Development of Age Normative Data for the Oral Language Sentence Imitation Screening Test

Melinda Diane Roos
Abstract

DEVELOPMENT OF AGE NORMATIVE DATA
FOR THE ORAL LANGUAGE SENTENCE IMITATION SCREENING TEST

By Melinda Diane Roos

The purpose of this investigation was to develop objective pass/fail age-related scores for the Oral Language Sentence Imitation Screening Test (OLSIST). A group of 60 children, ages five years zero months to five years eleven months, served as subjects. The Denver Articulation Screening Exam (DASE), and the Screening Test For Auditory Comprehension of Language (STACL) were used to determine if each child's articulation and receptive language abilities were within normal limits. Each subject's scores on the OLSIST were then compared with his/her score on the CELI to determine the degree of correlation in pass/fail performances on the two tests.

The results of this investigation indicated a strong correlation (.94) between the CELI and the OLSIST. Because of the high correlation found, a Z score was used to determine a cutoff score of seven for the OLSIST. A subject having from zero to seven errors is considered within normal range and requires no further testing. A total of eight or more errors indicates that the subject has failed this screening examination and would require additional testing.

This study indicates that the OLSIST is an efficient and reliable method of screening the expressive language abilities of children ages five years zero months to five years eleven months. The OLSIST
yields much the same screening information as the CELI but is more time-efficient.
DEVELOPMENT OF AGE NORMATIVE DATA
FOR THE ORAL LANGUAGE SENTENCE ImitATION SCREENING TEST

By
Melinda Diane Roos

A Thesis in Partial Fulfillment of the
Requirements for the Degree Master of Science
In the Field of Speech Pathology

May 1980
Each person whose signature appears below certifies that this thesis in his opinion is adequate, in scope and quality, as a thesis for the degree Master of Science.

Melvin S. Cohen, Associate Professor of Speech Pathology

Jack Hartley, Assistant Professor of Speech Pathology

Floyd Wood, Associate Professor of Counselor Education
# TABLE OF CONTENTS

## Chapter

1. **THE NATURE AND SCOPE OF THE PROBLEM** .......................... 1
   - THE PROBLEM .................................................. 4
   - HYPOTHESES .................................................... 4

2. **REVIEW OF THE LITERATURE** ............................................. 6
   - SPONTANEOUS LANGUAGE SAMPLES .................................. 6
   - STORY COMPLETIONS .............................................. 7
   - SENTENCE IMITATION .............................................. 7

3. **RESEARCH DESIGN AND PROCEDURES** ..................................... 11
   - POPULATION AND SAMPLE ......................................... 11
   - MATERIALS AND SOURCES ......................................... 12
   - METHODOLOGY ...................................................... 12

4. **RESULTS** ............................................................................. 15
   - ANALYSIS OF DATA ................................................... 15
     - *t* Test ................................................................. 15
     - Wilcoxon *t* Test - For Paired Ranks .......................... 16
     - Histogram and Frequency Distribution ....................... 16

5. **DISCUSSION** ........................................................................ 19

6. **SUMMARY AND CONCLUSIONS** ............................................. 23

REFERENCES ............................................................................... 24
Chapter 1

THE NATURE AND SCOPE OF THE PROBLEM

Three procedures are currently used to assess the linguistic performance of the language-impaired child: (1) spontaneous language sample, (2) story completion, and (3) elicited language sample. No single procedure can provide all the needed information about a child's expressive language abilities, but there is a definite need to use objective standards in describing and quantifying language skills of children (Carrow, 1974).

Although the spontaneous language sample yields a wide variety of information, various researchers have cited its limitations. Sharf (1972) reported three specific limitations: (1) This type of sampling is time-consuming for the examiner, particularly with the scoring procedures; (2) The child's understanding of the grammar, also called linguistic competence, is greater than actual spontaneous use of the language, or linguistic performance; and (3) There are differences between a child's linguistic competence and spontaneous production. Thus, the language sample only provides information concerning the language structures which the child uses, and does not explore the potential linguistic forms the child might be able to use in a different situation (Bliss, et al., 1977).

The "story completion" method was designed as an alternative procedure to the spontaneous language sample in the assessment of language development. This type of procedure involves the examiner
reading a "story" aloud to the subject, with an accompanying picture display. Each story ends with a question or a phrase, intended to elicit a target response. Eliciting a specific grammatical framework provides a more structured approach to language assessment, and allows easy scoring because of the objective fashion in which the test elicits the responses (Bliss, et al., 1977).

Several limitations exist with the story completion method. The grammatical concepts are tested only once or twice and this fails to give an in-depth sample of the child's linguistic capabilities. It has also been found that boys require more cuing than girls. This type of testing, which was discussed by Bliss (1977), requires more investigation, and limitations of this are not currently discussed in the literature.

The sentence imitation test provides useful information for the assessment of receptive and expressive language skills of children (Menyuk, 1964; Lennenberg, 1967). This is a reliable method of obtaining information about the child's grammatical system. It samples the child's linguistic capabilities and allows identification of the ways in which the child's performance differs from normal adult speech patterns.

Erwin (1964) found no differences between the grammar used by children in imitations of adult utterances and that which they used in spontaneous speech. McNeill (1970) stated:

The child tends to omit from the surface structure those linguistic elements which cannot be related to deep structures. That is, children will reproduce a sentence using the rules they know. They filter it through their own productive systems.
The Carrow Elicited Language Inventory, CELI, (Carrow, 1972) is a diagnostic tool which uses a sentence imitation task to measure children's grammatical structures. One limitation of this test is that some grammatical concepts are tested only once or twice. This fails to give an in-depth sample of the child's linguistic capabilities. Another shortcoming is related to the test's use of lengthy sentences, which assumes that the child's auditory memory is intact.

The Oral Language Sentence Imitation Screening Test, OLSIST, also uses the sentence imitation method to evaluate language performance. However, the OLSIST tests each grammatical category in greater depth than the CELI. The OLSIST is available in three forms: Stages III, IV, and V. The tests were constructed in order to be compatible with the test subject's auditory memory skills. This allows the clinician to assess the child's expressive language without the interference of insufficient auditory memory. If any child fails Stage V, it is suggested that Stage IV may be administered to determine language competency. Passing Stage IV subsequent to a failure on Stage V would indicate that sentence length was the determining factor in the initial failure.

The OLSIST also provides information about the core of expressive language structures within a given developmental stage. It appears to be an efficient screening method which should provide the speech-language pathologist with enough information to determine if the child's expressive language skills are within normal limits or if they warrant further testing. However, objective pass/fail scores have not been established for the instrument.
The developers, Zachman, et al. (1976, 1977) stated that the clinician can determine whether the child has passed, failed, or achieved a borderline performance by the total score obtained. The scoring instructions state:

**Pass:** Few test errors overall. Random distribution of errors. Demonstrates good understanding and usage of all test structures.

**Fail:** Numerous test errors overall. Distribution of errors may be throughout test or within many test structure categories. Consistently reduces length and/or simplifies complexity.

**Borderline:** Several test errors overall. Distribution of errors may be throughout test or within few test structure categories. Inconsistently or randomly reduces length and/or simplifies complexity. 

This subjective scoring procedure has been found to be a limitation of the test for the diagnostician who relies on percentile rankings or standard deviation scores in determining pass/fail criteria.

**THE PROBLEM**

The purpose of this study was to determine a pass/fail cutoff score for the Oral Language Sentence Imitation Screening Test and to evaluate the concurrent validity of the test using the Carrow Elicited Language Inventory as the comparison criteria.

**HYPOTHESES**

It is hypothesized:

1) That by equating scores on the OLSIST with scores on the CELI, a pass/fail cutoff score for the OLSIST can be determined which is comparable to the cutoff score on the CELI;
2) That the concurrent validity of the OLSIST in comparison with the CELI is positive as indicated by a Pearson Product Monument Correlation Coefficient in excess of 0.60 (Cronbach, 1970, p. 135).

It is assumed that when hypotheses 1 and 2 are tenable, evidence is sufficient to warrant the substitution of the OLSIST for deciding whether a child's language capabilities are within normal limits as usually determined by the CELI.
Chapter 2

REVIEW OF THE LITERATURE

The professional literature which discusses the usefulness of spontaneous language samples, story completion tests, and elicited language samples is varied.

SPONTANEOUS LANGUAGE SAMPLES

A spontaneous speech sample is one of the major diagnostic tools used to assess language performance. Lee and Canter (1971) claim that,

by analyzing a child's spontaneous, tape-recorded speech sample, a clinician can estimate to what extent the child has generalized the grammatical rules sufficiently to use them in verbal performance. p. 315

They observed that in spontaneous speech a child may be inconsistent in his use of grammatical forms which he accomplished within the simplicity of a structured test. Something more than standardized tests was needed to evaluate "the child's consistency and frequency of usage and his ability to combine many transformations into a single sentence in spontaneous speech." p. 316

Shriner (1969) describes variations in the quality and quantity of language samples resulting from differences in examiners, stimulus materials, elicitation situations, and sample sizes. This same view concerning language samples is shared by Wilson (1969). She states:

None of these methods is standardized with regard to either: (1) a definite set of instructions for the examiner and for
the examiner to give the subject, or, (2) a standardized set of stimulus materials easily available and convenient to use. Many of the methods do not appear to be designed to elicit representative speech and may actually encourage naming responses and short sentences. p. 95

**STORY COMPLETIONS**

In search of an improved design for studying language development, Bliss, et al. (1977) developed an alternate approach which incorporated the advantages of both the spontaneous sampling and the imitation methods. Using an objective procedure with easy scoring, the story completion method elicits specific grammatical structures which represent differing levels in linguistic development.

**SENTENCE IMITATION**

The performance data obtained through elicited imitation of a symmetrically developed sequence of sentences can provide the objective standard for describing and quantifying the language patterns of a child (Carrow, 1974). Imitation tests are considered to cover a wide range of linguistic complexity, and give information not only concerning what the child does say, but about that which he is capable of saying. "It allows analysis of forms over which the child has productive control but which may not occur in a particular language sample." (Carrow, 1974, p. 439).

Lutterman and Barr agree that the procedure of sentence repetition seems to "provide a valuable, simple tool to assess the linguistic performance of a language impaired subject." p. 29 It allows the individual's current level of performance to be compared with his own
previous productions. The interviewer is also able to evaluate grammatical performance estimates on each subject's linguistic competence.

Research studies on sentence imitation by Freedle, et al. (1970) and Rodd and Braine (1970) indicated that the child is assimilating and reproducing what he has heard. This yields useful information concerning linguistic competence. "Such research suggests that imitation could be used systematically as a tool for testing specific hypotheses about rules the child knows." (Rodd and Braine, 1970, p. 431).

Menyuk (1963b) examined grammatical capacity in children. In her study she assumed that language production is not an imitative function, and consequently did not rely on grammatical production for evaluation of grammatical capacity. She used the procedure of asking the children to repeat sentences containing syntactic structures found in the language they produce.

Menyuk (1964) again used sentence imitation in a study comparing the grammar of children having functionally deviant speech with those having normal speech. Children from each group were asked to repeat a list of sentences containing syntactic structures typically found in children's grammar. She observed that sentence length ranging from two to nine words did not seem to interfere with a child's normal ability to repeat sentences. However, it did interfere with the ability of children having language impairments to repeat sentences. Consequently it appears that, whereas the syntactic structure of a sentence facilitates the repetition ability of the normal child, it does not help the child who is language-impaired.

Other researchers agree with Menyuk (Erwin, 1964; McNiell, 1968;
and Odom, Leibert and Hills, 1968). The child is unable to imitate the appropriate features of language unless he has the important parts of syntax and grammar mastered on a receptive level.

Of those holding views contrary to the sentence imitation method, Brown and Fraser (1963) hypothesize that because children have such small memory spans, function words may be dropped because they carry little information and tend to be unstressed in speech. Children do not attempt to store these unstressed, low information words and will delete them when asked to imitate sentences. However, Scholes (1969) obtained results which suggest that relative stress may not explain children's deletions of function words in an immediate recall task which uses grammatical and ungrammatical word strings as stimuli.

The Oral Language Sentence Imitation Screening Test, OLSIST, is one example of the sentence imitation method used in expressive language assessment. Currently, the OLSIST does not have objective pass/fail criteria, but relies on generalized performance guidelines. Barrett (1980) gave several reasons why he and the other developers of the OLSIST decided to "eliminate" numerical cutoff scores: (1) Totaling up errors was felt to be time consuming for the examiner; (2) One cutoff score may not work for all populations tested; and (3) It was stated that "the clinician should exercise clinical judgment in the final classification of a child's performance." (Zachman, 1977, p. 8). If the child had several errors which were grouped in one or two categories he could possibly fail the test. The same number of errors distributed over a number of categories may be considered passing.
No single testing procedure can provide all of the needed information concerning a child's language performance. Therefore, a combination of tests should be available. The speech-language pathologist's background and available time will probably be influential in determining which tests will be administered to any given child.

Considering the need for more objective evaluation measurements, the establishment of objective pass/fail scores for the OLSIST may provide the speech-language pathologist with a more practical assessment tool.
Chapter 3

RESEARCH DESIGN AND PROCEDURES

The purpose of this investigation was to develop objective pass/fail age-related scores for the Oral Language Sentence Imitation Screening Test. A cutoff score may then be used to determine whether the child's language competency is "within normal limits" for chronological age or below the expected achievement level and in need of further evaluation. Each subject's score on the OLSIST was compared with his/her score on the Carrow Elicited Language Inventory to determine the degree of correlation in pass/fail performances on the two tests.

POPULATION AND SAMPLE

A group of 60 children who were not exhibiting speech or language difficulties, ages five years zero months to five years eleven months, with an equal distribution of males and females, served as subjects. The Denver Articulation Screening Exam, DASE, (Drumwright, 1971) and the Screening Test for Auditory Comprehension of Language, STACL, (Carrow, 1973) were used to ensure that each child's articulation and receptive language abilities were within normal limits. The subjects were from middle-socioeconomic backgrounds where Standard American English is the only language spoken in their homes. Those speaking English as a second language or where Spanish is spoken in the home were not included as subjects for this study.
MATERIALS AND SOURCES

The Denver Articulation Screening Exam and the Screening Test for Auditory Comprehension of Language were chosen to determine normalcy in this investigation. Both tests are easily and economically administered to identify those individuals who have a high probability of being handicapped in articulation or receptive language.

Stage V of the OLSIST was chosen for this study because of the need for an efficient language screening assessment tool for the speech-language pathologist to use with kindergarten children in the school system. Stage V is constructed of 20 sentences which test 23 morphological and grammatical language structures. The sentences range from six to thirteen morphemes in length, with a mean length of 8.5 morphemes. There are over 107 possible grammatical errors in 23 categories.

The Carrow Elicited Language Inventory (1972) consists of 51 sentences and one phrase. The sentences range in length from two to ten words, with an average length of six words. There are a total of 321 possible grammatical errors on the CELI. These errors are then categorized by type (substitution, omission, etc.). The number of grammatical errors then equals the number of "type" errors.

METHODOLOGY

The Denver Articulation Screening Exam and the Screening Test for Auditory Comprehension of Language were administered to 60 children ages five years zero months to five years eleven months, from middle-
socioeconomic backgrounds where Standard American English is the only language spoken in their homes. The children were divided into two groups. The order for test administration for group one was:

1) STACL,
2) DASE,
3) OLSIST,
4) CELI.

The order for group two was:

1) STACL,
2) DASE,
3) CELI,
4) OLSIST.

The difference in test order was to prevent artifacts from occurring because of an ordering effect.

A pilot study was conducted with three children in order to familiarize the examiner with testing procedures before the administration of testing began with the 60 research subjects.

Each subject was asked to imitate each OLSIST stimulus sentence after it was spoken by the examiner. All responses made by the child were tape-recorded on a Bell & Howell cassette recorder, and all scores were transcribed from the subject's actual face-to-face production. If word substitutions, word omissions, or changes in word order occurred, the examiner struck through the original word and wrote the subject's complete response above the target sentence on the score sheet. Errors were transferred to the "test structure" score sheet in order to locate the category where the error was made.

The CELI test administration consisted of the examiner's production of one stimulus sentence at a time, which the child attempted to reproduce on a direct imitation basis. The child's imitations of
the sentences were recorded on audiotape. All responses were transcribed onto a matrix-type for classification of the grammatical features.

Data from the OLSIST were analyzed, with scores being arranged in order from highest to lowest. All scores which were one standard deviation below the mean were considered failing and indicative of the student's requiring further testing.

Once the pass/fail scores were developed, each child's scores on the OLSIST and the CELI were compared to determine the degree of correlation.
Chapter 4

RESULTS

A group of 60 middle-socioeconomic status children between the ages of five years zero months to five years eleven months served as subjects for this study. The Denver Articulation Screening Exam, DASE, and the Screening Test for Auditory Comprehension of Language, STAACL, were used to determine whether the child's articulation and receptive language abilities were within normal limits. The 60 children who passed the two screening tests were then evaluated with the Carrow Elicited Language Inventory, CELI, and the Oral Language Sentence Imitation Screening Test, OLSIST.

Scores from the CELI and the OLSIST were arranged in order from highest to lowest, and compared by a $t$ Test, and a Histogram. It was found that a high positive correlation existed between the subject's scores on the CELI and the OLSIST.

ANALYSIS OF DATA

$t$ Test

A $t$ Test was used to test the hypothesis that there is a significant difference in the scores of the OLSIST and the CELI. The results of that test gave a degree of freedom of 59, which corresponds to a $t$ ratio of 7.32. The level of significance, $P < .001$, shows a highly significant relationship between the scores obtained on the CELI and the OLSIST. The data accounted for 0.88% of the total variance
between the scores through $\gamma = 0.94$. This reveals that there are very few (12%) errors that have not been accounted for.

**Wilcoxon t Test - For Paired Ranks**

The number of errors each subject made on the CELI was compared to the errors made on the OLSIST. This test revealed that all students, with the exception of one, scored more errors on the CELI than the OLSIST. This is not unusual, however, because there are $321 \times 2$ possible items on the CELI and 107 items on the OLSIST. As the number of errors increased on the CELI, the errors also increased on the OLSIST.

**Histogram and Frequency Distribution**

An interval size of 15 was used to show the frequency of distribution of scores on the CELI. A score interval size of five was used for the OLSIST.

**HISTOGRAM FOR CELI**

<table>
<thead>
<tr>
<th>N</th>
<th>F</th>
<th>CF</th>
</tr>
</thead>
<tbody>
<tr>
<td>0 - 14</td>
<td>17</td>
<td>17</td>
</tr>
<tr>
<td>15 - 29</td>
<td>20</td>
<td>37</td>
</tr>
<tr>
<td>30 - 44</td>
<td>11</td>
<td>48</td>
</tr>
<tr>
<td>45 - 59</td>
<td>3</td>
<td>51</td>
</tr>
<tr>
<td>60 - 74</td>
<td>2</td>
<td>53</td>
</tr>
<tr>
<td>75 - 89</td>
<td>1</td>
<td>54</td>
</tr>
<tr>
<td>90 - 104</td>
<td>3</td>
<td>57</td>
</tr>
<tr>
<td>105 - 119</td>
<td>0</td>
<td>57</td>
</tr>
<tr>
<td>120 - 134</td>
<td>1</td>
<td>58</td>
</tr>
<tr>
<td>135 - 149</td>
<td>0</td>
<td>58</td>
</tr>
<tr>
<td>150 - 164</td>
<td>0</td>
<td>58</td>
</tr>
<tr>
<td>165 - 179</td>
<td>1</td>
<td>59</td>
</tr>
<tr>
<td>180 - 194</td>
<td>1</td>
<td>60</td>
</tr>
</tbody>
</table>

N = number of errors
* = number of children
F = frequency
CF = cumulative frequency
These data show a high level of correlation between the OLSIST and the CELI. This indicates that the subject's performances on the two tests were quite similar. Also, the information gained from the CELI is much the same as that obtained from the OLSIST.

The histograms for the OLSIST and the CELI reveal that, of the 60 subjects examined, three children had an extremely large number of errors. Subject number 58 produced 128 errors on the CELI and 45 errors on the OLSIST. Subject 59 produced 172 errors on the CELI and 43 errors on the OLSIST. Subject 60 produced 182 errors on the CELI and 43 errors on the OLSIST. These scores are not typical of the population used in a norming sample. Development of a mean including these three atypical scores deceptively increased the mean for the CELI and the OLSIST. For this reason the lowest three scores were deleted and means were computed on a sample of 57 subjects.

The mean number of errors for the OLSIST is 7.45. One standard deviation is 5.47. The correlation between the CELI and the OLSIST,
based on a population of 57, is equal to 0.845. Because of the high correlation found between the CELI and the OLSIST, a Z score was used to determine the cutoff score for the OLSIST. A pass/fail cutoff score of seven errors is substantiated by this study.

The mean for the CELI is 28.91. One standard deviation is ±22. The suggested mean stated in the CELI manual is 14 for the age group of five years zero months to five years eleven months. The 14-point difference between the mean number of errors as stated in the CELI manual and that found by this study is explained by two factors. First, the CELI was normed on a sample size of 475 children while this study used a sample of 60. Second, there is a difference in the population samples examined. Carrow (1973) stated that the scores will vary from her standardized norms for tests given in different parts of the country and for tests administered to subjects of differing socioeconomic levels.

<table>
<thead>
<tr>
<th></th>
<th>N</th>
<th>MEAN</th>
<th>S.D.</th>
</tr>
</thead>
<tbody>
<tr>
<td>CELI</td>
<td>60</td>
<td>35.43</td>
<td>36.27</td>
</tr>
<tr>
<td>OLSIST</td>
<td>60</td>
<td>9.32</td>
<td>9.71</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>N</th>
<th>MEAN</th>
<th>S.D.</th>
</tr>
</thead>
<tbody>
<tr>
<td>CELI</td>
<td>57</td>
<td>28.91</td>
<td>22</td>
</tr>
<tr>
<td>OLSIST</td>
<td>57</td>
<td>7.46</td>
<td>5.47</td>
</tr>
</tbody>
</table>

N = number of subjects
S.D. = standard deviation
Chapter 5

DISCUSSION

The sentence imitation method has long been found to be a valuable tool in determining the level of a child's grammatical system. The Carrow Elicited Language Inventory was developed to provide a valid and standardized method of gaining grammatical information. However, the administration, transcription and scoring of this test is very time-consuming.

The OLSIST also uses a sentence imitation method of evaluation, but is designed to be an efficient method of screening for language problems in kindergarten-age children. The estimated time to administer and score the CELI is 45 minutes. A maximum of 10 minutes is needed for administration and scoring of Stage V of the OLSIST. A drawback of the OLSIST has been the lack of objective scoring guidelines to determine the child's level of grammatical usage.

This investigation was designed to develop objective pass/fail age-related scores for the OLSIST, and to compare these scores with the CELI to determine the degree of correlation between the two tests. Screening devices which yield a high degree of information in a minimum amount of time are particularly important in a school setting where large numbers of children must be screened for speech and language difficulties. The age group of five years zero months to five years eleven months, and Stage V of the OLSIST, were chosen for this study because the examiner considered there was a need for a reliable and
efficient expressive language screening method for the kindergarten age group in the schools.

Sixty children, ages five years zero months to five years eleven months with an equal distribution of males and females, were divided into two groups. Group one was given four tests in this order:

1) STACL,
2) DASE,
3) OLSIST,
4) CELI.

Group two was given the tests in this order:

1) STACL,
2) DASE,
3) CELI,
4) OLSIST.

The variation in test order was to prevent artifacts from occurring because of an ordering effect.

A t test was used to determine the degree of correlation of scores between the CELI and the OLSIST. The level of significance found, $P < .001$, shows a highly significant relationship between the scores of the two tests. These data accounted for .88% variance between the scores.

A Histogram which shows frequency of the distribution of scores revealed similar findings for the CELI and the OLSIST. The numbers of errors for both of the tests were grouped on the low end of the curve. As errors increased for the CELI they also increased for the OLSIST. All the children, with the exception of one, had more errors on the CELI than the OLSIST. Both the t test and the Histogram reveal a high level of correlation between the two tests. The skewed normal curve is the result of the subject sample.
The OLSIST manual states that "several errors" constitutes a borderline score. In the present investigation the data suggest students would be able to produce as many as seven errors before additional testing was warranted.

As these differences in error scores were considered large, the authors of the OLSIST were asked to comment. Barrett (1980) indicated that his study was based on the scores of children in the age five to seven years, while the present study examined children who were five years zero months to five years eleven months. Due to the present study's concentrating its sample at the lower end of the recommended age limits for Stage V, a higher number of errors would be likely to occur. Several reasons why the developers of OLSIST decided to "eliminate" numerical cutoff scores were explained by Barrett. Totaling errors was thought to be time-consuming for the examiner. Also, one cutoff score may not be appropriate for all populations tested.

The average time to administer, transcribe and score the CELI is approximately 45 minutes. Although the information gathered is valuable, there are many times when a quick sampling of a child's language capabilities is all that is desired in order to make a pass/fail decision.

The OLSIST yields not only the same type of information as the CELI, but also tests each grammatical category in greater depth. The OLSIST required approximately 10 minutes to administer and score. A tape recording of the administration of the sentences is not required as it is for the CELI.

A subject receiving a score of zero to seven errors on the
OLSIST is considered within a normal range for a child five years zero months to five years eleven months, and requires no further testing. A total of eight errors or more indicates that the subject has failed this screening examination and would require additional testing.

The histograms for the OLSIST and the CELI revealed that, of the 60 subjects, the scores of the three lowest children would need to be deleted. These scores are not typical of a norming sample, and falsely raise the mean for the two tests.

The results of this study, using a pass/fail cutoff score of seven for the OLSIST indicates that, of the 57 students tested, 38 passed and 19 failed. A passing or failing score on the CELI was based on the tenth percentile and below as suggested by the CELI manual. Based on the population of 57, 37 subjects passed and 20 failed.

The results of this study indicate a highly similar pass/fail ratio for the two assessment instruments.

The hypotheses stated for this study were met in all respects.
Chapter 6

SUMMARY AND CONCLUSIONS

The results obtained from this study indicate that the OLSIST is an efficient and reliable method of screening the language abilities of children from ages five years zero months to five years eleven months. The OLSIST yields much the same screening information as the CELI but is more time-efficient.

The CELI's raw scores are converted to percentile scores and stanine scores. Since the OLSIST is used as a screening device, stanine and percentile scores were not developed. Instead, a pass/fail scoring system at seven errors is supported by these data.

As with any diagnostic tool a borderline or failing score on the OLSIST should be used in conjunction with other tests for an adequate diagnosis to be made.

Suggestions for further study in this area would include using the OLSIST as a screening device on a district-wide basis in the schools. These same children would also be given the CELI to substantiate the degree of correlation and the numerical cutoff score developed by this study.
REFERENCES


Carrow, E. Carrow Elicited Language Inventory, 1973.


24


