagnostic tool. In order to ascertain how the difference found among the DSS scores might affect the diagnostic use of DSS, the eighteen subjects' individual DSS scores were converted to the percentiles provided by Lee (1974a) (Appendix D). Table II shows the subject's age, the individual DSS scores, and their correlating percentiles.

Examination of the data presented in Table II reveals that the differences in DSS scores obtained in this study were reflected by considerable variation when these scores were converted to percentiles. Eight of the eighteen subjects received one score with a percentile of 10 or less while their remaining two scores received higher percentiles. One subject received two scores below the 10th percentile with the remaining score higher than the 10th percentile. Two subjects received one score below the 10th percentile and one score above the 90th percentile.

When interpreting these percentiles, Lee (1974a) reported that children scoring close to the 10th percentile require further evaluation and those scoring one whole DSS point below the 10th percentile should receive interventional teaching. Based on these recommendations, of the eighteen subjects in the study, eight would require additional diagnostic testing and two would warrant interventional teaching. (It may also be noted that of the ten samples receiving DSS scores below the 10th percentile, eight of them were elicited using pictures.)

Lee (1974a) stated that

Clinicians should be cautioned not to make arbitrary decisions about enrolling or continuing a child for remedial training on the basis of a single DSS which falls at or below the 10th percentile line.

She advised that other factors such as social adjustment, intellectual functioning, motor functioning, etc. should be taken into consideration when making the decision to enter or continue a child in remedial training. Results of this study support the recommendation that DSS should not be used as a sole diagnostic measure. These results show that significant differences in the percentiles can occur, perhaps due to the stimulus materials used, which may lead the clinician to under- or overestimate a child's language abilities and his or her need for remedial training. An overestimation could lead to a child who warrants intervention being passed over for services while an underestimation might allow children into remedial training who have no need for it. This could cause serious difficulties with the speech-language clinician's effectiveness and accountability when making caseload decisions. It appears that a sole DSS score probably does not accurately represent a child's expressive use of syntax.

In conclusion, results of this study suggest that DSS scores tend to vary depending on the stimulus materials used to elicit the language samples. This variation could affect the reliability of DSS as the scores may vary depending on how the language sample was obtained. Different stimulus materials used with the same subject may produce DSS scores which are significantly different. This variation of scores may lead a speech-language clinician to an inadequate or erroneous description of a child's expressive language. This variability could also affect interpretation of past studies which have used DSS as these studies may not have controlled for the effects of stimulus materials in the design of the study. When using oral language sampling, clinicians and researchers should take into consideration the possibility

that stimulus materials may significantly affect a child's expressive language. It may be, as suggested by Wren (1985), that a variety of stimulus materials will need to be used in order to obtain a representative sample of expressive language.

CHAPTER V

SUMMARY AND IMPLICATIONS

Summary

Oral language sampling has been used as a clinical and research tool to describe a child's expressive language skills. Through the years, many methods have been presented with which to score or analyze utterances produced by a child during spontaneous conversation.

One method of analyzing a child's use of syntax, Developmental Sentence Scoring (DSS) by Lee and Canter (1971) was the focus of this study. The study which provided the normative data only revealed that toys, pictures, and stories were used to elicit the language samples. A review of the literature of oral language sampling revealed that different stimulus materials produce differing effects on the expressive output of children.

Therefore, the present study sought to determine the effect different stimulus material has on the language elicited from children. The purpose was to determine whether a significant difference exists among language samples elicited three different ways when analyzed using DSS. Eighteen children between the ages of 3.6 and 5.6 years were chosen to participate in the study. All of the children had normal hearing, normal receptive vocabulary skills and no demonstrated or suspected physical or social delays. Three language samples, each elicited by either toys, pictures, or stories, were obtained from each

child. For each sample, a corpus of 50 utterances was selected for analysis and analyzed according to the DSS procedure as described by Lee (1974a).

The means and standard deviations of the DSS scores were calculated for the samples elicited by each stimulus material. Two-tailed t-tests for related means were computed to determine if statistically significant differences exist among the language samples which were elicited by toys, pictures, or stories.

Results of this study showed that storytelling consistently elicited the higher mean DSS score, followed by, in descending order, toys and pictures. Analysis of results indicated that using stories ($\bar{x} = 8.90$) to elicit language produced a significantly higher mean DSS score than when pictures ($\bar{x} = 6.87$) were used to elicit language samples ($t = 3.42, p < .01$). A difference trending toward significance was also obtained between samples elicited by toys ($\bar{x} = 8.01$) and those elicited by pictures ($\bar{x} = 6.87$). As each child served as his or her own control and the same examiner and procedures were used to obtain each sample, it is proposed that these differences were primarily related to the stimulus materials used to elicit the language samples.

Implications

Clinical Implications

The results of this study suggest that different stimulus materials may significantly affect expressive language as analyzed by DSS. As it is not known which stimulus materials provided the DSS normative data, it would appear that the speech-language clinician should use the normative data cautiously. If the speech-language

clinician uses stimulus materials which are different from those used in the normative study, interpretation of the clinician's DSS scores may be erroneous or misleading. The speech-language clinician would be wise to follow Lee's (1974a) recommendation that DSS should not be used as the sole diagnostic tool to evaluate the language abilities of a child.

Due to these constraints, it might be that DSS would better serve as an indicator of progress rather than as a diagnostic tool. When using DSS as a clinical tool to document changes in treatment, therefore, it would increase reliability of the samples obtained if they were elicited in a uniform manner using the same materials each time. It was noted, however, that DSS scores repeated four times in a two-week period showed a trial effect with DSS scores increasing from the first session to the fourth (Koenigsknecht, 1974). The speech-language clinician should be aware of this trial effect when using DSS as a repeated measure.

Speech-language clinicians also need to carefully examine studies which have used DSS to determine how the language samples were obtained. If the stimulus materials varied significantly among the subjects in the study or were significantly different from the stimulus materials used in the DSS normative study, the validity of the study may be impaired. The speech-language clinician should examine the design of the study carefully before accepting the conclusions of the study.

While storytelling consistently received higher DSS scores than the other two methods, the assumption that this is the best task to elicit language should not be made. It is unclear how accurately storytelling represents spontaneous expressive language capabilities.

It may be that storytelling allows children to use language of a higher grammatical complexity than that used in spontaneous speech.

Regional differences in DSS normative data may be related to the fact that storytelling produced language samples with the higher DSS scores. Replicated DSS normative studies in the Portland, Oregon area for four, five, and six year olds (McCluskey, 1984; McNutt, 1985; Tilden-Browning, 1985) revealed that a significant difference exists among DSS scores obtained in Portland, Oregon and those obtained in the Midwest. The DSS scores obtained in the Midwest were significantly higher than those obtained in Oregon. In the study by Tilden-Browning (1985), it was noted that the majority of the corpus used for analysis was selected from utterances obtained during play with toys. McCluskey (1984) reported that most of the utterances selected for analysis were obtained from playing with toys and storytelling. McNutt (1985) obtained the corpus from playing with toys. While the DSS normative study did not specify from which tasks the corpus was selected, it did state that the last 50 utterances of the session were selected for analysis. In this normative study, the stimulus materials were always presented in the order of toys, followed by pictures and then stories. It may be that the majority of the utterances selected for analysis in the DSS normative study were obtained from the storytelling portion of the session. While geographical differences may have accounted for the differences observed between the DSS scores obtained in Oregon and the Midwest, use of different stimulus materials might also be responsible for some of the differences observed. It is possible that samples based on storytelling might need to be used in order for the DSS normative data to be used reliably.

Research Implications

Replications of this study might be conducted on specific age groups so that the scores could then be compared to DSS normative data to determine which condition produced language samples most like those obtained in the normative data (Koenigsnecht, 1974). It also might be beneficial to obtain storytelling samples only and determine how these compare with the original DSS normative data.

This study was run using children with normal language development. It would be worthwhile to determine if the same results would be obtained if the subject population consisted of language-disordered children as these are the children most likely to be enrolled in speech intervention.

As this study's sample size was small, more conclusive findings might also be obtained by conducting the study with a larger population.

While this study showed that an overall DSS score difference exists between the language samples, the samples were not analyzed according to grammatical categories. It would be useful to compare these samples regarding the type of grammatical forms used. For example, more and higher complexity pronouns might be used during storytelling than when playing with toys.

It is also possible that children of different age groups might show a significant difference in the way they respond to the stimulus materials. Koenigsnecht (1974) reported that the two-year-old subjects responded to toys better than pictures or storytelling while older children responded better to pictures and stories. Examination

of the raw data according to age groups (Appendix E) reveals that for all age groups, storytelling received the highest mean DSS score followed by toys and then pictures. A replica of this study using children of different age groups would help to clarify this issue.

As storytelling elicited the higher level grammatical forms in this study, it would be helpful to know how representative storytelling is of spontaneous expressive language skills.

As results of this study indicate that DSS scores may be significantly affected by different stimulus materials, future researchers need to be aware of the fact that the way language samples are elicited may significantly affect DSS scores. If the language samples are not elicited in the same manner for each subject, it would be difficult to ascertain if the DSS scores accurately represented the language capabilities of the subjects or were a reflection of the stimulus materials used to obtain the language sample. This could lead to inaccurate conclusions being drawn from the results of the study.

It appears that there are many variables which may affect the elicitation of oral language samples. Only through continued research will the speech-language clinician learn methods to accurately and reliably elicit language samples from children.

SELECTED BIBLIOGRAPHY

- AHMED, S.E. (1973). Linguistic analysis of children's speech: Effects of stimulus media on elicited samples. Unpublished Master's thesis, Kansas State University. In S. Barrie-Blackley, D. Musselwhite, and S. Rogister (1978). Clinical oral language sampling. Danville IL: The Interstate Printers and Publishers, Inc.
- ATKINS, C.P. and CARTWRIGHT, L.R. (1982). An investigation of the effectiveness of three language elicitation procedures on head start children. Language, Speech, and Hearing Services in Schools, 13, 33-36.
- BANKSON, N.W. (1977). The Bankson language screening test. Baltimore: University Park Press.
- BANTA, M. and DEMPSTER, A. (1972). The three little pigs. Racine WI: Western Publishing Co., Inc.
- BARRIE-BLACKLEY, S., MUSSELWHITE, C., and ROGISTER, S. (1978). Clinical oral language sampling. Danville IL: The Interstate Printers and Publishers.
- BLAXLEY, L., CLINKER, M, and WARR-LEEPER, G. (1983). Two language screening tests compared with developmental sentence scoring. Language, Speech, and Hearing Services in Schools, 14, 38-46.
- BROWN, R. (1973). A first language: The early stages. Cambridge: Harvard University Press.
- CARROW, E. (1974). A test using elicited imitation in assessing grammatical structure in children. Journal of Speech and Hearing Disorders, 39, 437-444.
- COWAN, A.A., WEBER, J., HODDINOTT, B.A., and KLEIN, J. (1967). Mean length of spoken response as a function of stimulus, experimenter, and subject. Child Development, 38, 191-203.
- DAVIS, G.A. (1983). A survey of adult aphasia. Englewood Cliffs NJ: Prentice-Hall, Inc.
- DUNN, L.M. and DUNN, L.M. (1981). Peabody picture vocabulary test - revised. Circle Pines MN: American Guidance Service.
- DUNN, L.M. and SMITH, J.O. (Eds.) (1967). Peabody language development kit. Circle Pines MN: American Guidance Service.

- FLUHARTY, N.B. (1978). Fluharty preschool speech and language screening test. Boston: Teaching Resources Corp.
- HAYNES, W.O. and MCCALLION, M.B. (1981). Language comprehension testing: The influence of three modes of test administration and cognitive tempo on the performance of preschool children. Language, Speech, and Hearing Services in Schools, 12, 74-81.
- HAYNES, W.O., PURCELL, E., and HAYNES, M.D. (1979). A pragmatic aspect of language sampling. Language, Speech, and Hearing Services in Schools, 10, 104-110.
- JOHNSON, W. (1944). Studies in language behavior. Psychological Monographs, 56, 1-111.
- JOHNSON, W., DARLEY, F.L., and SPRIESTERSBACH, D.C. (1963). Diagnostic methods in speech pathology. New York: Harper and Row, Publishers, Inc.
- KARNES, M. (1972). Game oriented activities for learning. "Situation pictures." Springfield MA: Milton Bradley Co.
- KINCAID, L. (1981). Goldilocks and the three bears. Cambridge, England: Brimax Books.
- KOENIGSKNECHT, R. (1974). Statistical information on developmental sentence analysis. In L. Lee (1974), Developmental sentence analysis. Evanston IL: Northwestern University Press.
- KRAMER, C., JAMES S., and SAXMAN, J. (1979). A comparison of language samples elicited at home and in the clinic. Journal of Speech and Hearing Disorders, 44, 321-330.
- LEE, L. (1974a). Developmental sentence analysis. Evanston IL: Northwestern University Press.
- LEE, L. (1974b). "Developmental sentence scoring" (Film). 20 min., color. Northwestern University Film Library.
- LEE, L. and CANTER, S.M. (1971). Developmental sentence scoring: A clinical procedure for estimating syntactic development in children's spontaneous speech. Journal of Speech and Hearing Disorders, 36, 315-340.
- Little Red Riding Hood. Racine WI: Western Publishing Co., Inc.
- LONGHURST, T.M. and FILE, J. (1977). A comparison of developmental sentence scores from head start children collected in four conditions. Language, Speech, and Hearing Services in Schools, 8, 54-64.

- LONGHURST, T.M. and GRUBB, S. (1974). A comparison of language samples collected in four situations. Language, Speech, and Hearing Services in Schools, 5, 71-78.
- LONGHURST, T.M. and SCHRANDT, T.A.M. (1973). Linguistic analysis of children's speech: A comparison of four procedures. Journal of Speech and Hearing Disorders, 38, (2), 240-249.
- MCCARTHY, D.A. (1954). Language development in children. In L. Carmichael (1954), Manual of child psychology. New York: Wiley.
- MCLEAN, J.E. and SNYDER-MCLEAN, L.K. (1978). A transactional approach to early language training. Columbus OH: Charles E. Merrill Publishing Company.
- MCCLUSKEY, K. (1984). Developmental sentence scoring: A comparative study conducted in Portland, Oregon. Unpublished Master's thesis, Portland State University.
- MCNUTT, E. (1985). A comparative study of the developmental sentence scoring normative data. Unpublished Master's thesis, Portland State University.
- NICE, M. (1925). Length of sentences as a criterion of a child's progress in speech. Journal of Educational Psychology, 16, 370-379.
- O'GRADY, A. and THROOP, F. The three billy goats gruff. Chicago IL: Rand McNally and Co.
- PETERSON, H.A. and MARQUARDT, T.P. (1981). Appraisal and diagnosis of speech and language disorders. Englewood Cliffs NJ: Prentice-Hall, Inc.
- ROBINSON, H.M., MONROE, M., and ARTLEY, A.S. (1962). We read pictures, we read more pictures, before we read. Chicago IL: Scott, Foresman.
- REGISTER, S.H. (1975). Three variations of the storytelling task for language sample elicitation. Unpublished Master's thesis, University of North Carolina at Chapel Hill. In S. Barrie-Blackley et al. (1978), Clinical oral language sampling. Danville IL: The Interstate Printers and Publishers, Inc.
- SHRINER, T.H. and SHERMAN, D. (1967). An equation for assessing language development. Journal of Speech and Hearing Research, 10, 21-40.
- SPECTOR, P.E. (1981). Research designs. Beverly Hills CA: Sage Publications, Inc.

- STALNAKER, L.D. and CREAGHEAD, N.A. (1982). An examination of language samples obtained under three experimental conditions. Language, Speech, and Hearing Services in Schools, 13, 116-120.
- STRANDBERG, T.E. and GRIFFITH, J. (1969). A study of the effects of training in visual literacy on verbal language behavior. Journal of Communication Disorders, 2, 252-263.
- TEMPLIN, M.C. (1957). Certain language skills in children (Institute of Child Welfare Monograph Series, No. 26). Minneapolis: University of Minnesota Press.
- TILDEN-BROWNING, S. (1985). A comparative study of the developmental sentence scoring normative data. Unpublished Master's thesis, Portland State University.
- UTLEY, J. (1950). What's its name? "The three Bears." Urbana IL: University of Illinois Press.
- WERNER, E. and KRESHECK, J. (1974). Structured photographic language test. Sandwich IL: Janelle Publications.
- WERNER, E.O. and KRESHECK, J.D. (1981). Variability on scores, structures, and errors on three measures of expressive language. Language, Speech, and Hearing Services in Schools, 12, 82-89.
- WREN, C.T. (1985). Collecting language samples from children with syntax problems. Language, Speech, and Hearing Services in Schools, 16, 83-102.

APPENDIX A

Dear Parent:

I am a graduate student at Portland State University and am conducting a study regarding language development in children. I am attempting to determine if children use different types of language when 1) playing with toys, 2) describing pictures, or 3) telling stories. I am currently looking for children to participate in this study.

If you agree, your child will receive a hearing screening and a picture vocabulary test; following this, your child will participate in three conversations with myself at your child's preschool. Your child's participation time will consist of approximately four 20-minute sessions.

Your child's name and performance will be kept confidential. You are free to withdraw your child from participation in this study at any time without jeopardizing your relationship with Portland State University.

Your child's participation in this study may help to broaden the knowledge of language development. This knowledge may be used to help children now and in future generations.

If you agree to let your child participate in this study, please sign the consent form below and return this form to your child's teacher by September 16, 1985.

This study is being run under the direction of Dr. Robert Casteel, Portland State University; if you have any questions, please feel free to call him at 229-3533 or myself at 239-9916. If you experience any problems that are a result of your participation in this study, please contact Victor C. Dahl, Office of Graduate Studies and Research, 105 Neuberger Hall, Portland State University, 229-3423.

Thank you,

Cheryl Dong
Graduate Student, Speech and Hearing Sciences
Portland State University

Child's Name _____

Child's Birthdate _____

I agree to let my child participate in this study as outlined above.

Parent/Guardian Signature: _____

Date: _____

APPENDIX B

DEVELOPMENTAL SENTENCE SCORING CATEGORIES
AND REWEIGHTED SCORES

| SCORE | INDEFINITE PRONOUNS OR NOUN MODIFIERS | PERSONAL PRONOUNS | MAIN VERBS | SECONDARY VERBS |
|-------|--|---|---|--|
| 1 | it, this, that | 1st and 2nd person: I, me, my, mine, you, your(s) | A. Uninflected verb: <i>I see you.</i> B. copula, is or 's: <i>It's red.</i> C. is + verb + ing: <i>He is coming.</i> | |
| 2 | | 3rd person: he, him, his, she, her, hers | A. -s and -ed: <i>plays, played</i> B. Irregular past: <i>ate, saw</i> C. Copula: <i>am, are, was, were</i> D. Auxiliary <i>am, are, was, were</i> | Five early-developing infinitives: <i>I wanna see (want to see)</i> <i>I'm gonna see (going to see)</i> <i>I gotta see (got to see)</i> <i>Lemme [to] see (let me [to] see)</i> <i>Let's [to] play (let [us to] play)</i> |
| 3 | A. no, some, more, all, lot(s), one(s), two (etc.), other(s), another B. something, somebody, someone | A. Plurals: <i>we, us, our(s), they, them, their</i> B. these, those | | Non-complementing infinitives: <i>I stopped to play.</i> <i>I'm afraid to look.</i> <i>It's hard to do that.</i> |
| 4 | nothing, nobody, none, no one | | A. can, will, may + verb: <i>may go</i> B. Obligatory do + verb: <i>don't go</i> C. Emphatic do + verb: <i>I do see.</i> | Participle, present or past: <i>I see a boy running.</i> <i>I found the toy broken.</i> |
| 5 | | Reflexives: <i>myself, yourself, himself, herself, itself, themselves</i> | | A. Early infinitival complements with differing subjects in kernels: <i>I want you to come.</i> <i>Let him [to] see.</i> B. Later infinitival complements: <i>I had to go. I told him to go. I tried to go. He ought to go.</i> C. Obligatory deletions: <i>Make it [to] go.</i> <i>I'd better [to] go.</i> D. Infinitive with wh-word: <i>I know what to get.</i> <i>I know how to do it.</i> |
| 6 | | A. Wh-pronouns: <i>who, which, whose, whom, what, that, how many, how much</i> <i>I know who came. That's what I said.</i> B. Wh-word + infinitive: <i>I know what to do. I know whom/ to take</i> | A. could, would, should, might + verb: <i>went come, could be</i> B. Obligatory does, did + verb C. Emphatic does, did + verb | |
| 7 | A. any, anything, anybody, anyone B. every, everything, everybody, everyone C. both, few, many, each, several, most, least, much, next, first, last, second (etc.) | (his) own, one, oneself, whichever, whoever, whatever <i>Take whatever you like.</i> | A. Passive with <i>get</i> , any tense Passive with <i>be</i> , any tense B. must, shall + verb: <i>must come</i> C. have + verb + en: <i>I've eaten</i> D. have got: <i>I've got it.</i> | Passive infinitival complement: With <i>get</i> : <i>I have to get dressed.</i> <i>I don't want to get hurt.</i> With <i>be</i> : <i>I want to be pulled.</i> <i>It's going to be locked.</i> |
| 8 | | | A. have been + verb + ing <i>had been + verb + ing</i> B. modal + have + verb + en: <i>may have eaten</i> C. modal + be + verb + ing: <i>could be playing</i> D. Other auxiliary combinations: <i>should have been sleeping</i> | Gerund: <i>Swimming is fun.</i> <i>I like fishing.</i> <i>He started laughing.</i> |

| SCORE | NEGATIVES | CONJUNCTIONS | INTERROGATIVE REVERSALS | WH-QUESTIONS |
|-------|--|---|---|--|
| 1 | it, this, that + copula or auxiliary is, 's, + not: It's <i>not</i> mine. This is <i>not</i> a dog. That is <i>not</i> moving. | | Reversal of copula: Isn't it red? Where <i>they</i> there? | |
| 2 | | | | A. who, what, what + noun: Who <i>is</i> in? What <i>is</i> he eating? What <i>book</i> are you reading? B. where, how many, how much, what ... do, what ... for: Where <i>did</i> it go? How <i>much</i> do you want? What <i>is</i> he doing? What <i>is</i> a hammer <i>for</i> ? |
| 3 | | and | | |
| 4 | can't, don't | | Reversal of auxiliary be: Is he coming? Isn't he coming? Was he going? Wasn't he going? | |
| 5 | isn't, won't | A. but B. so, and so, so that C. or, if | | when, how, how + adjective: When <i>shall</i> I come? How <i>do</i> you do it? How <i>big</i> is it? |
| 6 | | because | A. Obligatory do, does, did: Do they run? Does it bite? Didn't it hurt? B. Reversal of modal: Can you play? Won't it hurt? Shall I sit down? C. Tag question: It's fun, isn't it? It isn't fun, is it? | |
| 7 | All other negatives: A. Uncontracted negatives: I can <i>not</i> go. He has <i>not</i> gone. B. Pronoun-auxiliary or pronoun-copula contraction: I'm <i>not</i> coming. He's <i>not</i> here. C. Auxiliary-negative or copula-negative contraction: He <i>wasn't</i> going. He <i>hasn't</i> been seen. It <i>couldn't</i> be mine. They <i>aren't</i> big. | | | why, what if, how come how about + gerund: Why <i>are</i> you crying? What <i>if</i> I won't do it? How <i>come</i> he is crying? How <i>about</i> coming with me? |
| 8 | | A. where, when, how, while, whether (or not), till, until, unless, since, before, after, for, as, as + adjective + as, as if, like, that, than I know <i>where</i> you are. Don't come <i>till</i> I call. B. Obligatory deletion: I run faster <i>than</i> you [run]. I'm <i>as big as</i> a man [is big]. It looks <i>like</i> a dog [looks]. C. Elliptical deletions (score 0): That's <i>why</i> [I took it]. I know <i>how</i> [I can do it]. | A. Reversal of auxiliary have: Has he seen you? B. Reversal with two or three auxiliaries: Has he <i>been</i> eating? Couldn't he <i>have</i> waited? Could he <i>have been</i> crying? Wouldn't he <i>have been</i> going? | whose, which, which + noun: Whose <i>car</i> is that? Which <i>book</i> do you want? |

(from Lee, Developmental Sentence Analysis, 1974)

APPENDIX C

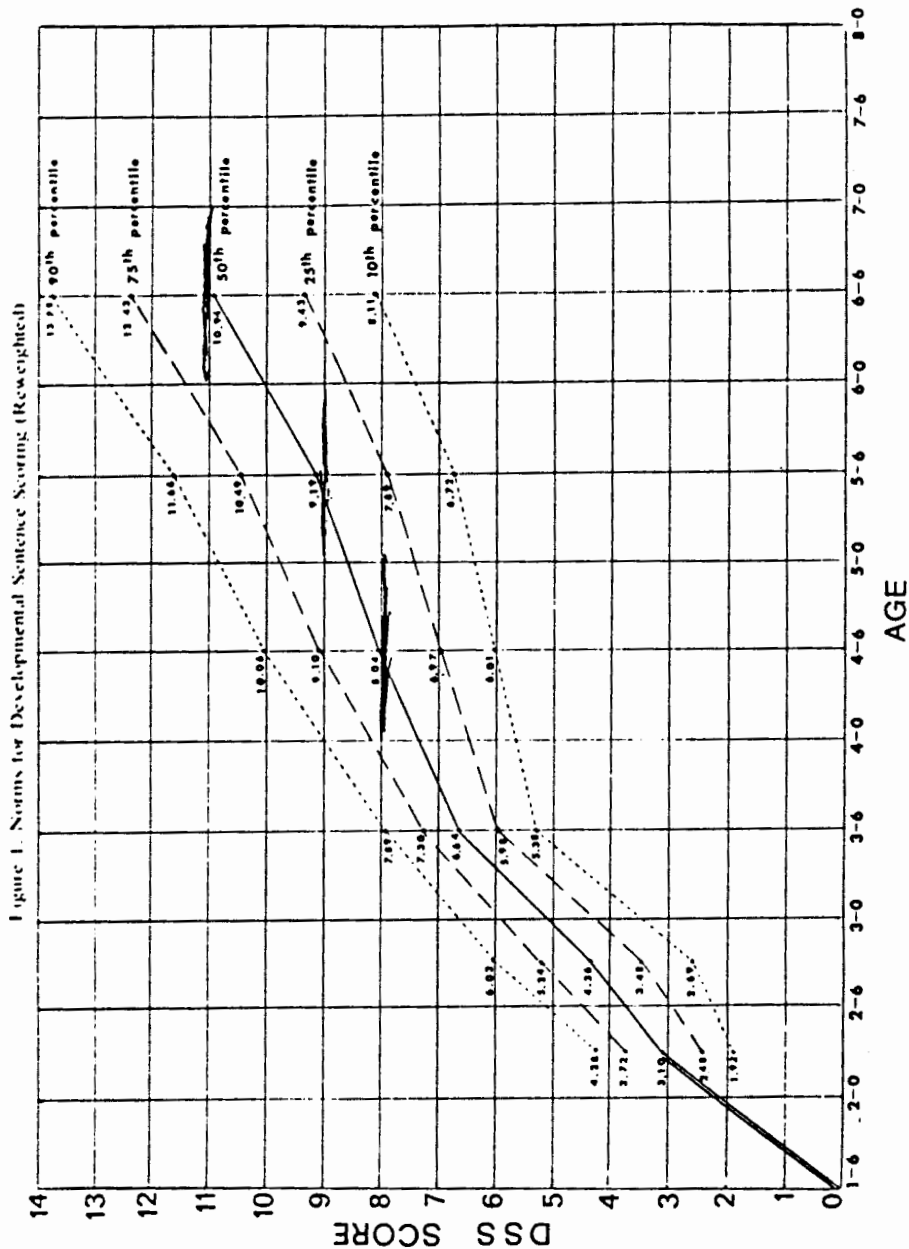
DEVELOPMENTAL SENTENCE SCORES (DSS) FOR EIGHTEEN
SUBJECTS IN SIX ORDERS AND THREE CONDITIONS

| Group | Order | Subject | | |
|-------|--------------|---------|--------|--------|
| | | 1 | 2 | 9 |
| I | Toys | 7.24** | 7.22 | 8.22** |
| | Pictures | 5.90 | 5.66 | 7.58 |
| | Storytelling | 7.24** | 12.58* | 7.54 |
| | | 3 | 7 | 8 |
| II | Pictures | 8.50 | 5.74 | 7.64 |
| | Toys | 11.58* | 7.76 | 7.24 |
| | Storytelling | 9.92 | 8.54* | 7.88* |
| | | 16 | 17 | 18 |
| III | Storytelling | 9.16* | 9.02* | 10.40* |
| | Toys | 8.88 | 7.58 | 7.96 |
| | Pictures | 5.82 | 5.28 | 6.10 |
| | | 10 | 11 | 12 |
| IV | Toys | 4.94 | 7.60 | 6.92 |
| | Storytelling | 6.28* | 11.88 | 8.18 |
| | Pictures | 6.02 | 13.30* | 9.58* |
| | | 4 | 5 | 6 |
| V | Pictures | 6.66 | 4.54 | 6.00 |
| | Storytelling | 8.78* | 5.92 | 13.90* |
| | Toys | 7.98 | 7.70* | 8.42 |
| | | 13 | 14 | 15 |
| VI | Storytelling | 6.46 | 10.04 | 6.64 |
| | Pictures | 5.08 | 7.82 | 6.46 |
| | Toys | 7.62* | 10.42* | 9.06* |

* = Highest DSS among the three scores for the individual subject.
** = Tied highest DSS.

APPENDIX D

LEE'S "NORMS FOR DEVELOPMENTAL SENTENCE SCORING"



(from Lee, Developmental Sentence Analysis, 1974a)

APPENDIX E

DSS SCORES AND MEANS ACCORDING TO AGE

| Age | Toys | Pictures | Storytelling |
|-----------|-------|----------|--------------|
| 3.9 | 7.76 | 5.74 | 8.54 |
| 3.9 | 7.24 | 7.64 | 7.88 |
| 3.6 | 8.22 | 7.58 | 7.54 |
| 3.6 | 4.94 | 6.02 | 6.28 |
| 3.9 | 7.62 | 5.08 | 6.46 |
| 3.9 | 10.42 | 7.82 | 10.04 |
| 3.6 | 7.96 | 6.10 | 10.40 |
| \bar{x} | 7.73 | 6.56 | 8.16 |
| 4.4 | 11.58 | 8.50 | 9.92 |
| 4.6 | 7.98 | 6.66 | 8.78 |
| 4.6 | 7.60 | 13.30 | 11.88 |
| 4.3 | 6.92 | 9.58 | 8.18 |
| 4.2 | 9.06 | 6.46 | 6.64 |
| 4.2 | 8.88 | 5.82 | 9.16 |
| 4.8 | 7.58 | 5.28 | 9.02 |
| \bar{x} | 8.51 | 7.94 | 9.08 |
| 5.1 | 7.24 | 5.90 | 7.24 |
| 5.6 | 7.22 | 5.66 | 12.58 |
| 5.0 | 7.70 | 4.54 | 5.92 |
| 5.0 | 8.42 | 6.00 | 13.90 |
| \bar{x} | 7.64 | 5.52 | 9.90 |