

ABSTRACT

LANGUAGE INTERVENTION EFFECTS ON NORM-REFERENCED AND CRITERION-REFERENCED TEST SCORES

by Amber Marie Pester

The assessment and treatment of language delay is a widely discussed topic among speech-language pathologists. Language delay is defined as a child's language that is similar to a younger child and will eventually arrive at the desired destination of normal development. Eighteen language-delayed children, enrolled in 5 months of language intervention, participated in this study. The purposes of this investigation were to (a) discern whether there were any significant differences in pre- versus post-intervention test score measurements of language skills, and (b) determine if there was a significant difference for a norm-referenced (Preschool Language Scale-3) versus a criterion-referenced (Receptive-Expressive Emergent Language Test-2) assessment instrument. Results demonstrated both expressive and total language skills were significantly different after 5 months of intervention. Pre-intervention disability levels were found not to inhibit a language-delayed child's improvements from intervention.

Language Intervention Effects of Norm-Referenced and Criterion-Referenced

Test Scores

A Thesis

Submitted to the

Faculty of Miami University

In the partial fulfillment of

The requirements for the degree of

Master of Arts

Department of Speech Pathology and Audiology

By

Amber Marie Pester

Miami University

Oxford, Ohio

2003

Advisor _____
Barbara Weinrich, Ph.D.

Reader _____
Doris Bergen, Ph.D.

Reader _____
Ann Glaser, M.S.

TABLE OF CONTENTS

CHAPTER I: INTRODUCTION

<i>Language Development</i>	1
<i>Language Intervention</i>	2
<i>Intervention Studies</i>	3
<i>Assessment Tools</i>	3
<i>Statement of the Problem</i>	4
<i>Purpose of the Study</i>	5
<i>Significance of the Problem</i>	5
<i>Chapter Summary</i>	6

CHAPTER II: REVIEW OF THE LITERATURE

<i>Language</i>	7
<i>Language Development</i>	9
<i>Language Delay</i>	10
<i>Assessment Tools</i>	11
<i>Norm-Referenced Tests</i>	11
<i>Criterion-Referenced Tests</i>	14
<i>Similarities Between Norm-Referenced and Criterion-Referenced Tests</i>	15
<i>Differences Between Norm-Referenced and Criterion-Referenced Tests</i>	16
<i>Studies Comparing Norm-Referenced and Criterion-Referenced Tests</i>	16
<i>Handicapped versus nonhandicapped</i>	16

<i>Test-retest comparison of screening measures</i>	<i>18</i>
<i>Comparing norm-referenced and criterion-referenced tests in school aged children.....</i>	<i>20</i>
<i>Criterion-referenced test with projected norms</i>	<i>22</i>
<i>Intervention.....</i>	<i>23</i>
<i>Three Purposes of Intervention</i>	<i>24</i>
<i>Effectiveness of Intervention.....</i>	<i>25</i>
<i>Shared reading intervention.....</i>	<i>25</i>
<i>Interactive focused stimulation intervention</i>	<i>25</i>
<i>Mentally handicapped children with intervention</i>	<i>26</i>
<i>Early intervention method effectiveness.....</i>	<i>28</i>
<i>Types of Intervention</i>	<i>29</i>
<i>Play-based intervention.....</i>	<i>30</i>
<i>Clinician-directed intervention</i>	<i>31</i>
<i>Child-centered intervention.....</i>	<i>31</i>
<i>Hybrid approach intervention</i>	<i>32</i>
<i>Studies Showing the Effectiveness of the Intervention Techniques</i>	<i>32</i>
<i>Focused-stimulation in children with language impairment</i>	<i>32</i>
<i>Efficacy of expansions and cloze procedures.....</i>	<i>33</i>
<i>Effects of treatment on linguistic and social skills in toddlers.....</i>	<i>35</i>
<i>Chapter Summary</i>	<i>37</i>

<i>Research Hypotheses</i>	37
CHAPTER III: METHODS AND PROCEDURES	
<i>Subjects</i>	38
<i>Confidentiality of Records</i>	38
<i>Procedures</i>	39
<i>Assessment Tools</i>	39
<i>Intervention</i>	40
<i>Experimental Design</i>	41
<i>Research Questions</i>	41
<i>Null Hypotheses</i>	42
<i>Statistical Analysis</i>	42
<i>Chapter Summary</i>	43
CHAPTER IV: RESULTS	
<i>Demographics</i>	44
<i>Subjects</i>	44
<i>Intervention</i>	44
<i>Descriptive and Inferential Statistics for Research Questions</i>	46
<i>Chapter Summary</i>	60
CHAPTER V: DISCUSSION	
<i>Discussion</i>	62
<i>Conclusions</i>	67

<i>Limitations</i>	68
<i>Implications for Future Research</i>	69
<i>Chapter Summary</i>	70
REFERENCES	72
APPENDIXES	
Appendix A Recruitment Letter	91
Appendix B Consent Form.....	92
Appendix C Subject Intervention Data Form.....	93
Appendix D PLS-3 Pre/Post Intervention Raw Scores	94
Appendix E REEL-2 Pre/Post Intervention Raw Scores	95

LIST OF TABLES

1. Table 1 Subject Demographics	45
2. Table 2 Receptive Data: Whole Sample PLS-3 Post/Pre Difference	48
3. Table 3 Receptive Language Difference Scores Between Pre-Versus Post- Intervention	49
4. Table 4 Expressive Data: Whole Sample PLS-3 Post/Pre Difference	49
5. Table 5 Expressive Language Difference Scores Between Pre-Versus Post-Intervention.....	50
6. Table 6 Total Data: Whole Sample PLS-3 Post/Pre Difference	51
7. Table 7 Total Language Difference Scores Between Pre-Versus Post- Intervention	52
8. Table 8 Receptive Data: Comparison of PLS-3 Improvement Between Norm-Referenced Groups.....	53
9. Table 9 Receptive Data: ANOVA Analysis of PLS-3 Improvement Between Norm-Referenced Groups	54
10. Table 10 Expressive Data: Comparison of PLS-3 Improvement Between Norm-Referenced Groups	54
11. Table 11 Expressive Data: ANOVA Analysis of PLS-3 Improvement Between Norm-Referenced Groups	55
12. Table 12 Total Data: Comparison of PLS-3 Improvement Between Norm- Referenced Groups.....	55

13. Table 13 Total Data: ANOVA Analysis of PLS-3 Improvement Between Norm-Referenced Groups	56
14. Table 14 Receptive Data: Whole Sample REEL-2 Post/Pre Difference ...	57
15. Table 15 Expressive Data: Whole Sample REEL-2 Post/Pre Difference..	58
16. Table 16 Total Data: Whole Sample REEL-2 Post/Pre Difference	59
17. Table 17 PLS-3 Sessions Attended/Test Score Differences	66
18. Table 18 REEL-2 Sessions Attended/Test Score Differences	67

LIST OF FIGURES

1. Figure 1 PLS-3 Receptive Language Difference Scores Between Pre-Versus Post-Intervention.....	48
2. Figure 2 PLS-3 Expressive Language Difference Score Between Pre-Versus Post-Intervention.....	50
3. Figure 3 PLS-3 Total Language Difference Scores Between Pre-Versus Post- Intervention.....	51
4. Figure 4 REEL-2 Receptive Language Difference Scores Between Pre-Versus Post-Intervention.....	58
5. Figure 5 REEL-2 Expressive Language Difference Scores Between Pre-Versus Post-Intervention.....	59
6. Figure 6 REEL-2 Total Language Difference Scores Between Pre-Versus Post- Intervention.....	60

ACKNOWLEDGEMENTS

This thesis would not have been possible without the commitment and support of many individuals. I would like to take this opportunity to thank my thesis committee members. In particular I would like to express my deepest gratitude to my committee chair, Dr. Barbara Weinrich. Her constant guidance, encouragement, and enthusiasm enabled me to complete this project. Additionally, her emotional support and editorial skills were immensely important for the completion of this thesis. I thank Dr. Doris Bergen and Ann Glaser for their suggestions, which enhanced the final written document. Without their guidance this project would not have been completed.

I would like to express my gratitude to Michael Hughes of the Statistical Consulting Center at Miami University. His statistical advice and services were vital to the completion of this project.

I would also like to acknowledge Nancy Gates, Barbara Weinrich, and the staff at Abilities First Foundation and Miami University Speech and Hearing Clinic for their cooperation in this matter. I greatly appreciate their assistance in recruiting and providing services to subjects for this project.

Finally, I am sincerely grateful for the encouragement provided by my family and friends throughout the past year. This thesis could not have been completed without their constant support.

CHAPTER I

Introduction

Language is a pertinent component of communication and typically utilizes words as a method of communication (Owens, 1996). Children are exposed to verbal and nonverbal language in their first few minutes of life and the exposure is continued throughout the rest their life. Bell and Ainsworth (1972) found that the degree of caregiver responsiveness to a child's verbalizations appears to be positively correlated with later language abilities. Other researchers have found within the first few months of life, infants are able to discriminate contrasting phonemes, different intonational patterns, and speech from nonspeech (Eimas, 1974; Hirschman & Katkin, 1974; Kearsley, 1973; Moffitt, 1971; Owens, 1996; Turnure, 1971).

Language Development

Children grow and continually learn language. As a newborn much of the communication is crying or speech-like sounds, usually when feeding (Owens, 1996; Paul, 1995). A child's ability to communicate his/her needs starts as early as birth. A child, between the ages of birth to 6 months, communicates with cries for assistance, coos single syllables, babbles strings of consonants, vocalizes to toys, and varies his/her voice in volume, pitch, and rate. Between 7 and 12 months, the child vocalizes more, imitates inflections, and eventually says his/her

first word(s). At 24 months a typical child has 200-300 words in his/her vocabulary. By 5 years, a child's vocabulary should include 2100-2200 words.

Unfortunately, not all children follow these stages of language development. In some cases, a child may just be lagging behind the other children his/her age and will eventually have age appropriate skills in a few months. However this is not always the case. There are some children who will continue to have difficulties using age appropriate language during their lifetime. Etiologies that affect children's lack of language development include: mental retardation, sensory deficits, environmental components, psychiatric disorders, and acquired disorders of communication function (Paul, 1995).

Language Intervention

Children, who are delayed in their language development, may need assistance before they can communicate and use language effectively. One possibility for these language-delayed children is enrollment in an early intervention program. This program may enhance the children's language development. Olswang and Bain (1991) included three major purposes of intervention. These were (a) to change or eliminate the underlying problem, (b) to make the child more normal on discrete aspects of language function by teaching specific behaviors, and (c) to teach compensatory strategies.

Intervention Studies

Intervention studies have shown positive results for children enrolled in early intervention programs. Dale, Crain-Thoreson, Notari-Syverson, and Cole (1996) found that children, ages 3 to 6 years, with language delays, increased their ability to produce what/who questions, open-ended questions, and expansions with intervention focusing on book reading. Pamplona and Ysunza (2000) found that children, ages 3 to 5 years, with cleft palate and accompanying language delay, made gains in both play and language development through participation in play therapy over a 12-month period. Logan and Caruso (1997) found that children, who stutter and are treated with intervention, have substantial and long-term improvements.

Assessment Tools

Children, thought to have a language delay, are assessed with norm-referenced tests and/or criterion-referenced tests and observed in interactions with others. Any of these assessment tools may help identify a child demonstrating a language delay. For the purposes of this study, a norm-referenced and criterion-referenced test will be analyzed. Norm-referenced tests have specific properties that allow a meaningful comparison of performance among children. These properties include clear administration and scoring criteria; validity, reliability, standardization, central tendency, standard error of measurement, and variability measures; and norm-referenced scores.

Paul (1995) proposed that criterion-referenced tests are procedures devised to examine a particular form of communicative behavior. Criterion-referenced tests do not reference to other children's achievement but only determine whether the child can attain a certain level of performance.

Once the child has been identified as having a significant language delay, the norm-referenced and criterion-referenced tests can be used to establish baseline function and identify targets for intervention. Another benefit to these two types of assessment is that the results will allow professionals to determine whether the goals have been achieved or mastered with intervention. Researchers (MacTurk & Neisworth, 1978; Oescher, Kirby, & Paradise, 1992; Summers, Larson, Miguel, & Terrell, 1996) have rarely measured the difference in the performance changes for norm-referenced versus criterion-referenced assessments made by language-delayed children following intervention.

Statement of Problem

Research (Fey, 1986; Paul, 1995) has identified that early intervention does produce language gains for language-delayed children. However, there are a limited number of research studies that focus on the differences between pre- versus post-intervention test scores with language-delayed children (Bradshaw, Hoffman, & Norris, 1998; Dale, Crain-Thoreson, Notari-Syverson, & Cole, 1996; Girolametto, Pearce, & Weitzman, 1996). Few studies compare the test scores of a norm-referenced test and a criterion-referenced test. Present studies fail to

identify if there are significant changes for norm-referenced and criterion-referenced test scores following intervention.

Purpose of the Study

The current research focused on the differences between a norm-referenced and criterion-referenced test. A study focusing on the differences between pre- versus post-intervention test scores discerned whether there were any differences in measurements of language skills. Additionally, the testing determined whether there was a significant difference for a norm-referenced versus a criterion-referenced assessment instrument. Knowledge of the differences allow professionals to better understand how these two assessment tools perform. Results enable professionals to weigh the strengths and weaknesses of each assessment tool and determine which would be more appropriate for their clients.

Significance of the Problem

This research study provided information regarding any differences in measurements of language skills following a designated intervention period. It determined whether there was a significant difference between the pre- versus post-intervention test scores of a norm-referenced versus a criterion-referenced assessment instrument. The data acquired from this study could aid in choosing the appropriate assessment tools for children with language delay.

Chapter Summary

In this chapter the term language was defined and the stages of language development were discussed. The role of early intervention and an overview of assessment tools were provided. The rationale for pre-versus post-intervention test scores for a norm-referenced and a criterion-referenced instrument for a language-delayed child was also introduced. The purposes of this study were identified as (a) to discern any significant differences in assessment measurements of language skills, for children with language delay, following intervention, and (b) to determine any significant differences for the pre- versus post-intervention test scores of a norm-referenced versus a criterion-referenced assessment instrument.

CHAPTER II

Review of Literature

The terms communication, speech, and language are used interchangeably, when actually each term is unique. Owens (1996) defined each term in a simple manner. Speech is defined as a verbal means of communication used to convey meaning and is the result of planning and executing specific motor sequences. “Language is defined as a socially shared code or conventional system for representing concepts through the use of symbols and rule-governed combinations of these symbols” (Owens, 1996, p. 8). Thirdly, communication is the combination of both speech and language. Communication is a process that participants use to exchange information and ideas, needs, and desires. For the purpose of this study, the emphasis will be on language and child language development.

Language

A more thorough definition of language comes from the American Speech- Language-Hearing Association (ASHA). ASHA (1983) defined language as

a complex and dynamic system of conventional symbols that are used in various modes for thought and communication. Modes for thought and communication are (a) language evolves within specific historical, social, and cultural contexts; (b) language is described by at least five parameters-

phonological, morphologic, syntactic, semantic, and pragmatic; (c) language learning and use are determined by the intervention of biological, cognitive, and psychological factors; and (d) effective use of language for communication requires a broad understanding of human interaction including such associated factors as nonverbal cues, motivation, and sociocultural roles (p. 949).

Language is a social interaction tool and has both rule-governed and generative properties (Owens, 1996; Paul, 1995; Weitzman, 1992). A rule-governed system includes grammars. Grammars are identified as the set of underlying operation principles or rules that describe the relationships between symbols that form the structure of language. These grammars allow users of language to comprehend and to create messages. Children learn the rules by actually using the language to encode and decode messages. Secondly, the generative system means to produce or bring in existence. With help from the linguistic rules, speakers are able to generate, or form meaningful utterances.

Language is comprised of several components: form, content, and use. “Form includes syntax, morphology, and phonology, the components that connect sounds or symbols with meaning. Content encompasses meaning or semantics, and use is pragmatics” (Owens, 1996, p. 17). Syntax is the form or structure of a sentence. Morphology concerns the internal organization of words, which is the morpheme. The morpheme is the smallest grammatical unit and is invisible

without violating the meaning. Phonology describes the structure, distribution, and sequencing of speech sounds and the shape of syllables. Semantics concerns the meaning or context of words and word combinations. Finally, pragmatics is concerned with the way language is used to communicate rather than the way language is structured.

Language Development

When children are born, they are not capable of communicating as a grown adult. Infants' communicative development is related to other developmental areas, such as attaining certain cognitive, social, and motor skills (Owens, 1996; Paul, 1995). An infant's maturation of certain neuromuscular structures and motor control also need to be present for the infant to communicate. The typical infant communicates through crying for the first 2 months, and then eventually starts to coo at about 3 months. Within the first 3 months of life, the infant is able to distinguish different sounds, turn the head when hears a voice, respond vocally to speech of others, and make predominantly vowel sounds. Between 4 to 6 months, the infant begins to babble strings of consonants, vary pitch, imitate tones, vocalize to toys, respond to name, and vocalize pleasure and displeasure (Paul, 1995). Typical developing children between the ages of 7 to 12 months begin to play vocally, recognize some words, use jargon, imitate adult speech, and speak one or more words.

After children say their first word, there is an explosion in their language. It is reported by Owens (1996) that a typical child has a vocabulary of 200-300 words at 24 months. Other characteristics of a child's language at the age of 24 months include use of short incomplete sentences, some prepositions, and some regular verb endings. At the age of 3 years, most children are eligible to begin preschool. Preschool is the time when children expand their language and learn the rules of language. Children begin to exhibit their abilities to form simple sentences, play with words and sounds, recount stories and the recent past, and follow three-step commands. Owens (1996) and Shipley and McAfee (1998) reported that a typical language-developing child would have up to 2,200 expressive and 20,000 receptive vocabulary words at the age of 5 years. After the age of 5 years, children start to form longer sentences, have good comprehension, and produce adult-like speech mannerisms.

Language Delay

According to ASHA (1982) a language disorder is the impairment or deviant development of comprehension and/or use of a spoken, written, and/or other symbol system. The disorder may involve the form, content, and/or use of language. The terms language "delay" and language "deviant" are frequently interchanged. However, there is a major discrepancy between them. Paul (1995) defined language "delay" as a child's language that is similar to a younger child and will eventually arrive at the desired destination of normal development. A

child who is 24 months old and has a vocabulary of five words with normal comprehension and cognition would be considered to have a language “delay”. This child is labeled with a language “delay” because he/she is expected to grasp normal development. Language “deviant” is defined by Paul (1995) as a child’s language development that is not just slower than normal but different in a qualitative way. An example of language “deviant” would be a child who demonstrates patterns of speech unlike those seen in younger normal children.

Assessment Tools

“The ultimate goal for the assessment process is to decide whether the child has a significant deficit in communication, and to describe that deficit, if identified, in as much detail as possible, relative to the normal sequence of language acquisition” (Paul, 1995, p. 20). There are three main procedures used to assess a young child’s language: norm-referenced tests, criterion-referenced tests, and observation. For the purpose of this paper, norm-referenced and criterion-referenced tests will be discussed. Law mandates the need for standardized measures when identifying children with language impairments (Merrell & Plante, 1997). Current practices suggest the need for data based analyses. Analyses will provide evidence to interpret test results and answer specific diagnostic questions.

Norm-Referenced Tests

According to Paul (1995) norm-referenced tests are formal assessments and have specific properties that allow a meaningful comparison of performance among children. These properties include clear administration and scoring criteria; validity, reliability, standardization, central tendency, standard error of measurement, and variability measures; and norm-referenced scores. Mills and Hambleton (1980) stated norm-referenced assessments are constructed to facilitate comparisons among individuals in relation to the performance of the normative group. A standardized test is also used when comparing a child to the norm. Standardization is defined as the process of administering a test under uniform conditions to each child who is tested (Montgomery & Connolly, 1987).

There are advantages and disadvantages for using norm-referenced tests. Norm-referenced tests will provide evidence regarding the existence of a problem, suggest a need for further assessment, and/or help document a need for the initiation or continuation of therapy (McCauley & Swisher, 1984). Montgomery and Connolly (1987) reported that norm-referenced tests were designed to delineate differences among individuals and used for diagnostic and placement purposes. Johnson and Martin (1980) concluded that norm-referenced tests spread out individuals along a continuum of performance in order to detect deviations from the average.

McCauley and Swisher (1984) noted disadvantages to norm-referenced tests if misused. A misused norm-referenced test can lead to (a) a mistaken understanding of an individual's problem, (b) an inappropriate and fruitless therapy program, and (c) an inaccurate conclusion regarding efficacy of therapy. Another disadvantage for norm-referenced tests is that the comparison of a test taker's score to the relative norms involves a comparison of estimated, rather than absolute, or true values.

Besides the disadvantages mentioned above, McCauley and Swisher (1984) reported four specific problems in the use of norm-referenced tests. The first problem is using age-equivalent scores as test summaries. "This problem concerns the relation of age-equivalent scores and the raw scores on which they are based" (Saliva & Ysseldyke, 1981, p. 67). With most norm-referenced tests, similar differences in age-equivalent scores are the result of smaller and smaller differences in raw scores (McCauley & Swisher, 1984). This problem is not necessarily based directly on evidence collected for children at that chronological age and can serve as a basis of misinterpretation. A second problem is the profile analysis. McCauley and Swisher (1984) stated the scores to be compared in a profile, on norm-referenced tests, are only estimates of the ideal or true scores one would obtain if the scores were free from measurement error. Performance on individual test items as indications of deficit is the third problem. That is, the small number of items on a norm-referenced test cannot adequately sample all of

the specific forms and developmental levels that might be appropriate. The fourth problem with using norm-referenced tests is the repeated testing as a means of assessing progress. The result is underestimation or overestimation of change, since the individuals are able to learn the items on the test. These problems demonstrate that norm-referenced tests provide incomplete and possibly misleading information for the formulation of language objectives and language analyses.

Criterion-Referenced Tests

Paul (1995) proposed that criterion-referenced tests are procedures devised to examine a particular form of communicative behavior. Criterion-referenced tests do not reference to other children's achievement but only determine if the child can attain a certain level of performance. Montgomery and Connolly (1987) stated that criterion-referenced tests document individual performance in relation to a domain of information or specific set of skills. Therefore, criterion-reference tests are designed to measure changes in successive performance in an individual. Criterion-referenced tests are used specifically for program planning and evaluating; however, they can also be standardized.

Much like the norm-referenced tests, criterion-referenced tests have their own advantages and disadvantages. One advantage for the criterion-referenced tests is their scoring procedures. This type of test is based on absolute rather than relative standards. Its primary use is to measure mastery of specific skills and test

items, based on known performance objectives associated with the tasks of interest. Criterion-referenced tests are sensitive to and can be used to measure the effects of instruction, based on task analysis, related directly to instructional objectives. Sensitivity is defined as the accuracy with which the test identifies children with language impairment as language impaired (Merrell & Plante, 1997). The ability to tie the test directly to the program objectives is another benefit of criterion-referenced tests. Freeman and Miller (2001) reported that criterion-referenced tests were consistently rated as the most useful assessment tool, both for understanding the child's abilities and needs, and for planning teaching responses to them. This assessment tool refers directly to the curriculum, and is likely to be considered comprehensible and relevant.

Although there are a number of advantages for criterion-referenced tests, there are a few disadvantages that need to be mentioned. One disadvantage includes the inability to assign age levels if not normed or administered in a standardized manner. MacTurk and Neisworth (1978) stated another disadvantage for the criterion-referenced tests is the lack of comparative interpretability.

Similarities Between Norm-Referenced and Criterion-Referenced Tests

Even though norm-referenced and criterion-referenced tests have many differences, there are a few similarities. For example, criterion-referenced and norm-referenced tests should demonstrate the same interrater and test-retest reliability (Montgomery & Connolly, 1987). Issues of validity, such as content,

concurrence, and predictive value, should also be similar between the two tests when administered.

Differences Between Norm-Referenced and Criterion-Referenced Tests

McCauley (1996) summarized the differences between norm-referenced and criterion-referenced tests in a simplistic way. The first difference is the fundamental purpose of both tests. The fundamental purpose of norm-referenced tests is to rank individuals, whereas the fundamental purpose of criterion-referenced tests is to distinguish specific levels of performance. A second difference is the test planning. Norm-referenced tests address a broad content and criterion-referenced tests address a clearly specific domain. Lastly, a third difference is how the individual's performance is summarized. With norm-referenced tests, performance is summarized meaningfully by using percentile ranks and standard scores; and criterion-referenced test performance is summarized meaningfully by using raw scores.

Studies Comparing Norm-Referenced and Criterion-Referenced Tests

Handicapped versus nonhandicapped. MacTurk and Neisworth (1978) studied the degree of correspondence between a presumably appropriate norm-referenced device and a project specific, criterion-referenced measure of child progress, as well as the utility of the two different measures in program planning, delivery, and modification. The 17 subjects in their study were divided into two groups, handicapped (7) and nonhandicapped (10). Handicapped was defined as

displaying significant but not profound dysfunctions in at least two of four basic developmental areas: communication, intellectual, social, and motor. Subjects ranged in age from 19 to 51 months. Inclusion criteria for these children were (a) be ambulatory, (b) be nonhandicapped sensorially, and (c) manifest no extreme disorders requiring close clinical or medical supervision.

Both groups of children were administered a norm-referenced and a criterion-referenced test on a quarterly basis. Results of the norm-referenced test were converted into a developmental quotient, a developmental rate, and a mean developmental quotient. Additionally, the criterion-referenced test's curriculum objectives were converted to a percentage of total objectives achieved.

Through the distribution of scores, the researchers found no significant relationships ($r = .08$ to $.43$, $p < .01$) between the total sample of subjects and nonhandicapped group with a Spearman Rank Correlation Coefficient. However, the handicapped group demonstrated a high correlation ($r = .93$ to $.99$, $p < .01$) between the developmental quotient and the criterion measures. The Wilcoxon Matched-Pairs Signed-Ranks Test was performed to test the feasibility of using the mean developmental quotients as a measure of child progress. This analysis indicated there were no significant differences between the pre- and post-developmental quotient for both nonhandicapped ($T = 25.5$, $p > .05$) and handicapped children ($T = 14.0$, $p > .05$).

Overall, the findings in this study demonstrated a high correlation (.93 to .99, $p < .01$) between the norm-referenced and criterion-referenced tests for the handicapped children. The norm-referenced test was highly correlated (.90 to .96, $p < .01$) with classroom measures and had utility as a diagnostic device for educational programming. This high correlation allows the norm-referenced assessment results to be a useful measure of a child's progress once in an educational program, and the concern of insensitivity is not warranted. These findings supported the following conclusions: (a) norm-referenced developmental quotient and criterion-referenced measures yielded approximately similar information for handicapped children, (b) classroom activities showed differential effects for each sample, and (c) mainstreaming did not have any apparent adverse effects on the normal children in the sample.

Test-retest comparison of screening measures. Summers, Larson, Miguel, and Terrell (1996) studied the performance stability of kindergarten children on two commonly used screening tests. The researchers hypothesized that the Clinical Evaluation of Language Fundamentals-Revised Screening Test (CELF-RST) by Semel, Wiig, and Secord (1989) and the Bankson Language Test-Second Edition, Screening (BLT-2S) by Bankson (1990) would identify the same children as failing and a child's test performance would be constant over time. Both tests measure a child's ability to apply his/her language competence on tasks that may be novel to the kindergarten child (Summers et al., 1996). Therefore, children

who were identified as failing the first screening would not demonstrate significantly better performance, by benefit of the kindergarten experience alone.

Both tests were administered to 101 kindergartners at the beginning of the school year and again 7 months later. The inclusion criteria were (a) intelligible speech, (b) passed a pure-tone and immittance screening test, (c) between 60 and 71 months of age, and (d) English as the primary language in the home. During the process of this investigation, no other services were provided to the kindergartners.

Results showed that the post-test scores on both screening measures reflected higher mean scores, standard deviations, and maximum scores than the pre-test scores. Pre-test results for the CELF-RST were mean score 11.81, standard deviation 5.60, and maximum score 23; and the BLT-2S were mean score 10.08, standard deviation 3.50, and maximum score 17. Post-test results for the CELF-RST were mean score 17.50, standard deviation 6.43, and maximum score 33; and the BLT-2S were mean score 12.69, standard deviation 3.33, and maximum score 20. A two-tailed t-test concluded there was a significant difference between mean scores for pre- and post-testing (BLT-2S, $t = -8.07$, $p < .001$; CELF-RST, $t = -12.94$, $p < .001$). Comparing pre- and post-test scores, researchers found approximately 71% of the children were accurately identified as passing or failing with the CELF-RST and 70% with the BLT-2S. This led researchers to believe that pre- and post-test differences seemed to reflect several

causal factors, such as maturation, skills assessed, time of administration, and psychometric variability.

Researchers believed skills assessed could be a causal factor because the CELF-RST and BLT-2S have different subtests. For example, the CELF-RST evaluated the child's ability to recall auditory information of varying length and language complexity, whereas the BLT-2S did not evaluate this language skill. Time of administration was also a concern to the researchers. Researchers thought with exposure to the formal school experience and additional time, many children initially failing the screening may have been able to perform at age appropriate levels later in the school year. A third causal factor of concern for researchers was psychometric variability. According to Walsh and Betz (1990) an error rate of 1% to 2% is to be expected with test construction, pre/post test correlation. However, Summers et al. (1996) had a 9% error rate with test construction, which exceeds the error rate reported by Walsh and Betz. With these causal factors, researchers suggested that kindergarten screening be postponed until later in the year to reduce false positives.

Comparing norm-referenced and criterion-referenced tests in school aged children. Oescher, Kirby, and Paradise (1992) provided additional post validation data relevant to practicality issues for a criterion-referenced statewide-achievement test (Louisiana Educational Assessment, LEAP) by using a widely known and respected norm-referenced test (California Achievement Test, CAT)

as a benchmark against which to evaluate the adequacy of the LEAP performance standards. LEAP was established to ensure the development and continuous review of grade-level standards in reading, writing, and mathematics for all public school grade levels. “Knowledge of scores on the norm-referenced test (CAT) and the criterion-referenced test (LEAP) would contribute to a meaningful interpretation of how acceptable performance, as defined by the state criterion level, relates to national performance levels” (Oescher, Kirby, & Paradise, 1992, p. 142).

Students in the third and tenth grades were selected to participate in this investigation (n=30,000). Each student was administered a norm-referenced test and a criterion-referenced test. Specific subscales (reading, language, and mathematics) of the norm-referenced test were administered and included in the total test scores.

Pearson product-moment correlation was used to express the results. Correlations between CAT and LEAP ranged from .70 to .80, identifying a high correlation and marked relationship between corresponding subscales on both tests. Discriminant function analysis was used to determine the extent to which CAT produced similar decisions about the students as LEAP. Comparing LEAP to CAT with the third graders, 13% of the students were mistakenly identified as failing the language subtest and 17% for the mathematics subtest on the CAT. This comparison of tests also identified 16% of tenth graders as failing the

language subtest and 15% failing the mathematics subtest on the CAT.

Altogether, students that were identified as failing the language and/or mathematics subtests on the CAT passed these subtests on the LEAP.

These results demonstrated that CAT and LEAP would not lead to similar decisions for all students. Researchers found the false positives and false negatives to be inconsistent. Even though differences were seen in this investigation, there were some moderately high correlations ($r = .70$ to $.80$) between similar content subtests on the CAT and LEAP.

Criterion-referenced test with projected norms. Behuniak and Tucker (1992) studied a different measurement approach to meet large-scale assessment and evaluation requirements. Their argument was for the development of a psychometric link between a criterion-referenced test and a norm-referenced test to eliminate the need to administer the norm-referenced test.

Researchers conducted the study in Connecticut, where the criterion-referenced test being used was developed. Participants' grade levels included 4th, 6th, and 8th. Academic areas assessed with both the criterion-referenced and norm-referenced tests were mathematics, reading, and language arts. Participants were asked to complete the entire criterion-referenced test and specific subtests on the norm-referenced test. Specific subtests were chosen from the norm-referenced test in order to match content between tests.

Results indicated that individual differences were similar when a criterion-referenced test was substituted for the norm-referenced subtests. An advantage to creating a link between the criterion-referenced and norm-referenced test is that the link can promote a focused, coherent assessment and evaluation system. Two other advantages were reported for the link: (a) allows the identification of standards that all students can actually attain and supports the implementation of an evaluation system consistent with that goal, and (b) encourages the clear articulation of the academic skills that students know or do not know and establishes this as a higher priority than a few points of measurement error in ranking students and their peers. However, researchers stated that precautions need to be taken when creating a criterion-referenced test and a norm-referenced test link. Limitations include: content coverage, method of content analysis, and error in individual scores.

Intervention

A child, who is thought to have a language delay, should be assessed with standard protocols to identify if a significant language delay is present. If a language delay is found, a child may benefit from intervention. Venes (2001) defined intervention as one or more actions taken in order to modify an effect. Olswang and Bain (1991) proposed three main purposes of intervention, which are to change or eliminate the underlying problem, to change the disorder, and to teach compensatory strategies.

Three Purposes of Intervention

Olswang and Bain (1991) stated that the first purpose of intervention is to change or eliminate the underlying problem. The intent is to render the child as a normal language learner, who will no longer need further intervention. However, this is not always possible. Unfortunately the underlying problem may be unclear, or incapable of being alleviated. More language-delayed children, with unknown etiologies or associated incurable conditions, are being seen for intervention. This uncertainty may impede progress and force interventionists to accept the child as a less-than-normal language learner.

The second purpose of intervention is to make the child more normal on discrete aspects of language function by teaching specific behaviors (Olswang & Bain, 1991). For example, a speech-language pathologist could teach the child to expand the number of words in sentences or to use language more flexibly and appropriately. This does not guarantee that the child will not need further help later on, but does make the child a better communicator. This purpose is one that is frequently used with language-delayed children.

Thirdly, Olswang and Bain (1991) stated intervention's purpose is to teach children compensatory strategies. In this approach, the speech-language pathologist is attempting to give the child tools to function better with the deficits he/she has. Compensatory strategies require a good deal of cognitive maturity.

Therefore, this approach is generally used to help older school-aged and adolescent children.

Effectiveness of Intervention

Shared reading intervention. Many researchers found that intervention is effective with language-delayed children (Ezell, Justice, & Parsons, 2000; Girolametto, Pearce, & Weitzman, 1996; McConkey, Jeffree, & Hewson, 1975; Ward, 1999). Ezell, Justice, and Parsons (2000) asked to what extent do children's emergent literacy skills improve after participation in a 5-week group book-reading intervention?

Participants in this study ranged in age from 2 years 8 months to 4 years to 2 months. Inclusion criteria required participants to be pre-school age and receiving speech-language therapy. Prior to beginning the book-reading program, participants were assessed with a norm-referenced test and informal assessment. A book-reading program was implemented for 3 weeks and each session lasted 1 hour. The fifth and final week, the participants were re-administered the norm-referenced test and informal assessments.

Researchers found that 3 of the 4 participants made notable gains in their acquisition of concepts about print and book reading. Therefore, children with communication disorders may increase their early literacy skills following shared reading intervention.

Interactive focused stimulation intervention. Girolametto, Pearce, and Weitzman (1996) hypothesized that enhanced or optimized levels of parental language input will provide increased opportunities for children with language delays to learn. A second hypothesis stated that children would increase their rate of utterances, use more words, and produce more advanced vocabulary and language than children in the control group.

Twenty-five participants were used in this study. Participants for this study ranged in age from 23 to 33 months. Inclusion criteria were: (a) performs at the single-word stage of language development, (b) has no major sensory impairments, (c) has no oral motor problems, (d) speaks English as the primary language, and (e) has no frank neurological problems, such as autism. Participants and their mothers were observed in two 15-minute, free play periods, followed by a criterion-referenced test. Participants' mothers were asked to complete a checklist about their child's behavior and language at home. The experimental group received intervention instructions, whereas the control group did not receive any intervention instructions. Eleven weeks of intervention were conducted by the participant's mother and took place at the participant's home. Researchers made home visits to provide feedback and strategies to the mothers. Following intervention, all mother-child dyads were re-assessed. Post-testing was performed in the same manner as the pre-testing.

Researchers found that mothers who received intervention instructions reduced their mean length of utterance and rate of words per minute. The participants who received intervention, made developmental gains in vocabulary, in the use of multiword phrases, and in grammatical complexity that were over and above the maturational changes made by the control group. Girolametto et al. (1996) also found the experimental group made significant gains in the structural complexity of their language.

Mentally handicapped children with intervention. McConkey, Jeffree, and Hewson (1975) hypothesized that mentally handicapped children, with language delay, would improve their language skills with their parents solely responsible for performing the language facilitation activities. A speech-language pathologist would introduce the parents to techniques appropriate for their child's needs.

Ten families participated in this study. Eight children were labeled with Down syndrome and the other two children's handicap was unknown. All children were administered a norm-referenced test prior to treatment. The children ranged in age from 40 to 69 months. Baseline was collected through developmental charts, a play session, and mental developmental scale. A speech-language pathologist selected the learning objectives to be targeted. Parents were taught strategies to use when performing intervention at home. The speech-language pathologist emphasized to the parents that the child would only benefit from regular intervention sessions. Parent-child sessions were audio taped at the

home and video taped when at the clinic. The home based intervention lasted for 5 months. When the intervention was complete, the children were re-administered the norm-referenced test.

Researchers found that family members, who were guided by speech-language pathologists, extended young mentally handicapped children's language development. In this particular study, all children showed progress. However, some children progressed more than others. It is worth mentioning that there were reasons to suspect that some children were not receiving enough intervention opportunities. In the single word category, four of the ten children made considerable progress with 10 to 14 weeks of intervention. The four children all began intervention with 0 to 10 words, and after intervention the children were producing more than 300 words. Two-word sentences also improved with these children. At the end of the intervention period, all the children were producing two-word sentences during the teaching games.

Early intervention method effectiveness. Ward (1999) conducted a study on the effectiveness of an early intervention method for delayed language development in young children. It was hypothesized that a group of language-delayed infants, detected by the researcher's screening in the first year of life, would demonstrate no impairment after 12 months of intervention. However, an untreated control group, receiving no intervention, would be referred for speech and language therapy at a later age.

Participants in the study were under the age of one year and were followed over the next two years of life. Infants who failed the researcher's linguistic screening test and showed signs of a developmental language delay were divided into two groups, experimental and control. Each participant was evaluated with a criterion-referenced test assessing both expressive and receptive language. The experimental group received intervention, while the control group did not. Intervention time period ranged from 2 to 8 months. One hundred and forty minutes was the average length of time for intervention in this study. Both groups were reassessed one and two years post-initial assessment.

At the one-year, post-intervention assessment, Ward (1999) found the experimental group to be linguistically within normal limits using the criterion-referenced test. A second follow-up assessment, at two years, continued to show the same results with expressive and receptive language skills; experimental group performed linguistically within normal limits. Before the study was complete, 30% of the control group was referred for speech and language therapy, while none of the experimental group were referred. Researchers' noted 24% of the control group and 5% of the experimental group were admitted into authority day nurseries after completion of the study.

Types of Intervention

There are many different types of intervention used with language-delayed children. Each intervention technique has advantages and disadvantages for a

practicing speech-language pathologist or researcher. It is important for practicing speech-language pathologists to determine the type of intervention they think will benefit the child the most. The correct intervention type will enable the speech-language pathologist to collect a high number of responses, and the child and clinician will establish good rapport, that is, the child will enjoy therapy. However, if the appropriate intervention technique is not provided, the clinician may experience some resistance from the child, response rate may be low, or the parents may feel frustrated with the slow progress the child is demonstrating. The therapy techniques may be trial and error before the right intervention type is found. However, it is in the best interest of the speech-language pathologist and child to determine the most appropriate and successful intervention type. For the purposes of this study, the following intervention techniques will be briefly discussed: (a) play-based, (b) hybrid approach (focused stimulation), (c) clinician-directed (drill and modeling), and (d) child-centered (self-talk, parallel talk, imitation, expansions, extensions, and buildups and breakdowns).

Play-based intervention. Paul (1995) described play-based intervention as a clinician arranging an activity so opportunities arise naturally for the child to provide target responses. The clinician spends his/her time engaging the child in a natural and enjoyable activity. With this approach, the clinician waits for the child to express wants, needs, or an idea. After the child uses a communication function, the clinician is able to model that same function with words.

Clinician-directed intervention. Fey (1986) referred to clinician-directed intervention as an approach that attempts to (a) make the relevant linguistic stimuli highly salient, (b) reduce or eliminate irrelevant stimuli, (c) provide clear reinforcement to increase the frequency of desired language behaviors, and (d) control the clinical environment so that intervention is optimally efficient in changing language behaviors. There are different techniques utilized with clinician-directed intervention. First, *drill* is defined as the clinician instructing the child in regards to the expected response and providing a training stimulus, such as a word or phrase to be repeated. Another technique is *modeling* and is described as the child listening to the training stimulus and then producing the target structure.

Child-centered intervention. With the child-centered approach the clinician is able to follow the child's lead and react to the child's behavior, placing the message in a communication context and giving the message linguistic meaning. There are many different techniques that can be used with the child-centered approach. First, *self-talk* is defined as the clinician describing his/her own actions as the clinician and child engage in parallel play. Another technique is *parallel talk* and is described as the clinician talking about the child's action, providing a running commentary. *Imitation* occurs when the clinician imitates exactly what the child says. *Expansions* and *extensions* are two similar techniques in that the clinician takes what the child says and adds grammatical

markers and semantic details that would make the utterance an acceptable adult utterance. Finally, *buildups* and *breakdowns* occur when a child's utterance is first expanded to a full grammatical form and then broken down into several phrase-sized pieces that overlap the content.

Hybrid approach intervention. The hybrid approach targets one or a small group of specific language goals (Paul, 1995). In this approach the clinician maintains a good deal of control in selecting activities and materials. The clinician also models and highlights the forms of language being taught to the child. One form of the hybrid approach is focused-stimulation. In focused-stimulation the clinician carefully arranges the context of interaction so that the child is tempted to produce utterances with obligatory contexts in the forms of language being targeted. This technique has specifically been effective with improving comprehension, as well as production of a language form (Fey, 1986).

Studies Showing the Effectiveness of the Intervention Techniques

Focused-stimulation in children with language impairment. Fey and Cleave (1993) conducted a study using 30 children with marked delays in the development of grammar, over an 18-week time period. Three groups were formed with ten children in each. The three groups were labeled as the clinician treatment group, the parent treatment group, and the delayed-treatment group. Focused-stimulation was employed by both the clinicians and parents.

All the children were pre- and post-tested with speech/language samples. Procedures for focused-stimulation included: (a) an activity was created so that the child would have many opportunities to hear and attempt productions of the specific targets, (b) the target form was modeled by the clinician, (c) contexts were manipulated so that the target could be modeled in highly salient contexts, (d) requests for elaboration were used frequently, and (e) techniques were employed in stories made for the child.

Both treated groups exhibited significant gains in grammatical expressions. In contrast, the delayed-treatment group, who were withheld from treatment, made no gains in grammatical expression. The benefits of focused-stimulation treatment were especially noteworthy in the post-speech/language sampling. Researchers found that the clinician treatment group performed significantly better than the delayed-treatment group ($F= 4.70$, $p= .04$) on the main verb score per sentence. However, the two treatment groups did not differ in this area.

Efficacy of expansions and cloze procedures. Bradshaw, Hoffman, and Norris (1998) studied two 4-year-old boys who were identified as language delayed after administering the *Clinical Evaluation of Language Fundamentals-Preschool* (Secord, Wiig, & Semel, 1992). Participants received two types of storybook reading intervention. These interventions were (a) cloze and expansion procedures, and (b) question-model answer conditions. The intervention

administered was randomly selected and evenly distributed throughout the study. Each intervention was performed for a 15-minute time period during 12 sessions. Intervention was held three times a week.

The cloze and expansion intervention goal was to enhance the children's production of story interpretations by expanding labels into descriptions and descriptions into interpretations. A cloze example included the clinician stating, "The cow is taking a bath because....", requiring the child to fill-in the answer. An expansion example was as follows: pointing to a character on a page and requiring a label, enabled the clinician to expand the child's utterance into a phrase or sentence. Following the expansion, the clinician proceeded to provide the child with questions that required interpretation.

The question-model answer conditions provided the child with an equal number of opportunities to label, describe, and interpret information. An example of this situation was the clinician asking, "What is this?" while pointing, or "Which rose is the biggest?" If the child did not answer, the clinician would model the answer in a sentence. This intervention approach differed from cloze and expansion intervention in that the question-model answer condition produced zero opportunities for cloze and expansion procedures.

Researchers found both children produced more responses when cloze and expansion procedures were used. Child A averaged 46 utterances with cloze and expansion procedures and 37.3 during the question-model answer conditions,

while child B averaged 44 utterances with cloze and expansion procedures and 34.2 during the question-model answer conditions. The frequencies of interpretations were also compared between the two intervention types. Child A produced more interpretations in the cloze and expansion procedures (mean=19.7) versus the question-model answer conditions (mean=12). Child B demonstrated similar results with the cloze and expansion procedures (mean=15.3) versus the question-model answer conditions (mean=8). Finally, researchers discussed the complexity of syntactic structures and morphological forms (i.e. multiple verbs, clauses containing both a subject noun phrase and predicate verb phrase, conjunctions, past tense verbs, and modal verbs). Both children demonstrated an increase in the complexity of their responses with both interventions. However, the children produced more complex utterances with cloze and expansion procedures, with subject noun phrase and predicate verb phrase and past tense verbs increasing the most over the 12 sessions.

Effects of treatment on linguistic and social skills in toddlers. Robertson and Weismer (1999) studied 21 late-talking toddlers, ranging in age from 21 to 30 months. Two groups were formed with the toddlers, 11 in the experimental group and 10 in the control group. All toddlers were pre-tested and demonstrated delays in the acquisition of language, yet normal development in other areas. Toddlers in the control group received no intervention, while those in the experimental group received intervention for a 12-week time period. Procedures included child-

centered intervention, focusing on stimulation of vocabulary development and use of 2- or 3-word combinations in social contexts. Both the experimental and control groups of children were post-tested following the 12-week intervention period.

During intervention, the children developed their own routines. To increase the opportunities for linguistic input or to encourage the use of language, the clinician manipulated the child's routine. While the child was in his/her routine, the clinician was able to promote linguistic growth through, parallel talk and recasting. Parallel talk was defined as the clinician provides a verbal description of the child's actions. For example, "Hug the bear" or "Sweep the floor". Recasting was defined as a repetition of the child's utterance with modification of modality or voice (i.e., Child: "Want juice", Clinician: "You want some juice?").

Results demonstrated that the children receiving intervention produced a significantly higher mean length of utterance (1.32 vs. 1.09) than the children in the control group ($F=10.33$, $p=.003$). Total number of words ($F=46.83$, $p=.01$), number of different words ($F=41.05$, $p=.01$), lexical repertoire ($F=46.86$, $p=.01$), percentage of intelligible utterances ($F=24.44$, $p=.01$), and socialization ($F=12.15$, $p=.003$) all demonstrated significant differences between the experimental and control groups. These significant differences between groups

demonstrated to the researchers that intervention focused on routines and scripts does improve a late-talking toddlers language.

Chapter Summary

The studies reviewed provide useful information pertaining to norm-referenced and criterion-referenced tests. In some cases, attempts have been made to discriminate the similarities and differences of norm-referenced and criterion-referenced tests. However, none of these studies discerned any differences in norm-referenced and criterion-referenced tests following language intervention. For this reason, it is very difficult to apply the information provided by previous studies to clinical settings. In order to obtain current descriptions of differences in norm-referenced and criterion-referenced tests following language intervention, further research is necessary.

Research Hypotheses

1. There will be a significant difference between pre- versus post-intervention test scores on a norm-referenced test for language-delayed children following 5 months of intervention.
 - a. There will be a significant difference in the degree of improvement on the norm-referenced test between the groups of children and their pre-intervention disability levels.

2. There will be a significant difference between pre- versus post-intervention test scores on a criterion-referenced test for language-delayed children following 5 months of intervention.

CHAPTER III

Methods and Procedures

Subjects

The participants in this study consisted of 18 language-delayed children enrolled in speech-language therapy in Southwestern Ohio. Therapy was conducted at four different sites. Sites included two private practices, a university speech and hearing clinic, and an agency for developmental disabilities. Participants' chronological ages ranged from 25 months to 54 months at onset of the study. For inclusion in this study, each participant adhered to the following criteria: (a) severity of language impairment ranged from standard scores greater than 1 to 2.5 standard deviations below the mean in receptive, expressive, or total language skills, (b) received intervention for 5 months, and (c) spoke English as his/her primary language. All subjects were identified as having a language delay, and not a language deviance.

Confidentiality of Records

Participants were given a number to protect confidentiality of their records. Each participant was referred to by this number throughout the study. Data was stored in locked files. Only the faculty advisor and researcher had access to these files. Coded data was kept on computer disk, not on hard drives of community accessible computers.

Procedures

The parents of the potential subject, matching the inclusion criteria, were given a recruitment letter explaining the study's purpose and the expectations of the subject (Appendix A). Prior to any test administration the parents of the eligible children received a brief synopsis of the study and signed a consent form (Appendix B). Parents were assured of data confidentiality. To assess the effect of intervention on norm-referenced and criterion-referenced test scores, each language-delayed child was administered the Preschool Language Scale-3 (PLS-3) (Zimmerman, Steiner, & Pond, 1992) and the Receptive-Expressive Emergent Language Test (REEL-2) (Bzoch & League, 1991) before a 5-month intervention period. A licensed speech-language pathologist and a graduate researcher administered the tests. Evaluations allowed the participants to have predetermined long and short-term objectives for intervention. These objectives were continued throughout the study. Each participant's intervention program was determined by his/her speech-language pathologist. After 5 months of intervention, each participant was re-administered the PLS-3 and REEL-2.

Assessment Tools

The PLS-3 (Zimmerman et al., 1992) assesses the receptive and expressive language skills in infants and young children (2 weeks to 6 years, 11 months). The PLS-3 has two subscales that focus on different aspects of communication. The auditory comprehension subscale focuses on attention abilities and the expressive

communication subscale focuses on social communication and vocal development. Both subscales evaluate a child's knowledge of language content with tasks that focus on vocabulary, concepts, and language structures. Results determine whether deficiencies are primarily receptive or expressive, or if there is an overall delay/disorder in communication.

The REEL-2 uses observational information, with a checklist, to identify receptive and expressive language skills typically developed in infants and toddlers up to 3 years of age (Bzoch & League, 1991). Two primary uses of REEL-2 are to provide descriptions of the present developmental status of young children and to assist with setting intervention goals. The REEL-2 addresses three functional language systems. These functional language systems are receptive, expressive, and inner language. To assess the three functional language systems, the REEL-2 focuses on four contexts: phonemic, morphemic, syntactic, and semantic. Results from the REEL-2 are used to diagnose delays in language skill development and provide information to parents regarding the stages of learning to stimulate in the home.

Intervention

Subjects' clinicians were asked to document the type of intervention they used with their particular subject. Clinicians recorded (a) type of intervention, (b) subject's goals, and (c) length of session on a data form (Appendix C) following each therapy session. Several types of interventions were provided to the subjects

in this study. These interventions included: (a) play-based, (b) clinician-directed, (c) child-centered, and (d) hybrid approach. Play-based intervention relies on natural experiences to provide opportunities for target responses. The clinician-directed intervention attempts to reduce or eliminate irrelevant stimuli and provides clear reinforcement to increase the frequency of the desired language behavior. With client-directed intervention the clinician follows the child's lead in therapy and expands and corrects the child's message, creating linguistic meaning. The hybrid intervention approach allows the clinician to lead and provide models of correct language. The majority of the language intervention was 30-minute, individual sessions.

Experimental Design

A pre- versus post-intervention research design was used to evaluate if there were significant differences in (a) norm-referenced test scores, and (b) criterion-referenced test scores following 5 months of intervention. This research design was used to determine if the degree of improvement on the norm-referenced test was related to the initial pre-intervention disability level, as defined by the standard score scale.

Research Questions

1. Will language-delayed children demonstrate significant gains on pre-versus post-intervention norm-referenced test scores following 5 months of intervention?

a. Will there be a significant difference in the degree of improvement on the norm-referenced test between the groups of children and their pre-intervention disability levels.

2. Will language-delayed children demonstrate significant gains on pre-versus post-intervention criterion-referenced test scores following 5 months of intervention?

Null Hypotheses

1. There will be no significant difference between pre- versus post-intervention test scores on a norm-referenced test for language-delayed children following 5 months of intervention.

a. There will be no significant difference in the degree of improvement on the norm-referenced test between the groups of children and their pre-intervention disability levels.

2. There will be no significant difference between pre- versus post-intervention test scores on a criterion-referenced test for language-delay children following 5 months of intervention.

Statistical Analysis

Participants' test scores from the pre- and post-intervention were analyzed using a paired t-test to determine if there was a significant difference between pre-versus post-intervention test scores. In addition, for sub-hypothesis (1a), a one-way analysis of variance (ANOVA) was used to determine if there was a

significant degree of improvement on the norm-referenced test in relation to the pre-intervention disability level, as defined by the standard score scale. The alpha significance level was set at $p < 0.05$.

Chapter Summary

In this chapter the method for participant selection was outlined. The sample size of 16 language-delayed children enrolled in speech therapy was stated. Research procedures for data collection and experimental design were outlined. This information was analyzed with a pre- versus post-intervention research design. Participants' test scores from pre- versus post-intervention were analyzed using a paired t-test, with an alpha significance level of $p < 0.05$. Additionally, a one-way ANOVA was used to determine if there was a significant degree of improvement on the norm-referenced test in relation to the pre-intervention disability level, as defined by the standard score scale. The alpha significance level was set at $p < 0.05$.

CHAPTER IV

Results

Demographics

Subjects

Eighteen subjects participated in this investigation. Two subjects were lost to attrition during the 5-month intervention period. Therefore 16 subjects were used for statistical analysis. Subjects were recruited from (a) Miami University Speech and Hearing Clinic (N = 5) and a private practice (N = 1) in Oxford, Ohio; and, (b) Abilities First Foundation (N = 6) and a private practice (N = 4) in Middletown, Ohio. Inclusion criteria for this study were: (a) having a language impairment with standard scores ranging from greater than 1 to less than 2.5 standard deviations below the mean in receptive, expressive, or total language skills, (b) receiving intervention for 5 months; and, (c) speaking English as their primary language. To qualify for the investigation, subjects had to receive a standard score between 55 and 85 on one or more of the PLS-3 subscales: (a) receptive, (b) expressive, or (c) total language. Of the sixteen subjects, 9 were males and 7 were females. Subjects ranged in age from 25 months to 54 months (M = 33) at the onset of the investigation. Each subject received intervention for 5 months. Participants received intervention from a licensed speech-language pathologist.

Intervention. Therapy goals and therapeutic intervention varied for each subject. Each subject had at least one expressive language goal. Twelve of the 16 subjects had goals overtly targeting receptive language skills. Clinicians used a variety of child-centered, clinician-directed, and play-based approaches throughout the investigation.

Table 1 shows the number of sessions each child attended over the 5-month intervention period ($M = 16$). Of the sixteen subjects, one child was seen once every other week. The other fifteen children were scheduled to attend therapy once or twice a week. The number of sessions attended did not reflect cancellations due to illness, vacations, or time conflicts. Additionally, Table 1 provides the subjects' pre-intervention disability levels (PLS-3) for receptive, expressive, and total language skills. Appendix D contains the subjects' pre- and post-intervention raw scores for the norm-referenced test (PLS-3). Appendix E contains the subjects' pre- and post-intervention raw scores for the criterion-referenced test (REEL-2).

Table 1

Subject Demographics

Subject Number	Number of Sessions Attended	Pre-Intervention Level for Receptive Language	Pre-Intervention Level for Expressive Language	Pre-Intervention Level for Total Language
1	27	70-85	70-85	70-85
2	32	70-85	55-69	55-69
3	17	86-100	70-85	86-100

Table 1 con't.

Subject Demographics

Subject Number	Number of Sessions Attended	Pre-Intervention Level for Receptive Language	Pre-Intervention Level for Expressive Language	Pre-Intervention Level for Total Language
4	14	70-85	70-85	70-85
5	11	70-85	70-85	70-85
6	9	70-85	55-69	70-85
7	9	86-100	70-85	70-85
8	8	86-100	70-85	86-100
9	11	86-100	70-85	86-100
10	13	70-85	70-85	86-100
11	28	70-85	70-85	70-85
12	23	70-85	55-69	70-85
13	13	70-85	55-69	55-69
14	8	55-69	55-69	55-69
15	14	55-69	55-69	55-69
16	10	55-69	70-85	55-69

Descriptive and Inferential Statistics for Research Questions

Research Question 1: Will language-delayed children demonstrate significant gains on pre- versus post-intervention norm-referenced test scores following 5 months of intervention?

In order to analyze this question, a statistical paired t-test was used to determine if pre- versus post-intervention norm-referenced (PLS-3) test scores were significantly different. The pre- and post-intervention test scores of the PLS-3 were analyzed to determine the change from pre- to post-test scores.

The mean and standard deviation for pre-versus post-intervention were determined. The pre-versus post-intervention test scores were analyzed to determine the significance level. This analysis was computed for all the subjects combined (N = 16). As seen in Table 2 there was no significant difference between subjects' pre-versus post-intervention (PLS-3) receptive language test scores.

Table 2

Receptive Data: Whole Sample PLS-3 Post/Pre Difference

Subjects	df	Mean	SD	t	p
16	15	3.3125	10.713	1.24	0.2352

Figure 1 and Table 3 demonstrate the difference between pre-versus post PLS-3 receptive language standard scores. Of the 16 subjects, 7 subjects showed a decline or no improvement, whereas 9 subjects demonstrated improvement.

Figure 1. PLS-3 Receptive Language Difference Scores Between Pre-Versus Post-Intervention

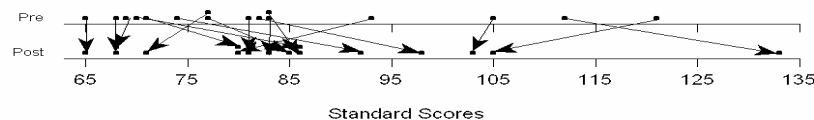
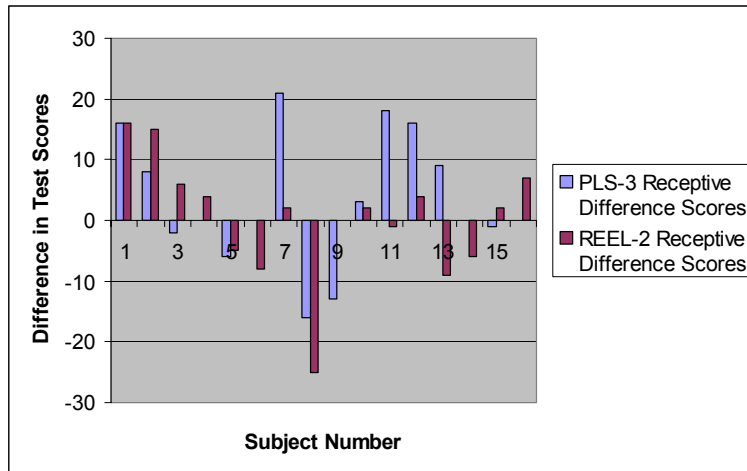


Table 3

Receptive Language Difference Scores Between Pre-Versus Post-Intervention



The mean and standard deviation for pre-versus post-intervention were determined. The pre-versus post-intervention test scores were analyzed to determine the significance level. This analysis was computed for all the subjects combined (N = 16). As seen in Table 4 there was a significant difference between subjects' (PLS-3) pre-versus post-intervention expressive language test scores.

Table 4

Expressive Data: Whole Sample PLS-3 Post/Pre Difference

Subjects	df	Mean	SD	t	p
16	15	9.5625	11.343	3.37	0.0042

Figure 2 and Table 5 demonstrate the difference between pre- versus post PLS-3 expressive language standard scores. Of the 16 subjects, 4 subjects showed a decline or no improvement, whereas 12 subjects demonstrated improvement.

Figure 2. PLS-3 Expressive Language Difference Scores Between Pre-Versus Post-Intervention

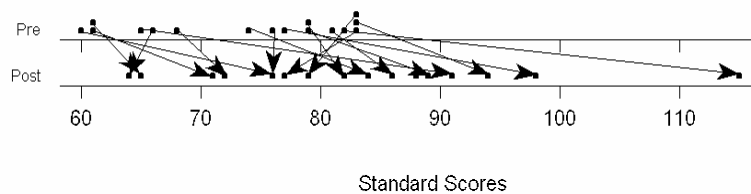
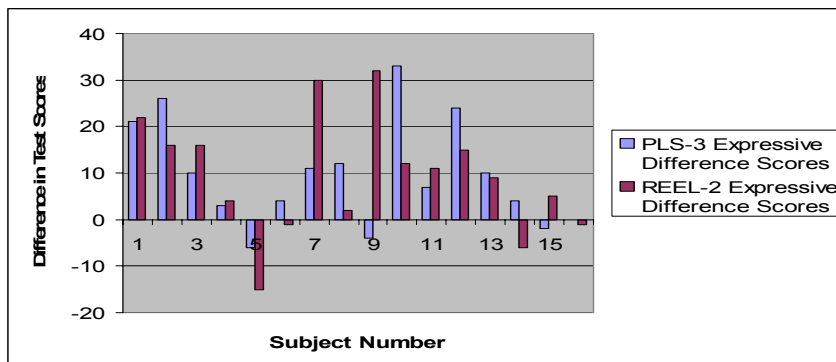


Table 5

Expressive Language Difference Scores Between Pre-Versus Post-Intervention



The mean and standard deviation for pre-versus post-intervention were determined. The pre-versus post-intervention test scores were analyzed to determine the significance level. This analysis was computed for all the subjects combined (N = 16). As seen in Table 6 there was a significant difference between subjects' pre-versus post-intervention (PLS-3) total language test scores.

Table 6

Total Data: Whole Sample PLS-3 Post/Pre Difference

Subjects	df	Mean	SD	t	p
16	15	7.1875	10.628	2.71	0.0163

Figure 3 and Table 7 demonstrate the difference between pre-versus post PLS-3 total language standard scores. Of the 16 subjects, 5 subjects showed a decline or no improvement, whereas 11 subjects demonstrated improvement.

Figure 3. PLS-3 Total Language Difference Scores Between Pre-Versus Post-Intervention

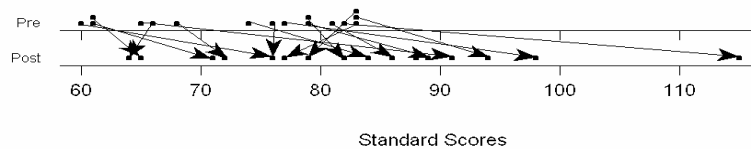
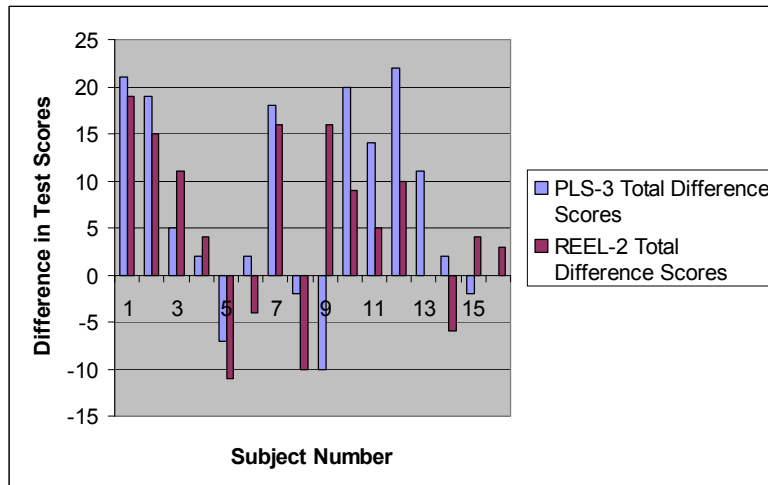


Table 7

Total Language Difference Scores Between Pre-Versus Post-Intervention



The t-test analysis determined that the pre- versus post-intervention PLS-3 receptive language standard scores were not significantly different for all subjects combined ($t(15) = 1.24, p = 0.2352$). However, the t-test analysis did determine that pre- versus post-intervention expressive language ($t(15) = 3.37, p = 0.0042$) and total language ($t(15) = 2.71, p = 0.0163$) standard scores were significant for all subjects combined. The alpha level was set at .05 for all statistics analyzed.

Research Question 1a: Will there be a significant difference in the degree of improvement on the norm-referenced test between the groups of children and their pre-intervention disability levels?

To answer this research question, a one-way analysis of variance (ANOVA) was used to determine if there was a significant difference in the

degree of improvement on the norm-referenced test (PLS-3) in relation to the subjects' pre-intervention disability level.

Subjects' were divided into groups based on their pre-intervention norm-referenced test scores (PLS-3). Groups were formed using standard scores categorized by standard deviation from the mean. For receptive and total language, individuals were divided into three separate standard score categories: (a) 55-69, (b) 70-85, and (c) 86-100. Expressive language standard scores determined two groups: (a) 55-69, and (b) 70-85.

Table 8 presents the receptive data comparison of PLS-3 improvement between norm-referenced groups. Least squares mean and probability data are noted for each group.

Table 8

Receptive Data: Comparison of PLS-3 Improvement Between Norm-Referenced Groups

Subject Groups	Least Squares Mean	p
55-69 (N = 3)	-0.3333	0.9567
70-85 (N = 9)	7.11111	0.0618
86-100 (N = 4)	-2.5000	0.6399

Table 9 shows the receptive data comparison of PLS-3 differences in improvement between norm-referenced groups. Statistical ANOVA values are reported.

Table 9

Receptive Data: ANOVA Analysis of PLS-3 Improvement Between Norm-Referenced Groups

Subjects	df	F	p
16	15	1.40	0.2817

Table 10 presents the expressive data comparison of PLS-3 improvement between norm-referenced groups. Least squares mean and probability data are noted for each group.

Table 10

Expressive Data: Comparison of PLS-3 Improvement Between Norm-Referenced Groups

Subject Groups	Least Squares Mean	p
55-69 (N = 6)	11.00000	0.0369
70-85 (N = 10)	8.700000	0.0336

Table 11 shows the expressive data comparison of PLS-3 differences in improvement between norm-referenced groups. Statistical ANOVA values are reported.

Table 11

Expressive Data: ANOVA Analysis of PLS-3 Improvement Between Norm-Referenced Groups

Subjects	df	F	p
16	15	0.15	0.7087

Table 12 presents the total data comparison of PLS-3 improvement between norm-referenced groups. Least squares mean and probability data were noted for each group.

Table 12

Total Data: Comparison of PLS-3 Improvement Between Norm-Referenced Groups

Subject Groups	Least Squares Mean	p
55-69 (N = 6)	8.66666	0.0771
70-85 (N = 6)	8.66666	0.0771
86-100 (N = 4)	2.7500	0.6272

Table 13 shows the total data comparison of PLS-3 differences in improvement between norm-referenced groups. Statistical ANOVA values are reported.

Table 13

Total Data: ANOVA Analysis of PLS-3 Improvement Between Norm-Referenced Groups

Subjects	df	F	p
16	15	0.43	0.6597

The one-way ANOVA determined that the pre-intervention disability level measured by the PLS-3 did not demonstrate significant improvement for receptive language ($F(2,13) = 1.40, p = 0.2817$), expressive language ($F(1,14) = 0.15, p = 0.7087$), and total language ($F(2,13) = 0.43, p = 0.6597$) test scores for the group comparisons.

The least squares mean revealed some change in the receptive language scores ($p = 0.0618$) for the 70-85 level group. There was some change in total language scores for the 55-69 level group ($p = 0.0771$) and the 70-85 level group ($p = 0.0771$). The alpha level was set at .05 for all statistics analyzed.

Research Question 2: Will language-delayed children demonstrate significant gains on pre- versus post-intervention criterion-referenced test scores following 5 months of intervention?

In order to analyze this question, a statistical paired t-test was used to determine if pre- versus post-intervention criterion-referenced (REEL-2) test scores were significantly different. The pre- and post-intervention test scores were analyzed to determine the change from pre- to post-test scores.

The mean and standard deviation for pre-versus post-intervention were determined. The pre-versus post-intervention test scores were analyzed to determine the significance level. This analysis was computed for all the subjects combined (N = 16). As seen in Table 14 there was no significant difference between the subjects' pre-versus post-intervention REEL-2 receptive language test scores.

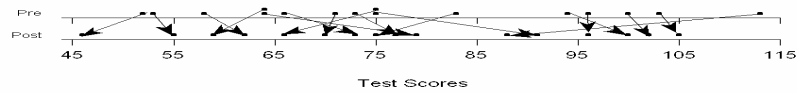
Table 14

Receptive Data: Whole Sample REEL-2 Post/Pre Difference

Subjects	Mean	SD	t	p
16	25	9.8014	.10	0.9201

Figure 4 demonstrates the difference between pre-versus post REEL-2 receptive language scores. Of the 16 subjects, 6 subjects showed a decline or no improvement, whereas 10 subjects demonstrated improvement.

Figure 4. REEL-2 Receptive Language Difference Scores Between Pre-Versus Post-Intervention



The mean and standard deviation for pre-versus post-intervention were determined. The pre-versus post-intervention test scores were analyzed to determine the significance level. This analysis was computed for all the subjects combined (N = 16). As seen in Table 15 there was a significant difference between subjects' pre-versus post-intervention REEL-2 expressive language test scores.

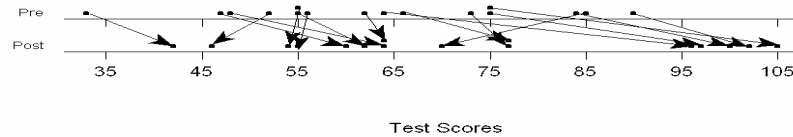
Table 15

Expressive Data: Whole Sample REEL-2 Post/Pre Difference

Subjects	Mean	SD	t	p
16	9.4375	12.58	3.00	0.0090

Figure 5 demonstrates the difference between pre-versus post REEL-2 expressive language scores. Of the 16 subjects, 7 subjects showed a decline or no improvement, whereas 9 subjects demonstrated improvement.

Figure 5. REEL-2 Expressive Language Difference Scores Between Pre-Versus Post-Intervention



The mean and standard deviation for pre-versus post-intervention were determined. The pre-versus post-intervention test scores were analyzed to determine the significance level. This analysis was computed for all the subjects combined ($N = 16$). As seen in Table 16 there was a significant difference between subjects' pre-versus post-intervention REEL-2 total language test scores.

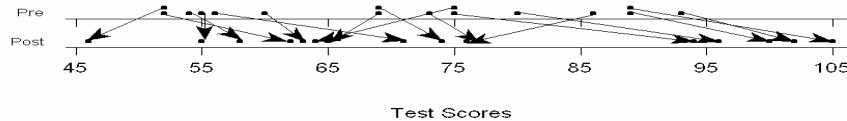
Table 16

Total Data: Whole Sample REEL-2 Post/Pre Difference

Subjects	Mean	SD	t	p
16	5.0625	9.4125	2.15	0.0481

Figure 6 demonstrates the difference between pre-versus post REEL-2 total language scores. Of the 16 subjects, 7 subjects showed a decline or no improvement, whereas 9 subjects demonstrated improvement.

Figure 6. REEL-2 Total Language Difference Scores Between Pre-Versus Post-Intervention



The t-test analysis determined that the pre- versus post-intervention REEL-2 receptive language scores were not significantly different for all subjects combined ($t(15) = .10, p = 0.9201$). However, the t-test analysis did determine that pre- versus post-intervention expressive language ($t(15) = 3.00, p < 0.0090$) and total language ($t(15) = 2.15, p < 0.0481$) test scores were significant for all subjects combined. The alpha level was set at .05 for all statistics analyzed.

Chapter Summary

This chapter presented the statistical results of the data analysis. Research question one determined that there was significant difference between pre-versus post-intervention with regards to the norm-referenced expressive and total language standard scores. Additionally, no significant difference was found between pre-versus post-intervention with regards to the norm-referenced receptive language test scores. No significant difference was found in the degree

of improvement on the norm-referenced test in relation to the pre-intervention disability level. Finally research question three's analyses found a significant difference between pre-versus post-intervention with regards to the criterion-referenced expressive and total language test scores. No significant difference was found between pre-versus post-intervention with regards to the criterion-referenced receptive language test scores.

CHAPTER V

Discussion

This study investigated the effects that 5 months of intervention had on test scores of language-delayed children. The amount of language-delay was assessed with a norm-referenced test (PLS-3) and a criterion-referenced test (REEL-2). Receptive, expressive, and total language skills were measured with both assessment instruments. Finally, subjects were divided into classification groups to determine if there was a significant degree of improvement on the norm-referenced test in relation to the pre-intervention disability level. The subjects were divided by standard scores into classification groups. The subjects' receptive and total language standard scores were divided into 3 groups: (a) 55-69, (b) 70-85, and (c) 85-100; whereas, subjects' expressive language standard scores were divided into 2 groups: (a) 55-69 and (b) 70-85.

The results of the present study demonstrated that 5 months of intervention with language-delayed children yielded significant gains in the areas of expressive and total language skills on both the norm-referenced and criterion-referenced tests. The post-intervention test scores revealed that language-delayed children obtained higher scores following 5 months of intervention. These findings suggested that language-delayed children's expressive and total language scores on a norm-referenced test and a criterion-referenced test improved

significantly after a 5-month intervention period. However, receptive language did not demonstrate significance.

The present findings are consistent with McConkey, Jeffree, and Hewson (1975) who reported that young mentally handicapped children, ranging in age from 40 to 59 months, experienced improved language development after 10 to 14 weeks of intervention. Researchers reported that 4 of the 10 children produced 0 to 10 words before intervention began and after intervention they produced more than 300 words. Furthermore, Ward (1999) reported that children under the age of 1 year, who received 2 to 8 months of intervention, did not require further intervention. Ward (1999) found that the children who received intervention were linguistically within normal limits using a criterion-referenced test one-year, post-intervention assessment. Both studies indicated that children receiving intervention for language delay do improve their language skills over a certain period of time on a criterion-referenced test. The present results support the above notions by demonstrating that language-delayed children do improve their expressive and total language skills on a criterion-referenced test after 5 months of intervention.

The current investigation demonstrated that receiving language intervention for 5 months significantly affected a language-delayed child's ability to communicate with others. From these findings it can be inferred that language intervention does allow language-delayed children the opportunities to refine their

language skills. Ward's (1999) findings verified this study's results by providing data which stated that 30% of the participants, in the control group, who were not receiving early language intervention, were referred for speech and language therapy by their second birthday. None of the children, who received early intervention, required further speech and language therapy by their second birthday. Furthermore, Ezell, Justice, and Parsons (2000) found language-delayed children who received a specified book-reading program for 3 weeks made notable gains in their acquisition of concepts about print and book reading on a norm-referenced test.

The current investigation revealed no significant differences in receptive language skills for either the norm-referenced test or the criterion-referenced test. This finding does not agree with the research presented by Ward (1999) and McConkey, Jeffree, and Hewson (1975). Ward (1999) reported the receptive language scores increased from 65.8 to 99.8 for group 1, and 66 to 101 for group 2 on the criterion-referenced test. These increases in scores were found to be significant ($t = 2.701$, $p < 0.01$). McConkey, Jeffree, and Hewson (1975) found parents producing opportunities for language facilitating schemes, increased their child's ability to identify and locate objects from less than 10 to more than 100. One possible cause for this finding is that the clinicians, who provided the treatment, devoted more time for targeted goals dealing with expressive language skills. Additionally, language-delayed children may develop expressive language

skills at a faster rate, while receptive language skills develop after the expressive language skills are learned.

Analysis of subject data revealed that 50% of the subjects had a standard score of 80 or higher in receptive language skills, whereas only 31% of the subjects scored an 80 or higher in expressive language skills. This would suggest that the majority of subjects presented with low average or greater receptive language skills. Therefore, significant gains with receptive language standard scores would not be expected.

The results of the current study demonstrated that the majority of subjects produced both positive or negative difference scores (Tables 3, 5, and 7) following intervention for the norm-referenced and criterion-referenced tests. However, several subjects produced opposite polarity differences for the norm-referenced versus criterion-referenced tests. That is, the norm-referenced difference score was positive and the criterion-referenced difference score was negative, or vice versa. This occurred for 3 subjects within receptive language scores, 4 subjects within expressive language scores, and 4 within total language scores.

Tables 17 and 18 present subject data regarding sessions attended and pre versus post-intervention test score differences. Children who attended 17 or more sessions demonstrated improvement in all three skill areas: receptive, expressive, and total language on the PLS-3 (Table 17). Children who attended less than 17

sessions demonstrated variable improvement. Children who attended 14 or more sessions demonstrated improvement in all three skill areas: receptive, expressive, and total language on the REEL-2 (Table 18). Children who attended less than 14 sessions demonstrated variable improvement. This would suggest that the greater amount of therapy intervention demonstrates a positive effect upon change in all areas of language.

Table 17

PLS-3 Sessions Attended/ Pre-Post Test Score Differences

Number of Sessions Attended	Receptive Scores	Expressive Scores	Total Scores
8	O	+	O
8	-	+	O
9	+	+	+
9	O	+	O
10	O	O	O
11	-	-	-
11	-	-	-
13	+	+	+
13	+	+	+
14	O	O	O
14	O	O	O
17	+	+	+
23	+	+	+
27	+	+	+
28	+	+	+
32	+	+	+

Key:

+ post-intervention scores increased

- post-intervention scores decreased

O post-intervention scores no change

Table 18

REEL-2 Sessions Attended/ Pre-Post Test Score Differences

Number of Sessions Attended	Receptive Scores	Expressive Scores	Total Scores
8	-	-	-
8	-	O	-
9	-	O	-
9	O	+	O
10	+	O	O
11	-	-	-
11	O	+	+
13	O	+	+
13	-	+	O
14	O	+	+
14	+	+	+
17	+	+	+
23	+	+	+
27	+	+	+
28	O	+	+
32	+	+	+

Key:

+ post-intervention scores increased

- post-intervention scores decreased

O post-intervention scores no change

Subject data revealed that there was not a significant difference in the degree of improvement on the PLS-3 between the groups of children and their pre-intervention disability levels. These results implied that children with a severe language delay did improve as much as the children who only had a mild language delay. One reason for this finding could be the small sample size

included in this investigation. Additionally, the subjects were divided into smaller groups, based upon their pre-intervention performance level (Table 1). Groups included as few as 3 subjects. This small number indicated a higher rate of variability; therefore, no significance was found between the pre-intervention disability levels. Even though significant gains were not found between norm-referenced groups, there was either some indication of change or significance within several groups (Tables 8, 10, and 12). These included (a) receptive language, 70 to 85 group ($p = 0.0618$); (b) expressive language, 50 to 69 group ($p = 0.0369$) and 70 to 85 group ($p = 0.0336$); and, (c) total language, 50 to 69 group ($p = 0.0771$) and 70 to 85 group ($p = 0.0771$).

The current investigation found expressive language and total language scores to be significant for both the norm-referenced (PLS-3) and criterion-referenced (REEL-2) tests. Additionally, there was stronger improvement for expressive language scores on the norm-referenced test (PLS-3) ($t = 3.37$) versus the criterion-referenced test (REEL-2) ($t = 3.00$) (Tables 4 and 15). Similar findings occurred for total language scores on the norm-referenced test (PLS-3) ($t = 2.71$) produced versus the criterion-referenced test (REEL-2) ($t = 2.15$) (Tables 6 and 16).

As confirmed in the present research findings and past research studies, language intervention is viable for enhancing language skills in language-delayed children. Although significant results were not found for receptive language skills

on either the norm-referenced test or the criterion-referenced test, it would be beneficial for future studies to include a larger sample size to see if these results are reliable.

Conclusions

The following conclusions about the research were made.

1. Language-delayed children receiving 5 months of intervention presented significant gains on the expressive and total language skills of a norm-referenced test.
2. Language-delayed children receiving 5 months of intervention did not present significant gains on the receptive language skills of a norm-referenced test.
3. There was not a significant difference in the degree of improvement between any group of language-delayed children on the norm-referenced test in relation to their pre-intervention disability level.
4. Language-delayed children receiving 5 months of intervention presented stronger improvement for expressive and total language skills on the norm-referenced versus criterion-referenced test.
5. Language-delayed children receiving 5 months of intervention presented significant gains on the expressive and total language skills of a criterion-referenced test.

6. Language-delayed children receiving 5 months of intervention did not present significant gains on the receptive language skills of a criterion-referenced test.

Limitations

The present study included several limitations. First, the study was limited in subject size. Sixteen subjects were a small proportion of the language-delayed children in the southwestern Ohio region. Therefore, the statistics should be interpreted cautiously, keeping in mind that only a small segment of language-delayed children in the southwestern Ohio region was represented.

Subjects included in the current study sought treatment for language-delay at one of four participating speech and hearing clinical settings. Within these practices, 9 speech-language pathologists provided treatment to the subjects. It is expected that some variation in treatment may result from differences in the speech-language pathologists' clinical interpretations of the language-delay. The variation in treatment goals occurred when clinicians considered the severity of the language delay, the age of the child, and whether the child presented other language or cognitive deficits.

Another factor that limited the study was the number of treatment sessions a child received over the 5-month period ($M = 16$). For example, one child received therapy twice a week for 5 months, whereas another child received

therapy once every other week for 5 months. The total number of sessions for the child receiving intervention twice a week for 5 months was 32. The child receiving intervention once every other week for 5 months had a total of 8 sessions. This amount of time does not include cancellations due to illness or time conflicts. Therefore, the statistics should be interpreted cautiously, keeping in mind the length of intervention was 5 months, but the number of sessions ranged from 8 to 28 sessions between the subjects.

For receptive language skills, 4 subjects were within the first standard deviation below the mean, 9 within the second standard deviation, and 3 within the third standard deviation. For expressive language skills, 10 subjects were within the second standard deviation below the mean and 6 within the third standard deviation (Table 1). With this in mind, the speech-language pathologist had to determine which language skills were necessary for the child to be a successful communicator. Therefore, one area of language skills could have increased more dramatically than the other language skills not overtly targeted.

Implications for Future Research

The findings of the current study support further research into the effectiveness of language intervention. Fey (1986) noted that findings of intervention effectiveness studies are limited to the time and place of the investigation. This study investigated the changes that occurred in a population over time. Completion of similar studies with other language-delayed children

may lead to a better understanding of the effectiveness of intervention after a specific time period.

Further research is also needed to investigate the relationship that pre-intervention disability levels have on the effectiveness of intervention. In this particular study, there was not any significance difference in degree of improvement between children with mild, moderate, or severe language delay. Therefore, other investigations may concentrate on pre-intervention disability levels with a larger sample size. These investigations may lead to further understanding of how language intervention should be used with children having varying degrees of language delays, and the time period necessary to produce significant language improvements.

Additional studies that merit investigation should include comparisons of pre- versus post-intervention test scores on other norm-referenced and criterion-referenced tests with language-delayed children. Examination of both norm-referenced and criterion-referenced tests may provide information about how the two assessment tools are similar or different in measuring the child's receptive, expressive, and total language skills.

Due to the many variables that are associated with language delay and intervention, no one study could thoroughly examine all the benefits of intervention with language-delayed children. However, multiple research efforts

may result in a greater understanding of intervention with language-delayed children and the associated variables.

Chapter Summary

This chapter provided discussion, conclusions, limitations of the current research, and implications for future research. It was concluded that language intervention had a significant impact on language-delayed children's expressive and total language skills. Additionally, it was found that pre-intervention disability levels did not show significant between-group differences in the degree of improvement with 5 months of intervention. The limitations of the current study were presented and should be considered when interpreting the results. Further research should investigate comparisons of pre- versus post-intervention test scores on other norm-referenced test and criterion-referenced tests. Research should also consider identifying pre-intervention disability levels and the effects it may have on developing language skills.

References

American Speech-Language-Hearing Association. (1983). Committee on Language. Definition of language, *Asha*, 25, 44.

- American Speech-Language-Hearing Association. (1982). Committee on Language, Speech, and Hearing Services in the Schools. Definitions: Communicative disorders and variations. *American Speech-Language-Hearing Association, 24*, 949-950.
- Bankson, N.W. (1990). *Bankson language test-2 screen*. Austin, TX: Pro-Ed.
- Behuniak, P., & Tucker, C. (1992). The potential of criterion-referenced tests with projected norms. *Applied Measurement in Education, 5*(4), 337-353.
- Bell, S., & Ainsworth, M. (1972). Infant crying and maternal responsiveness. *Child Development, 43*, 1171-1190.
- Bradshaw, M.L., Hoffman, P.R., & Norris, J.A. (1998). Efficacy of expansions and cloze procedures in the development of interpretations by preschool children exhibiting delayed language development. *Language, Speech, and Hearing Services in Schools, 29*, 85-95.
- Bzoch, K.R., & League, R. (1991). *Receptive-expressive emergent language test-2*. Austin, TX: PRO-ED, Inc.
- Dale, P.S., Crain-Thoreson, C., Notari-Syverson, A., & Cole, K. (1996). Parent-child book reading as an intervention technique for young children with language delays. *Topics in Early Childhood Special Education, 16*(2), 213-235.

- Eimas, P. (1974). Linguistic processing of speech by young infants. In R. Schiefelbusch & L. Lloyd (Eds.), *Language Perspectives-Acquisition, Retardation, and Intervention*. Baltimore: University Park Press.
- Ezell, H.K., Justice, L.M. & Parsons, D. (2000). Enhancing the emergent literacy skills of pre-schoolers with communication disorders: A pilot investigation. *Child Language Teaching and Therapy*, 16(2), 121-136.
- Fey, M.E. (1986). *Language intervention with young children*. Newton, MA: Allyn & Bacon.
- Fey, M.E., & Cleave, P.L. (1993). Two approaches to the facilitation of grammar in children with language impairment: An experimental evaluation. *Journal of Speech and Hearing Research*, 36(1), 141-168.
- Freeman, L., & Miller, A. (2001). Norm-referenced, criterion-referenced, and dynamic assessment: What exactly is the point? *Educational Psychology in Practice*, 17(1), 3-16.
- Girolametto, L., Pearce, P., & Weitzman, E. (1996). Interactive focused stimulation for toddlers with expressive vocabulary delays. *Journal of Speech and Hearing Research*, 39, 1274-1283.
- Hashway, R.M. (1998). *Assessment and evaluation of developmental learning: Qualitative individual assessment and evaluation models*. Westport: Praeger Publishers.

- Hirschman, R., & Katkin, E. (1974). Psychophysiological functioning, arousal, attention, and learning during the first year of life. In H. Reese (Ed.), *Advances in Child Development and Behavior*. New York: Academic Press
- Johnson, D.L., & Martin, S. (1980). Criterion-referenced testing: New wine in old bottles. *Academic Therapy*, 16(2), 167-173.
- Kearsley, R. (1973). The newborn's response to auditory stimulation: A demonstration of orienting and defense behavior. *Child Development*, 44, 489-507.
- Logan, K.J., & Caruso, A.J. (1997). Parents as partners on the treatment of childhood stuttering. *Seminars in Speech and Language*, 18(4), 309-326.
- MacTurk, R.H., & Neisworth, J.T. (1978). Norm referenced and criterion based measures with preschoolers. *Exceptional Children*, 45(1), 34-39.
- McCauley, R.J. (1996). Familiar strangers: Criterion-referenced measures in communication disorders. *Language, Speech, and Hearing Services in Schools*, 27, 122-130.
- McCauley, R.J., & Swisher, L. (1984). Use and misuse of norm-referenced tests in clinical assessment: A hypothetical case. *Journal of Speech and Hearing Disorders*, 49, 338-348.

- McConkey, R., Jeffree, D.M., & Hewson, S. (1975). Involving parents in extending the language development of their young mentally handicapped children. *British Journal of Disorders of Communication*, 14(3), 203-218.
- Merrell, A.W., & Plante, E. (1997). Norm-referenced test interpretation in the diagnostic process. *Language, Speech, and Hearing Services in Schools*, 28, 50-58.
- Mills, C.N., & Hambleton, R.K. (1980). *Guidelines for reporting criterion-referenced test score information*. Paper presented at the annual meeting of the American Educational Research Association, Boston, MA, April 7-11, ED 189130.
- Moffitt, A. (1971). Consonant cue perception by twenty and twenty-four week old infants. *Child Development*, 42, 717-731.
- Montgomery, P.C., & Connolly, B.H. (1987). Norm-referenced and criterion-referenced tests: Use in pediatrics and application to task analysis of motor skill. *Physical Therapy*, 67(12), 1873-1876.
- Oescher, J., Kirby, P.C., & Paradise, L.V. (1992). Validating state-mandating criterion-referenced achievement tests with norm-referenced test results for elementary and secondary students. *Journal of Experimental Education*, 60(2), 141-150.
- Olswang, L., & Bain, B. (1991). Intervention issues for toddlers with specific language impairments. *Topics in Language Disorders*, 11, 69-86.

- Owens, R.E., Jr. (1996). *Language development: An introduction*. Needham Heights: A Simon & Schuster Company.
- Pamploma, C., & Ysunza, A. (2000). Active participation of mothers during speech therapy improved language development of children with cleft palate. *Scand J Plast Reconstr Hand Surg*, 34, 231-236.
- Paul, R. (1995). *Language disorders from infancy through adolescents: Assessment and intervention*. St. Louis: Mosby-Year Book Inc.
- Roberston, S.B., & Weismer, S.E. (1999). Effects of treatment on linguistic and social skills on toddlers with delayed language development. *Journal of Speech, Language, and Hearing Research*, 42(6), 1234-1248.
- Salvia, J., & Yesseldyke, J. (1981). *Assessment in special and remedial education* (2nd ed.). Boston: Houghton-Mifflin.
- Secord, W., Wiig, E., & Semel, E. (1992). *Clinical evaluation of language fundamentals-preschool*. San Antonio, Tx: The Psychological Corp.
- Semel, E., Wiig, E., & Secord, W. (1989). *Clinical evaluation of language fundamentals-revised screening test*. San Antonio, TX: The Psychological Corp.
- Shipley, K.G., & McAfee, J.G. (1998). *Assessment in speech-language pathology* (2nd ed.). San Diego: Singular Publishing Group Inc.

- Summers, P.A., Larson, G.W., Miguel, S.A., & Terrell, S.L. (1996). Test-retest comparisons using the CELF-RST and BLT-2S with kindergartners. *Language, Speech, and Hearing Services in Schools, 27*, 324-329.
- Turnure, C. (1971). Response to voice of mother and stranger by babies in the 1st year. *Developmental Psychology, 4*, 182-190.
- Venes, D. (Ed.). (2001). *Taber's cyclopedia medical dictionary* (19th ed.). Philadelphia: F.A. Davis Company.
- Walsh, W.B., & Betz, N.E. (1990). *Tests and assessments* (2nd ed.). Englewood Cliffs, NJ: Prentice-Hall.
- Ward, S. (1999). An investigation into the effectiveness of an early intervention method for delayed language development in young children. *International Journal of Language & Communication Disorders, 34*(3), 243-264
- Weitzman, E. (1992). *Learning language and loving it*. Toronto: A Hanen Centre Publication.
- Zimmerman, I.L., Steiner, V.G., & Pond, R.E. (1992). *Preschool language scale-3*. San Antonio: The Psychological Corporation Standardization Harcourt Brace Jovanovich Inc.

Appendix A

Recruitment Letter

August 1, 2002

Dear Parents:

I am currently a 2nd year graduate student in the Department of Speech Pathology and Audiology at Miami University. In order to finish my Master of Arts in Speech-Language Pathology, I am collecting data for my master's thesis. I am conducting a research study on the effectiveness of intervention, as measured by a norm-referenced and criterion-referenced test. Your child is already receiving intervention, and with your permission will be given two language tests prior to and following 20 weeks of intervention. The pre- and post-intervention test results will demonstrate if there is a significant difference between your child's performance on the norm-referenced and/or criterion-referenced tests following 5 months of intervention. The procedures for involving human subjects in research were approved by Miami University's Institutional Review Board for Human Subjects Research on July 8, 2002.

The purpose of this letter is to request your child's participation in this study. Participation is on a voluntary basis. Refusal to participate in this study, will not affect the services the child is receiving in any way. Your child may refuse to answer any questions at any time throughout the study. Your child's attendance at the therapy sessions is required for participation in this study. All data will be treated confidentially and used for research purposes only.

Although any treatment may have potential side effects, language intervention has been used for years to assist children with language-delays without any reported side effects. Administration of norm-referenced and criterion-referenced tests and language intervention are non-invasive and have not been reported as uncomfortable. The only inconvenience to you and your child is scheduling two appointments outside of your child's regular therapy sessions for administration of the tests.

Testing will occur at a time other than the regular therapy session, at no charge for testing. Time of testing will be scheduled after the agreement to participate. Testing time will vary for each child. Results of the free evaluation will be given to the parents.

Please contact Amber Pester, graduate student researcher, at (513) 664-5775 or Dr. Barbara Weinrich, faculty advisor, at (513) 529-2548 with any questions or concerns regarding any portion of this study. Any questions about your child's

rights as a research participant can be forwarded to the Miami University Office for the Advancement of Scholarship and Teaching (513) 529-3734.

Thank you for your cooperation. Your child's participation in this study is greatly appreciated.

Sincerely,

Amber Pester, B.S.
Graduate Student Researcher

Barbara Weinrich, Ph.D.
Faculty Advisor

Appendix B

Consent Form

I, _____, hereby authorize Amber Pester, graduate student researcher, and Barbara Weinrich, faculty researcher, to analyze my child's test results regarding language development. The purpose of this study is to determine if there is a significant difference between pre- and post-intervention test scores for a norm-referenced test and a criterion-referenced test. This study involves my child consistently attending his/her therapy sessions and completing two language tests, which will be administered during assigned therapy sessions. My child's licensed speech-language pathologist will be administering the language tests.

I have read the information above and acknowledge that this study has been explained to me throughout its entirety. The principal investigator has informed me that there are no outside risks associated with this study. I understand that I may telephone Amber Pester, graduate student researcher, at (513) 664-5775 or Dr. Barbara Weinrich, faculty advisor, at (513) 529-2548 should I have any additional questions. I understand that I may also call the Miami University Office for the Advancement of Scholarship and Teaching at (513) 529-3734 with questions about my child's rights as a research participant.

I understand that any information about my child obtained for this study will be kept strictly confidential and that my child will not be identified in any report or publication.

I understand that my child is free to refuse to participate in this study or to withdraw at any time.

My signature below indicates that I freely agree to allow my child to participate in this investigational study.

(Signature of Parent/Guardian)

(Date)

Appendix C

Subject Intervention Data Form

1. Subject's identification number (1-30): _____
2. Date of pre-intervention test administration: _____
3. Date of post-intervention test administration: _____
4. Dates of intervention period: _____ to _____
5. Number of sessions subject attended: _____
6. Length of sessions: _____
7. Check the type(s) of intervention used
 - * play based _____
 - * clinician-directed _____
 - drill _____
 - drill-play _____
 - modeling _____
 - * whole language _____
 - * hybrid _____
 - focused stimulation _____
 - vertical structuring _____
 - milieu teaching _____
 - * child-centered _____
 - self-talk _____
 - parallel talk _____
 - imitation _____
 - expansions _____
 - extensions _____
 - buildups and breakdowns _____
 - * other(s) describe: _____
8. Describe a typical session.
9. List the subject's short-term objectives for intervention period.
 - _____
 - _____
 - _____

Appendix D

PLS-3 Pre/Post Intervention Raw Scores

Subject Number	PLS-3 Pre-Intervention Receptive Language Test Scores	PLS-3 Post-Intervention Receptive Language Test Scores	PLS-3 Pre-Intervention Expressive Language Test Scores	PLS-3 Post-Intervention Expressive Language Test Scores	PLS-3 Pre-Intervention Total Language Test Scores	PLS-3 Post-Intervention Total Language Test Scores
1	82	98	77	98	77	98
2	77	85	65	91	68	87
3	105	103	79	89	91	96
4	83	83	79	82	79	81
5	77	71	83	77	78	71
6	81	81	68	72	72	74
7	112	133	83	94	97	115
8	121	105	74	86	97	95
9	93	80	83	79	87	77
10	83	86	82	115	81	101
11	74	92	81	88	75	89
12	70	86	60	84	61	83
13	71	80	61	71	62	73
14	65	65	61	65	59	61
15	69	68	66	64	64	62
16	68	68	76	76	69	69

Appendix E

REEL-2 Pre/Post Intervention Raw Scores

Subject Number	REEL-2 Pre-Intervention Receptive Language Test Scores	REEL-2 Post-Intervention Receptive Language Test Scores	REEL-2 Pre-Intervention Expressive Language Test Scores	REEL-2 Post-Intervention Expressive Language Test Scores	REEL-2 Pre-Intervention Total Language Test Scores	REEL-2 Post-Intervention Total Language Test Scores
1	75	91	75	97	75	94
2	64	79	48	64	56	71
3	94	100	84	100	89	100
4	73	77	73	77	73	77
5	64	59	85	70	75	64
6	83	75	56	55	69	65
7	103	105	75	105	89	105
8	113	88	62	64	86	76
9	96	96	64	96	80	96
10	100	102	90	102	93	102
11	71	70	66	77	69	74
12	58	62	47	62	52	62
13	75	66	33	42	55	55
14	52	46	52	46	52	46
15	53	55	55	60	54	58
16	66	73	55	54	60	63