

11-1-1985

A Comparative Study of the TEEM and the Morphological Aspects of the BLST and TOLD-P

Kathleen Marie Zuehlsdorff
Portland State University

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



Part of the [Speech Pathology and Audiology Commons](#)

Let us know how access to this document benefits you.

Recommended Citation

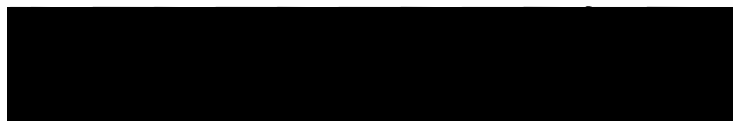
Zuehlsdorff, Kathleen Marie, "A Comparative Study of the TEEM and the Morphological Aspects of the BLST and TOLD-P" (1985). *Dissertations and Theses*. Paper 3525.
<https://doi.org/10.15760/etd.5409>

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

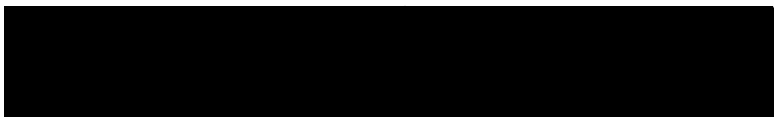
AN ABSTRACT OF THE THESIS OF Kathleen Marie Zuehlsdorff for the Master of Science in Speech Communication, with an emphasis in Speech-Language Pathology, presented November 1, 1985.

Title: A Comparative Study of the TEEM and the Morphological Aspects of the BLST and the TOLD-P.

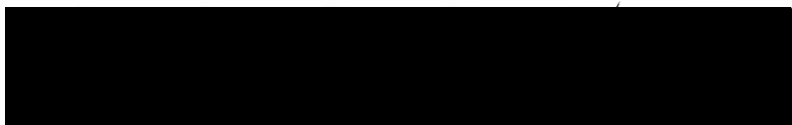
APPROVED BY MEMBERS OF THE THESIS COMMITTEE:



Robert V. Casteel, Chairman



Joan McMahon



Steve A. Brannan

The administration of standardized tests is an important method used by speech-language pathologists in the diagnosis of speech and language problems. Test validity is an important consideration in selecting a measurement tool. The construct validity, or trait measurement, of a test necessitates the accumulation of information from many sources. Correlation with other instruments is one

important procedure used to establish construct validity.

The purpose of this investigation was to examine the construct validity of a new test, which purports to measure morphology, entitled Test for Examining Expressive Morphology (TEEM) (Shipley, Stone and Sue, 1983).

Additional tests of expressive morphology, the Bankson Language Screening Test (BLST) (Bankson, 1977) and the Test of Language Development-Primary (TOLD-P) (Newcomer and Hammill, 1982) were utilized to determine the association of the TEEM with two highly-researched instruments.

Seventy-two preschool, kindergarten, and first-grade children comprised the subjects of this study. All children were chosen randomly from the Portland metropolitan area public schools and preschools. Each subject demonstrated hearing within normal limits and exhibited normal voice quality, fluency, and overall intelligibility. The TEEM, the morphology section of the BLST, and Subtest V of the TOLD-P (Grammatical Completion) were administered to all subjects in one session. Results were recorded as raw scores for all three instruments.

The Pearson product-moment correlation coefficient was used to analyze the raw scores collectively and by age group. Analysis of the data collectively revealed high-moderate correlations between the TEEM and the TOLD-P and between the TEEM and the BLST. With two high-moderate

correlations, the construct validity of the TEEM was supported. In addition, an overall high correlation was demonstrated by the BLST and the TOLD-P.

Mean scores and standard deviations obtained on each instrument by age group revealed a considerable overlap between scores of all ages on the TEEM. Similarly, a considerable overlap between scores of preschool and kindergarten children was shown by both the BLST and the TOLD-P. Scores from the first-grade group did not overlap with younger groups on the BLST, but did overlap with the preschool group on the TOLD-P.

Correlation coefficients by age group per instrument supported the construct validity of the TEEM. A stable, high-moderate association between the TEEM and the TOLD-P was shown for all age groups examined. Moderate correlations were also obtained between the TEEM and the BLST across ages. There appeared to be an undoubtedly high correlation between the BLST and the TOLD-P at the preschool level, but only moderate correlations were exhibited by these tests for the other ages.

Although the results obtained on overlapping test items did not appear to significantly influence the overall correlations obtained, differences between morpheme categories tested may have been a factor affecting the outcome. Three additional considerations which may have

affected the outcome are: 1) fewer items administered to some subjects due to the ceiling suggested by the TOLD-P instruction manual, 2) examiner variability, and 3) test format variables, specifically the absence of visual input to accompany auditory stimuli on the TOLD-P.

A COMPARATIVE STUDY OF THE TEEM
AND THE MORPHOLOGICAL ASPECTS
OF THE BLST AND TOLD-P

by

KATHLEEN MARIE ZUEHLSDORFF

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

MASTER OF SCIENCE IN SPEECH COMMUNICATION:
with an emphasis in
SPEECH-LANGUAGE PATHOLOGY

Portland State University
1985

TO THE OFFICE OF GRADUATE STUDIES AND RESEARCH:

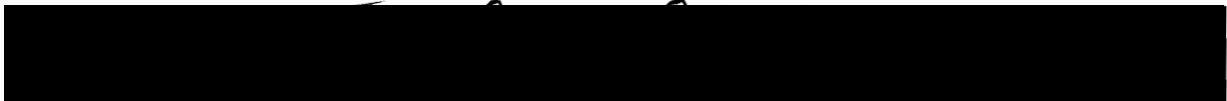
The members of the Committee approve the thesis of
Kathleen Marie Zuehlsdorff presented November 1, 1985.



Robert L. Casteel, Chairman


Joan McMahon


Steve A. Brannan

APPROVED:


Theodore G. Grove, Head, Department of Speech Communication


Jim F. Heath, Interim Dean, Graduate Studies and Research

ACKNOWLEDGMENTS

With great joy I take this opportunity to express my gratitude to the many people who influenced my work on this thesis. First and foremost, a heartfelt thank you to Dr. Casteel for his expert guidance and the countless hours he spent with me reviewing and revising this manuscript. I am especially grateful for the emotional support he generously offered, not only during this project, but throughout my entire graduate program. It has been an honor and a privilege to study with him.

A warm thanks is extended to the other members of my committee, Joan McMahon and Dr. Brannan. Their insights and valuable suggestions helped to refine this study. I particularly appreciated their support and warm congratulations at the conclusion of my orals. I also want to thank Dr. Grove for sharing his expertise in research design and statistical analysis and Jomar Lococo for her valuable suggestions during the initial stages of the study.

To my close friends and study partners, Cheryl Dong and Cathy Pew--thanks not only for your input into this thesis, but for all the times we spent studying, sharing, and encouraging one another.

I am fortunate to have the love and support of two

terrific families. I am particularly grateful to my parents, Richard and Joan Coffman, who instilled in me a love for learning when I was young. To my wonderful sisters, Michelle and Glenda, thank you for listening and always being there for me. And to my dearest friends, Barb and Jim Sieke, who have given me endless hours of companionship and laughter, thank you. You are all very precious to me.

Lastly, and most of all, I thank my loving husband, Gary. His daily encouragement inspired me to "hang in there," especially during the difficult times. His sensitivity and confidence in me provided a continual source of strength; and without him, this work would never have been completed. Gary, with great pleasure and love I dedicate this thesis to you.

TABLE OF CONTENTS

	PAGE
ACKNOWLEDGMENTS	iii
LIST OF TABLES	viii
CHAPTER	
I INTRODUCTION AND STATEMENT OF PURPOSE . .	1
Introduction	1
Statement of Purpose	3
Definition of Terms	4
II REVIEW OF THE LITERATURE	7
Literature Concerning Morphology . .	7
Literature Concerning the TEEM . .	13
Purpose for Using the Test	
Validity	
Reliability	
Test Scores by Age Level	
Literature Concerning the BLST . . .	16
Purpose for Using the Test	
Validity	
Reliability	
Test Scores by Age Level	
Literature Concerning the TOLD-P . .	19
Purpose for Using the Test	
Validity	
Reliability	
Test Scores by Age Level	
Summary	25

CHAPTER		PAGE
III	METHODS AND PROCEDURES	27
	Subjects	27
	General Criteria for Selection	
	Data Collection: Sample Selection .	28
	Physical Setting Hearing Screening Speech Sample	
	Instrumentation	30
	Test for Examining Expressive Morphology Bankson Language Screening Test Test of Language Development- Primary	
	Examiner Reliability	35
	Data Collection: Test Administration	36
	Testing Scoring the TEEM, BLST, and TOLD-P	
	Item Analysis	37
	Data Analysis	38
IV	RESULTS AND DISCUSSION	42
	Results	42
	Discussion	50
V	SUMMARY AND IMPLICATIONS	59
	Summary	59
	Implications	61
	Clinical Research	
	BIBLIOGRAPHY	64

	PAGE
APPENDIX A: PERMISSION LETTER	68
APPENDIX B: TEEM SCORE FORM	70
APPENDIX C: BLST SCORE FORM	72
APPENDIX D: TOLD-P SCORE FORM	74
APPENDIX E: RAW TEST SCORES	76

LIST OF TABLES

TABLE		PAGE
I	Order of Acquisition of Word Formation Rules	9
II	Mean Ratings of Fifty Professionals Reflect- ing on the Nature of the Contents of the TOLD-P Subtests and Composite Scores .	21
III	Correlation Coefficients Depicting the Relationship Between the TOLD-P Subtests and Criterion Tests at Three Age Levels	23
IV	Order of Test Administration	37
V	Item Analysis	39-40
VI	Means and Standard Deviations for all Subjects	43
VII	Pearson r Correlation Coefficients for all Subjects	44
VIII	Means and Standard Deviations for all Subjects By Grade Level	45
IX	Pearson r Correlation Coefficients for all Instruments by Age Group	47
X	Number of Correct Responses to Overlapping Items by Age Level	48-49

CHAPTER I

INTRODUCTION AND STATEMENT OF PURPOSE

INTRODUCTION

A critical component of any program for the communicatively handicapped is the appraisal and diagnosis of speech and language disorders. Speech-language pathologists employ a number of methods to assess speech and language problems, including the use of standardized tests (Peterson and Marquardt, 1981). Instrumentation utilized to assess speech and language skills in children must be appropriate. Test materials must be administered in the child's native language, cannot be racially or culturally discriminatory, must be given by qualified personnel, and must be valid and reliable (Dublinske and Healey, 1978; Neidecker, 1980; Mcloughlin and Gullo, 1984).

To facilitate diagnosis, language is often viewed in terms of receptive and expressive abilities. Language is further divided into the areas of phonology, morphology, semantics, syntax and pragmatics (Dale, 1976). These language segments are of equal importance clinically and develop interdependently. Morphology is of particular interest since this parameter defines the rules by which the

smallest meaningful units of our language are combined (Wiig and Semel, 1984).

An important consideration in selecting a measurement tool is the validity of the instrument. Validity is defined as how well a particular test does what it purports to do. In other words, tests are considered valid to the extent that they serve their function. One of the primary functions of a test is trait measurement, termed as construct validity (Peterson and Marquardt, 1981). The construct validity of a test requires the gradual accumulation of information from many sources. Two important methods used to establish construct validity are factor analysis and correlation with other tests (Peterson and Marquardt, 1981).

Recently a new test which attempts to measure morphology has been developed entitled, Test for Examining Expressive Morphology (TEEM) (Shipley, Stone and Sue, 1983). While other subtests of morphological ability have been devised prior to the TEEM, specific items evaluating morphology are typically buried among tests which measure several language parameters. Examples of such instruments include the Bankson Language Screening Test (BLST) (Bankson, 1977), and the Test of Language Development-Primary (TOLD-P) (Newcomer and Hammill, 1982). The TEEM consists of items selected to measure morphological skills only and requires

a relatively short administration time. Since the BLST and the TOLD-P were devised to examine several areas of language, both require a much lengthier time period to administer.

As new tests are put on the market, research is necessary to establish validity. A successful procedure for doing so is to correlate results from a new test of language ability with other valid, reliable instruments. Since the TEEM is a new test of morphology development, available only since 1983, little research has been done in correlating it with other tests.

STATEMENT OF PURPOSE

The purpose of the proposed investigation is to examine the construct validity of the TEEM by determining the correlation of results from the TEEM with two additional tests of expressive morphology in a group of preschool, kindergarten, and first grade children. This study will seek to answer the following primary question:

What is the association between results obtained on the TEEM and results obtained on two different instruments of expressive morphology?

The following set of questions addresses the association between the three test instruments with respect to expressive morphology and will be used to answer the primary question:

1. What is the correlation between expressive morphology as measured by the TEEM and expressive morphology as measured by the BLST?
2. What is the correlation between expressive morphology as measured by the TEEM and expressive morphology as measured by the TOLD-P?
3. What is the correlation between expressive morphology as measured by the BLST and expressive morphology as measured by the TOLD-P?

DEFINITION OF TERMS

Bound Morpheme--is defined as a morpheme that cannot stand alone to convey meaning. Instead, these morphemes must be attached to free morphemes before meaning occurs. Bound morphemes provide specific meaning to indicate person, number, possession, tense, and degree and include prefixes and suffixes. For example, "ed" is a bound morpheme used to indicate past tense in the word "jumped" (Van Riper, 1978; Shipley, Stone and Sue, 1983).

Construct Validity--construct validity of an instrument refers to the degree to which the instrument measures a specific psychological trait. Construct validity is established by the gradual accumulation of information from many sources (Peterson and Marquardt, 1981).

Expressive Ability--is defined as the ability to produce a meaningful message as in speaking, writing or gesturing (Van Riper, 1978).

Free Morpheme--is defined as a morpheme that can stand alone to convey meaning. For example, words such as "dog," "run," and "table" are considered free morphemes (Shipley, Stone and Sue, 1983).

Morpheme--is defined as a sound or smallest possible combination of sounds that conveys meaning (Van Riper, 1978; Shipley, Stone and Sue, 1983).

Morphology--is defined as the study of how morphemes are combined to form words (Wiig and Semel, 1984).

Phoneme--is a family or group of speech sounds whose components share the same distinctive sound features (Weiss, Lillywhite and Gordon, 1980).

Phonology--involves the sound system of language and is defined as the study of sounds that comprise language and the rules for using sounds (Weiss, Lillywhite and Gordon, 1980).

Pragmatics--is defined as the rules governing the use of language in context (Bates, 1976).

Receptive Ability--is defined as the ability to understand a spoken, written, or gestural message (Van Riper, 1978).

Reliability--refers to the consistency with which a test measures an ability (Newcomer and Hammill, 1982). High reliability of an instrument suggests that similar outcomes are obtained when the instrument is administered to the same subject on two or more occasions (Peterson and Marquardt, 1981).

Semantics--is defined as the meaning of language. Semantics involves both the meaning of individual lexical items and the meaning of sentences as determined by the meanings of individual words and the structure of the sentence (Dale, 1976; Weiss, Lillywhite and Gordon, 1980).

Syntax--refers to the word order of utterances and thus is the study of the arrangement of words and the rules for ordering words (Weiss, Lillywhite and Gordon, 1980).

Validity--Validity of an instrument is defined as the ability of the instrument to measure what it purports to measure. A highly valid test instrument accurately evaluates the attribute for which it was designed (Peterson and Marquardt, 1981).

CHAPTER II

REVIEW OF THE LITERATURE

LITERATURE CONCERNING MORPHOLOGY

A review of the literature concerned with the study of morphology revealed that research has been directed toward the acquisition and development of morphology in young children, the morphological development of children with language impairment, and evaluation instruments. Because a complete review of morphology is beyond the scope and purpose of this study, selected topics have been chosen. In order to narrow the focus, a definition of morphology will be followed by a brief history of normal acquisition and development. A rationale for understanding morphological development will also be provided.

Morphology is defined as the study of the rules for combining morphemes into words. Morphemes refer to the smallest units of meaning in the English language (Vogel, 1977; Weiss, Lillywhite and Gordon, 1980; Wiig and Semel, 1984). Morphemes are one of two types, free or bound. Free morphemes can occur alone, and such words as "car," "table," and "house" are examples of free morphemes. Bound morphemes do not occur alone, but are attached to free morphemes and

as such alter the meaning of the free morpheme. For example, the "s" in the word "cars" is a bound morpheme and changes the free morpheme "car" by denoting pluralization. Assessment of morphological ability is typically limited to bound morphemes. On the other hand, free morphemes are evaluated in the area of semantic ability and are not the primary consideration in morphological assessment (Vogel, 1977; Shipley, Stone and Sue, 1983; Wiig and Semel, 1984).

The acquisition and normal development of morphology in young children has been investigated by a number of researchers (Brown, 1973; Carrow, 1973; de Villiers and de Villiers, 1973). Carrow (1973) reported the age ranges and order of acquisition of selected morphemes which are depicted in Table I.

In an extensive investigation by Brown (1973), the spontaneous speech of three children was studied; and the order of acquisition of a group of 14 morphemes was analyzed. Brown noted that these morphemes generally develop after the two- and three-word stages of language, which Brown denoted as Stage II. He also observed a stable order of acquisition in these children when he established a criterion of 90 percent correct usage in an obligatory context. Brown ranked the order of acquisition as follows:

TABLE I
ORDER OF ACQUISITION OF WORD FORMATION RULES

Word Formation Rule	Application	Age Range of Acquisition (in yrs)		
		By 75%	to	By 90%
Regular Noun Plurals	Balls	3-6	to	6-0
	Coats	5-6	to	6-6
	Chairs	6-6	to	7-0+
Present Progressive Tense	Running	3-0	to	3-6
	Going	3-6	to	5-6
Adjective Forms				
Comparative Superlative	Smaller	4-0	to	5-0
	Biggest	3-0	to	3-6
Noun Derivation				
-er	Hitter	3-6	to	5-0
	Painter	4-0	to	6-0
	Farmer	5-0	to	6-6
-man	Fisherman	5-6	to	6-0
-ist	Bicyclist	7-0	to	7-0+
Adverb Derivation				
-ly	Easily	7-0	to	7-0+

1. Present progressive tense ending, i.e., "-ing"
2. Preposition "in"
3. Preposition "on"
4. Regular noun plural, e.g., "dogs"
5. Past tense irregular, e.g., "went"
6. Noun possessive, e.g., "Adam's chair"
7. Uncontractible copula forms of "to be," e.g., "am," "is," and "are"
8. Articles "a" and "the"
9. Past tense regular, e.g., "Eve walked"
10. Third person regular ending, e.g., "-s" in "Eve walks"
11. Third person irregular forms, e.g., "does"
12. Uncontractible auxiliary, e.g., "I will help"
13. Contractible copula forms of "to be," e.g., "It's big"
14. Contractible auxiliary, e.g., "I'll get it"

Brown (1973) concluded that the order of acquisition is not related to the frequency order of morphemes in the parent's speech, but that both semantic and grammatical complexity determine the order of acquisition.

In a cross-sectional study of the acquisition of morphemes in children's language, de Villiers and de Villiers (1973) supported the conclusion that the order of acquisition may best be predicted by some combination of

grammatical and semantic complexity and perceptibility in speech. Mean length of utterance appeared to be a better predictor of the acquisition of morphemes in the early stages of language development than chronological age. There did not appear to be a relationship between frequencies in parental speech and the child's order of acquisition. These researchers ultimately concluded that it is possible that no single factor can be considered of primary importance in determining the acquisition of morphemes (de Villiers and de Villiers, 1973). Although there were differences in the rank order of the 14 morphemes between the results of Brown and those of de Villiers and de Villiers, the positive correlation was .87 (Carrow-Woolfolk and Lynch, 1982).

Since all English words are composed of one or more morphemes, knowledge of the rules for producing meaningful words is imperative. Skilled speakers use morphology to alter the meaning of root words and to produce the semantic distinctions of number, case, verb tense, third person singular verbs, and comparisons (Wiig and Semel, 1984). Morphology is also used to extend or modify the meaning of root words by the addition of a prefix and/or suffix (Wiig and Semel, 1984).

English speakers need to understand morphology in order to use the language efficiently and effectively. Speech-

language pathologists must be knowledgeable about the normal acquisition and development of morphology in young children in order to appraise, diagnose and treat language impaired children (Wiig and Semel, 1984).

Utilizing a creative and ingenious research design, Berko (1958) investigated the knowledge of morphology in four- to seven-year-old children. The format required children to apply a variety of derivational and inflectional word endings to nonsense words associated with nonsense pictures. An example of a test item was a line drawing of a bird-like figure accompanied by the following text: "This is a wug. Now there is another one. There are two of them. There are two ____." The child was asked to supply the missing word. The rationale for using nonsense words rather than real words was that the task required children to use their knowledge of morphological rules and not just imitation or memory. Comparison of the responses of preschoolers with those of first graders demonstrated that the children acquired the knowledge of morphological rules in an orderly and predictable fashion. The responses also supported the notion that the children first learned a general rule and then modified their knowledge by learning more specific rules (Berko, 1958).

Over the years, however, the effect of using nonsense stimuli has been questioned. Subsequent research (Newfield

and Schlanger, 1968; Dever, 1972) has indicated that children respond more accurately to test items using real words than to nonsense stimuli. These findings have led several authorities (Dever, 1972; Peterson and Marquardt, 1981) to question the validity of nonsense paradigms for assessment of morphologic skills. The authors of the TEEM (Shipley, Stone and Sue, 1983) have emphasized this point to support their development of real word test items.

LITERATURE CONCERNING THE TEEM

Purpose for Using the Test

Shipley, Stone and Sue (1983), in addressing bound morphemes, stated that the Test for Examining Expressive Morphology (TEEM) was designed to "help clinicians evaluate expressive morpheme development with children whose language skills range from three to eight years of age." A group of 89 possible test items was compiled from Stone's (1950) word list for primary readers and Weiss and Lillywhite's (1976) list of 1,000 words children learn first. Fifteen normal three-year-old children were presented with these words to evaluate their familiarity with the list. Test items were selected from words found to be the most familiar to these children.

The authors noted that while other instruments have been developed to examine morphology, many of the tests have

not been standardized, do not yield age level distinctions, are unavailable for general distribution, or do not provide an in-depth sample of morphological abilities. The TEEM was created to aid educators and clinicians in gaining knowledge about morphological development as part of the total language testing data. More specifically, the TEEM was developed to: (1) sample a variety of bound morphemes effectively and efficiently, (2) use an expressive sentence completion model, and (3) differentiate among age levels.

Validity

Stone (1980) initially evaluated the TEEM with 40 normal children between the ages of 3-0 and 7-0 years. The goal of this prestandardized testing was to estimate test validity and reliability and to determine whether children's responses differed at each age level. Three types of validity were considered: content validity, construct validity, and concurrent validity.

According to the author, content validity was presumed within the test construction because a sentence-completion model with familiar stimulus items was used, test items sampled a variety of allomorphs from six major morpheme types, and morphologic structures that children develop between the ages of 2-0 and 8-0 years were chosen. Construct validity was estimated by a correlation of the TEEM scores of each child with his/her chronological age in

months. A Pearson r of .87 ($p < .05$) between TEEM scores and age was reported. Concurrent validity was evaluated by a comparison of the subjects' TEEM scores with the Peabody Picture Vocabulary Test (Dunn, 1965). A Pearson r of .84 ($p < .05$) was reported between these two instruments (Stone, 1980).

Reliability

From the initial 40 subjects (Stone, 1980), the TEEM was evaluated for intratester and intertester reliability. The test was readministered to 12 randomly-selected children after an interval of 7-14 days. A Pearson r of .94 ($p < .05$) was found between test scores from the two test administrations. Intertester reliability was estimated when another speech-language pathologist gave the TEEM to 12 different, randomly-selected children from the original group. An r of .95 was calculated for the test scores obtained by different clinicians (Stone, 1980).

Test Scores by Age Level

The Neuman-Keul's multiple range analysis of variance was utilized to analyze results obtained on the TEEM from four age levels. Significant differences were found in ages of the initial sample population (Stone, 1980). In addition, a follow-up standardization study of the TEEM with 500 children revealed significant differences among the

subjects' scores at each age level. Critical significant differences were noted for every six-month interval between the ages of 3-0 to 5-0 and for every twelve-month interval between the ages of 5-0 and 8-0 years (Sue, 1981). A Pearson r was computed between age and performance on the TEEM. An r of .83 ($p < .05$) was found between the 500 subjects' test scores and their ages in months. An analysis of scoring patterns by gender revealed no significant difference in response between males and females.

LITERATURE CONCERNING THE BLST

Purpose for Using the Test

According to Bankson (1977), the Bankson Language Screening Test (BLST) was developed to: "provide a means by which a number of psycholinguistic as well as perceptual skills could be surveyed in children in a relatively short period of time." Such an instrument is necessary for determining those areas of language in need of further analysis by diagnostic tests. The BLST is comprised of a battery of subtests organized into five general categories and is specifically designed to screen expressive language skills. The general categories include items related to semantic knowledge, syntactic rules, morphological rules, visual perception and auditory perception.

Validity

A comparison was made of scores among 70 children on the BLST and three other widely-used instruments to determine concurrent validity. The following Pearson product-moment correlations were reported:

BLST and the <u>Peabody Picture Vocabulary Test</u> (Dunn, 1965)	r = .54
BLST and the <u>Boehm Test of Basic Concepts</u> (Boehm, 1971)	r = .62
BLST and the <u>Test for Auditory Comprehension of Language</u> (Carrow, 1973)	r = .64

The moderate correlation coefficients obtained suggest that the BLST is measuring a behavior similar to the three comparative tests. However, the test author stressed that the BLST tests behaviors other than those tested on the other measures. Bankson noted that the primary difference between the BLST and the three other comparative tests was that the former test assessed expressive language, while the latter tests assess receptive language. Content validity was based upon the premise that the test items were chosen as representative of the types of skills that speech-language clinicians diagnose and manage (Bankson, 1977).

Further correlations were determined between selected subtests of the BLST and tests designed to assess similar language parameters. The morphology and syntactic subtests were correlated with the Developmental Sentence Scoring (Lee, 1974) with a resulting r value of .76. This

correlation was statistically significant beyond the .01 level (Bankson, 1977).

In a more recent investigation (Blaxley, Clinker and Warr-Leeper, 1983), the performance of 90 children between the ages of four and six years on the BLST and the Fluharty Preschool Speech and Language Screening Test (Fluharty, 1978) was compared with their performance on the Developmental Sentence Scoring (Lee, 1974). The primary purpose of the study was to determine the accuracy of these screening instruments in detecting language impairments. The researchers concluded that the BLST was moderately accurate in identifying kindergarten children in need of further testing who also placed below the tenth percentile on the Developmental Sentence Scoring. Overall, the BLST was considered to be a more accurate screening instrument than the Fluharty Preschool Speech and Language Screening Test; however, test administration time was cited as a drawback in large-scale screening projects when time is limited (Blaxley, Clinker and Warr-Leeper, 1983).

There has been no research conducted to examine the validity of the morphology subtests of the BLST exclusively.

Reliability

Two methods were utilized to determine the reliability of the BLST. Test-retest reliability was evaluated by administering the instrument to a group of 70 children

twice. The second testing occurred one week after the first administration. Point-to-point item analysis revealed reliability at the .94 level. Furthermore, the Kuder-Richardson 20 Test indicated a .96 overall reliability index (Bankson, 1977).

Test Scores by Age Level

Normative data on the BLST were established on a sample population of 637 children between the ages of 4-1 and 8-0 years. The children were members of two preschool classes, 19 public school classes, and two parochial school classes. The classes were chosen as representative of the socioeconomic levels present in the semi-rural counties adjacent to the Washington, D.C. metropolitan area. All children were grouped into six-month age intervals. Like the TEEM, the BLST is more sensitive to developmental differences at the lower end of the established age range (Bankson, 1977).

LITERATURE CONCERNING THE TOLD-P

Purpose for Using the Test

According to Newcomer and Hammill (1982), the Test of Language Development-Primary (TOLD-P) was designed to:

1. identify children who are significantly below their peers in language proficiency;

2. determine children's specific strengths and weaknesses in language skills;
3. document children's progress in language as a consequence of special intervention programs; and
4. serve as a measurement device in research studies involving language behavior.

A linguistic model was chosen as a theoretical base for the instrument. The authors incorporated a number of theoretical perspectives into their conceptual model on which the seven subtests were formulated. Semantics, syntax, morphology and phonology are assessed both receptively and expressively to provide a multi-faceted test of linguistic abilities in 4-0 to 8-11 year old children. It is important to note that all references made to the TOLD-P are based on the 1982 revised edition unless otherwise noted. The original Test of Language Development was devised in 1977.

Validity

The authors reported that special consideration was given in selection of the test items to insure adequate content validity. In addition, a follow-up study was conducted in which 50 professionals were asked to rate the TOLD-P according to the degree to which the subtests measured aspects of two dimensions of language: semantics or syntax and receptive or expressive. Each dimension was rated on a nine-point scale. On one scale, a low score

indicated that the rater thought that the test format was more a measure of semantics than syntax; on the other scale, a low score indicated that the test was seen as more a measure of listening than speaking. A score of five indicated that the rater believed that the test measured the constructs about equally. The mean ratings of the 50 professionals were computed (see Table II). Results of the ratings provided additional support for the content validity of the test (Newcomer and Hammill, 1982).

TABLE II

MEAN RATINGS OF FIFTY PROFESSIONALS REFLECTING ON THE NATURE OF THE CONTENTS OF THE TOLD-P SUBTESTS AND COMPOSITE SCORES

Subtests	Semantics, Syntax	Listening, Speaking	Description
Picture Vocabulary	2	2	Semantics, Listening
Grammatical Understanding	6	2	Syntax, Listening
Oral Vocabulary	2	7	Semantics, Speaking
Sentence Imitation	7	6	Syntax, Speaking
Grammatical Completion	7	6	Syntax, Speaking

Numerous studies have been conducted to evaluate the concurrent validity of the TOLD-P. The authors described at length an investigation in which the TOLD-P subtests were

correlated individually and collectively with six different instruments, the results of which are detailed in Table III.

From the data detailed above, one may conclude that the concurrent validity of the TOLD-P is supported in all but the Grammatic Understanding subtest (Subtest III) for ages four and eight (Newcomer and Hammill, 1982).

In another study, Weber (1982) correlated the TOLD (1977) Grammatic Completion subtest with two other tests of morphology, the ITPA Grammatic Closure (Kirk, McCarthy and Kirk, 1968) and the Berry-Talbott Language Tests: Comprehension of Grammar (Berry, 1966), using 60 learning-disabled and 60 normal children. He reported correlation coefficients of .70 and .64.

The test manual described in great detail a number of studies concerning construct validity of the TOLD-P. These investigations are too numerous to outline individually, but it is important to note that all of the correlations strongly supported the validity of the instrument. Subtest interrelationships, relationship of the TOLD-P to tests of intelligence, relationship of the TOLD-P to tests of school readiness and achievement, factorial analysis of scores, and group differentiation were the areas discussed (Newcomer and Hammill, 1982).

Two recent studies have investigated the usefulness of the TOLD (1977) with children who speak Black American

TABLE III

CORRELATION COEFFICIENTS DEPICTING THE RELATIONSHIP BETWEEN THE TOLD-P SUBTESTS AND CRITERION TESTS AT THREE AGE LEVELS

TOLD-P Subtests	Criterion Tests	Age			Total
		Four	Six	Eight	
Picture Vocabulary	PPVT	.54	.73	.68	.80
Oral Vocabulary	WISC - Vocabulary	.67	.72	.76	.79
Grammatical Understanding	NSST - Receptive	.13*	.76	.47*	.70
Sentence Imitation	DTLA - Related Syllables	.71	.78	.68	.84
Grammatical Completion	ITPA - Grammatical Closure	.55	.49	.78	.78
Word Discrimination	Wepman Auditory Disc.	.62	.70	.57	.69
Word Articulation	Templin-Darley	.86	.73	.64	.84
Total Raw Score	TACL	.63	.72	.73	.79

* The reliability of the criterion test (NSST-Receptive) was too low to permit much correlation, even with a correction formula applied.

PPVT: The Peabody Picture Vocabulary Test (Dunn, 1965)
 WISC: Wechsler Intelligence Scale for Children (Wechsler, 1974)
 NSST: Northwestern Syntax Screening Test (Lee, 1971)
 DTLA: Detroit Tests of Learning Aptitude (Baker and Ireland, 1967)
 ITPA: The Illinois Test of Psycholinguistic Abilities (Kirk, McCarthy and Kirk, 1968)
 Wepman Auditory Disc.: Auditory Discrimination Test (Wepman, 1968)
 Templin-Darley: The Templin-Darley Tests of Articulation (Templin and Darley, 1980)
 TACL: Test for Auditory Comprehension of Language (Carrow, 1973)

English. Weber's (1982) study described above noted that a major limitation in all three of tests was a lack of scoring procedures for Black English speakers. Wiener, Lewnau and Erway (1983) conducted a study, also based on the original TOLD, designed to collect normative data on children who speak Black English. This investigation concluded that the TOLD (1977) is less a test of language development and more a test of knowledge of a second dialect for Black English speakers. Validity of the TOLD with children whose primary language is not standard American English has been questioned (Weber, 1982; Wiener, Lewnau, and Erway, 1983).

Reliability

A number of studies were cited by the authors of the TOLD-P to evaluate test reliability. Internal consistency was studied in a group of 250 children, 50 children in each age category. A statistical analysis was made for each of the subtests as well as the composite scores and overall language score. While there was some question as to the internal consistency of three of the subtests individually, the internal consistency of the instrument on the whole ranged between .87 and .95, depending on the age level. In another study, the internal consistency reliability of the TOLD-P was determined with communicatively impaired children. The coefficient associated with the total score was .95. Test-retest reliability was also evaluated in a

different group of 21 children. The Pearson product-moment coefficient of the total test was .99 (Newcomer and Hammill, 1982).

Test Scores by Age Level

Raw score means and standard deviations for the TOLD-P subtests at different age levels were based on the performance of 1,836 children in the standardization sample. Results clearly indicated that children perform progressively better on all subtests as they grow older. Twenty-one t-tests were calculated to assess the significance of the observed differences. In all cases, the resulting values were significant at or beyond the .05 level of confidence ($p < .05$). Standard scores and percentiles are provided at six-month age intervals between 4-0 and 8-11 years (Newcomer and Hammill, 1982).

SUMMARY

A review of the literature was conducted pertaining to the acquisition and development of morphology in children and to the three test instruments utilized in this investigation. Particular emphasis was placed on the definition of morphology, on a brief history of normal morphological acquisition, and on a rationale for understanding morphological development. Berko's early research in the area of morphology was described,

accompanied by a statement regarding the effectiveness of nonsense stimuli in comparison to lexical item stimuli. Statements were reported by test authors about the purpose, validity, reliability, and test scores by age level for each of the instruments.

CHAPTER III

METHODS AND PROCEDURES

SUBJECTS

General

Seventy-two preschool, kindergarten and first grade Caucasian children comprised the subjects of this study. Children whose ages occurred within a six-month age range of their birthdate were initially selected. The age range of the preschool group was from four years ten months to five years four months (\bar{X} = five years one month). The age ranges of the kindergarten and first grade groups were from five years ten months to six years four months (\bar{X} = six years) and from six years nine months to seven years three months (\bar{X} = seven years), respectively. All children were chosen randomly from the Portland metropolitan area public schools and preschools. A hearing screening was completed and a speech sample was taken from each child. Testing was administered immediately after the hearing screening and speech sample if the child met all of the criteria for selection.

Criteria for Selection

In order to be included in the investigation, each

subject met the following criteria:

1. passed a hearing screening test at 25dB for the pure tone frequencies of 500, 1000, 2000, and 4000 Hz in the better ear;
2. did not have a history of chronic otitis media within the previous six months, as reported by the parent;
3. had no obvious visual or physical limitations;
4. had normal speech intelligibility accompanied by fluency and voice quality within normal limits, as evaluated by the examiner during a three-minute speech sample. Two judges evaluated tape-recorded speech samples of 10% of the subjects to confirm examiner reliability.
5. did not receive previous or current speech-language services;
6. parent permission slips were signed and returned by the parent or guardian.

DATA COLLECTION: SAMPLE SELECTION

Physical Setting

All subjects were tested within their respective schools. Each child went with the examiner from the classroom to another room to be tested. In all cases, the examiner and the child were the only occupants in a given

room, and all rooms were reasonably quiet and adequately lighted. The physical setting was the same for the hearing screening, the language sample, and the test administration.

Hearing Screening

A hearing screening was conducted by the examiner with a portable Maico 2B audiometer. Pure tone frequencies of 500, 1000, 2000, and 4000 Hz were tested at 25dB. Each child chosen as a subject passed all of the frequencies in at least one ear. All subjects were screened for hearing immediately prior to being tested with the TEEM, BLST and TOLD-P. Furthermore, parent permission slips indicated a negative history for chronic otitis media within the previous six months.

Speech Sample

A three-minute speech sample was elicited with a variety of pictures and toys for each child. The objective was to evaluate fluency, voice quality and speech intelligibility. All subjects were judged to be within normal limits for these parameters by the examiner. In a pilot study, seven speech samples were tape-recorded and evaluated for these parameters by the examiner and two judges. Both of the judges were second-year graduate students in speech-language pathology. One hundred percent agreement was obtained between the evaluations of the

examiner and those of the two judges. None of the children evaluated by the judges participated in the study.

INSTRUMENTATION

Test for Examining Expressive Morphology

The TEEM (ShIPLEY, Stone and Sue, 1983) is a verbal, sentence completion test and is intended for administration on an individual basis. It was designed to evaluate the morphological skills of children between 3-0 and 8-0 years of age. The test has 54 plates, each containing one to three black and white line drawings. The examiner is instructed to sit opposite the child and place the test book with the stimulus page facing the child. The verbal stimulus phrases face the examiner. The examiner reads the first sentence and points to the first picture. A motion is made to the adjoining picture, and the next phrase is read. For example, the examiner points to the first picture and says, "Here is a boat." Then the examiner points to the adjoining picture and says, "Here are two ____." The subject responds by completing the sentence. Five practice items are provided. The test is untimed, but requires approximately 6-10 minutes for administration of all the items.

Test items cover the following morphemes: present progressive, regular and irregular plurals, possessives,

third person singulars, regular and irregular past tense, regular and irregular comparatives, and regular and irregular superlatives.

The TEEM should not be administered to children who do not respond well to the task. The authors state that bound morpheme errors given in response to practice items should not be corrected; they would consider such a practice to be "teaching the test" (p. 15). However, according to the authors, if:

you suspect the child possesses the morpheme in question, it may be useful to sample a few unrelated items. Point to nontest items in the environment, and ask, "What are these? What are they doing?" Once sampled, the practice items can be continued.

Responses to each of the items should be written out as fully as possible on the score sheet. Minimally, the regular allomorphs, which are given in brackets on the score sheet, should be noted. All irregulars should be completely transcribed. For example, if the child used the plural form of "dog" correctly, a /z/ would be recorded. If the child said "foots" for "feet," the entire word would be written. This provides a useful base for analysis. The stimulus may be repeated, but failure to respond after the second attempt is considered "no response" and marked NR on the score sheet. In determining the child's raw score, an NR is counted as an error. A raw score is calculated by the total number of correct items.

The manual states that the raw scores can be converted to age-level approximations and provides a table for doing so. For example, a child with a raw score of 19 is said to possess an approximate age-level score between 3-0 and 3-6 years, since his/her ability to score on the TEEM is similar to that of the average child in this age bracket. In addition, mean scores and standard deviations for the TEEM are detailed.

Bankson Language Screening Test

The BLST (Bankson, 1977) is a verbal test and is also intended for administration on an individual basis. The test was designed to screen the expressive language skills of children between 4-0 and 8-0 years of age. The test has 27 plates, each containing a variety of color drawings.

The BLST consists of a battery of 17 subtests organized into five general categories: semantics, morphology, syntax, visual perception, and auditory perception. Each subtest consists of nine items. Again, the examiner sits opposite the child, and the test booklet with the stimulus page faces the child.

Since morphology is the primary area of interest in this investigation, only the 27 items from the morphology subtests were given. The morphology subtests require about 5-7 minutes to administer as compared to 25 minutes for the entire test. All items of the morphology subtests were

administered. The child was asked to give a one-word response to the question posed by the examiner. A sentence-completion format, similar to the TEEM, is also used by the BLST.

Specific directions for the administration and scoring of each of the 17 subtests are included opposite each test plate. Each item is scored as either correct or incorrect, and the following scoring guidelines were recommended by the author:

1. it is permissible to repeat the directions if the child appears confused;
2. the examiner should have a plain sheet of paper available since certain items require the pictures to be covered;
3. it is important to adhere to the suggested model for demonstration items in order to elicit the desired responses, particularly on the morphology subtests;
4. pointing responses are not acceptable;
5. in some cases, more than one answer can be scored as correct. The examiner's judgment regarding appropriateness of the response is considered.

Scores obtained on each of the various subtests may be graphed on the language profile sheet at the end of the score sheet. Raw scores can be converted to a percentile rank by six-month age intervals. Bankson indicated that children who score at or below the 30th percentile need further language assessment and probable intervention. In addition, a table of means and standard deviations for each

of the subtests is provided.

Test of Language Development-Primary

The TOLD-P (Newcomer and Hammill, 1982) is a multi-faceted test of linguistic skills designed for children between 4-0 and 8-11 years of age. The TOLD-P includes five subtests and two supplemental subtests and requires a minimum of 30 minutes to administer. Subtest I (Picture Vocabulary) is a 25-item receptive subtest in which the subject responds by pointing to one of four possible pictures per plate. Subtest II (Oral Vocabulary) is a 20-item expressive subtest in which the subject is asked to define common English words that are spoken by the examiner. Subtest III (Grammatical Understanding) is a 25-item receptive subtest which evaluates syntax. Expressive syntax and grammar are evaluated in Subtest IV (Sentence Imitation). The last subtest, Subtest V (Grammatical Completion), examines the child's ability to recognize and use common morphological forms. A cloze format is utilized which requires the examiner to read unfinished sentences. The child supplies the missing morphological marker. Subtests I, II, III, and IV and the supplemental subtests (Word Discrimination and Word Articulation) were omitted since their focus is on parameters of language other than morphology.

Subtest V was administered, and testing began with the

first item of the subtest. No training items were provided. Testing terminated when the child missed five items in succession. All items above this ceiling were scored as incorrect. Correct responses earned one point, while incorrect responses were denoted with a zero. The total number of points for each subtest was calculated, and this value was designated as the raw score. The raw score was noted at the bottom of the subtest and was recorded on the profile chart of the answer sheet. Subtest V required about 5-8 minutes to administer.

Raw scores can be converted to several variables by use of the norm tables provided at the end of the manual. For each subtest, age scores can be determined as well as standard scores and percentiles for different ages by six-month intervals. An overall Spoken Language Quotient is derived by adding the standard scores of each of the five subtests. Listening, speaking, semantic, and syntactical quotients may also be determined.

EXAMINER RELIABILITY

An important consideration in this study was interjudge reliability. Since the examiner was the only judge, an analysis of the examiner's ability to determine accurate and inaccurate morpheme production was necessary. To do this, all test items were administered to five subjects, and their

responses were recorded. A random sample of 25 items was then selected by an individual not involved in the judging and dubbed onto a second tape. The random sample of test items was scored by the examiner and two judges, all second-year graduate students in speech-language pathology. The scores of the first judge and the examiner revealed 96% agreement. Ninety-two percent agreement was obtained between the scores of the examiner and those of the second judge. In addition, 88 percent agreement was attained by the two outside judges. As a result of this procedure, accurate reliability of the examiner to judge morpheme production in young children was established.

A further consideration was intrajudge reliability. To determine this, the random sample of 25 tape-recorded items was scored again by the examiner and the first judge six months after the initial scoring. The examiner and the judge obtained 100 percent agreement with their original scores. Therefore, intrajudge reliability over time was considered to be 100 percent. The second judge was not available to participate in the second scoring.

DATA COLLECTION: TEST ADMINISTRATION

Testing

The TEEM, BLST, and TOLD-P were administered in one session. The break between the tests varied, depending on

the needs of each child. The average time for the hearing screening, speech sample, and testing varied between 25-35 minutes per child, and the tests were counterbalanced in order of presentation, as shown in Table IV.

TABLE IV
ORDER OF TEST ADMINISTRATION

ORDER #	TEEM	BLST	TOLD-P
1	1	2	3
2	1	3	2
3	2	1	3
4	3	1	2
5	2	3	1
6	3	2	1

The 72 subjects were divided into six groups of twelve each with four preschool, four kindergarten and four first graders in each group. Each group received one of the test orders as depicted above.

Scoring the TEEM, BLST, and TOLD-P

The TEEM, BLST, and TOLD-P were scored according to their respective instruction manuals. Results were recorded as raw scores for all three instruments.

ITEM ANALYSIS

A detailed item analysis of the TEEM, BLST, and TOLD-P

was conducted in order to determine the number of identical, overlapping items among the three tests. The number of morphemes examined by the TEEM, BLST, and TOLD-P is 54, 27, and 30 respectively. Of the 54 items on the TEEM, six items (11 percent) overlap with the BLST. Of the 27 items on the BLST, six items (22 percent) are replicated on the TEEM. A comparison of the TEEM and TOLD-P revealed that of the 54 items on the TEEM, eight items (14.8 percent) overlap with the TOLD-P, and 26.6 percent of the 30 items on the TOLD-P are replicated on the TEEM. Similarly, a comparison of the BLST and TOLD-P revealed an overlap of four items. Of the 27 items on the BLST, 24.8 percent overlap with the TOLD-P; and of the 30 items on the TOLD-P, 13.3 percent overlap with the BLST.

It is also important to note that three items (10 percent) on the TOLD-P test a category of morphemes not found on either of the other two tests (derivational nouns). Furthermore, approximately 33 percent of the BLST consists of free morphemes which have no relationship to items on either the TEEM or the TOLD-P. A complete item analysis is detailed in Table V.

DATA ANALYSIS

The set of questions asked in the statement of purpose required several analyses of the data. Nonetheless, all the

TABLE V
ITEM ANALYSIS

Category	TEEM (1)	BLST (2)	TOLD-P (3)	Item Overlap		
				1&2	1&3	2&3
Present	writing	running	playing	1	0	1
Progressive	reading crawling swinging	reading swimming	riding swimming jumping			
Plurals						
/s/	blocks cakes	books		0	0	0
/z/	dogs flowers	pennies		0	0	0
/ez/	brushes watches houses	boxes	dressess	0	0	0
/vz/	leaves knives calves		leaves	0	1	2
Irr	feet children teeth women	children men	mice men women	1	1	1
Possessives						
/z/	monkey's cowboy's		mother's woman's boy's children's	0	0	0
/s/	rabbit's cat's			0	0	0
/ez/	mouse's witch's nurse's			0	0	0
Third Person Singular						
/z/	climbs drives	runs reads	drives	0	1	0
/s/	eats walks		eats	0	1	0
/ez/	washes pushes dances			0	0	0

TABLE V (continued)

Category	TEEM (1)	BLST (2)	TOLD-P (3)	Item Overlap		
				1&2	1&3	2&3
Past Tense						
/d/	combed emptied	smiled climbed	played	0	0	0
/t/	dropped roped		cooked	0	0	0
/ed/	melted planted counted			0	0	0
Irr	drawn cut drank caught fed		threw rode drew eaten	0	0	0
Adjectives						
/er/	smaller shorter longer bigger	bigger	bigger smaller	1	2	1
/est/	smallest shortest longest biggest	biggest	smallest	1	1	0
Irr	better best	better best	more most best	2	1	1
Derivational Nouns			singer painter drummer	0	0	0
Future Tense		will fall		0	0	0
Pronouns		her them him she they he hers theirs his		0	0	0
			TOTAL	6	8	4

questions were concerned with how the data from one instrument correlated with the data from one of the other tests.

The Pearson product-moment correlation coefficient was obtained for all data analyzed by use of the Honeywell 6640 computer at Portland State University. Correlations were made between the TEEM raw score and the raw score of the morphology section of the BLST. Correlations were made between the TEEM raw score and the raw score of Subtest V on the TOLD-P. Correlations were also obtained between the morphology section raw score of the BLST and the raw score of Subtest V of the TOLD-P. Correlations were obtained to analyze the data collectively and by age group.

CHAPTER IV

RESULTS AND DISCUSSION

RESULTS

The purpose of this study was to examine the construct validity of the TEEM by determining the correlation of results from the TEEM with two additional tests of expressive morphology in a group of preschool, kindergarten, and first grade children. Raw scores were obtained on the TEEM, BLST, and TOLD-P for each subject. The Pearson product-moment correlation coefficient was used to analyze the data, collectively and by age group. Data analysis of the total scores for each instrument will be presented first, followed by the results obtained for each group.

In total, 72 children were tested with the TEEM, BLST, and TOLD-P, and a total mean score and standard deviation were calculated per instrument. Table VI illustrates the mean scores and standard deviations obtained for all subjects. (See Appendix E for the raw data.)

Table VI shows a mean score of 43.4722 for the TEEM, with a standard deviation of 4.0175. The total number of items on the TEEM is 54, the range was 30 to 51, and the median score was 41. Next, a mean score of 22.0278 and a

TABLE VI
MEANS AND STANDARD DEVIATIONS
FOR ALL SUBJECTS

Test	Mean	Standard Deviation
TEEM	43.4722	4.0175
BLST	22.0278	2.6429
TOLD-P	20.4444	3.6420

standard deviation of 2.6429 were obtained on the BLST from a total of 27 possible items. The range on the BLST was 13 to 27, and the median score was 20. Lastly, the TOLD-P yielded a mean score of 20.4444 and a standard deviation of 3.6420 from a possible 30 items. The range on the TOLD-P was 11 to 28, and the median score was 20.

The Pearson product-moment correlation coefficient (Pearson r) was used to answer the primary question: what is the association between results obtained on the TEEM and results obtained on two different instruments of expressive morphology? Table VII outlines the correlations obtained between each of the tests, utilizing all test data.

The correlation between expressive morphology as measured by the TEEM and expressive morphology as measured by the BLST is a positive .6461. The correlation between expressive morphology as measured by the TEEM and expressive morphology as measured by the TOLD-P is a positive .6795.

TABLE VII
PEARSON r CORRELATION COEFFICIENTS
FOR ALL SUBJECTS

Test	TEEM	BLST	TOLD-P
TEEM		.6461 p = 0.000*	.6795 p = 0.000
BLST	.6461 p = 0.000		.7113 p = 0.000
TOLD-P	.6795 p = 0.000	.7113 p = 0.000	

* p < .001

Lastly, the correlation between expressive morphology as measured by the BLST and expressive morphology as measured by the TOLD-P is a positive .7113.

In addition to results obtained for all subjects collectively, the data were statistically analyzed by age group. Twenty-four children were tested in each grade level, and mean scores and standard deviations were established per instrument. Table VIII depicts the mean scores and standard deviations obtained for each category.

Table VIII reveals an interesting pattern of mean scores and standard deviations for each instrument by age group. The TEEM mean scores increase by approximately two items as age increases; however, there is a substantial overlap

TABLE VIII
 MEANS AND STANDARD DEVIATIONS FOR ALL SUBJECTS
 BY GRADE LEVEL

Test	Mean	Standard Deviation
<u>TEEM</u>		
Preschool	41.2917	3.5322
Kindergarten	43.2083	4.0538
Grade 1	<u>45.9167</u>	<u>3.0175</u>
Total	43.4722	4.0175
<u>BLST</u>		
Preschool	20.7500	2.8780
Kindergarten	21.5833	2.3204
Grade 1	<u>23.7500</u>	<u>1.7004</u>
Total	22.0278	2.6429
<u>TOLD-P</u>		
Preschool	19.3333	3.7144
Kindergarten	18.9583	2.9559
Grade 1	<u>23.0417</u>	<u>2.7894</u>
Total	20.4444	3.6420

between groups as illustrated by the standard deviation scores. Similarly, the mean scores of the BLST increase as age increases, but an overlap is only apparent between the preschool group and the kindergarten group. In contrast, the mean scores for the TOLD-P do not follow a pattern of increasing with age, since the mean score for the kindergarten group was lower than the mean score for the

preschoolers. There is an overlap in standard deviation scores for the TOLD-P between the first two groups and between the preschool and first-grade groups, but no apparent overlap between the kindergarten and first-grade groups.

The Pearson product-moment correlation coefficient was also used to calculate the association between each test for each age category. Table IX outlines the correlations obtained between each of the instruments by subject group.

Table IX reveals a stable, high-moderate correlation between the TEEM and the TOLD-P for all age categories. Moderate correlations were also obtained between the TEEM and the BLST for all groups; however, the correlations varied from a positive .4405 to .6013. A high correlation was calculated between the BLST and the TOLD-P at the preschool level; nonetheless, low-moderate correlations were obtained for the kindergarten and first-grade groups.

Lastly, a detailed analysis of overlapping test items was completed in order to determine how many items missed on one test were correct on another instrument. In order to understand the relationship between overlapping test items, a comparison of the correct number of responses for a given item on one test and the correct number of responses for the same item on a different test is critical. Table X provides information regarding the number of correct and incorrect

responses for all overlapping test items.

TABLE IX
PEARSON r CORRELATION COEFFICIENTS FOR ALL INSTRUMENTS
BY AGE GROUP

Test	TOLD-P	BLST
<u>TEEM</u>		
Preschool	.6318 p = 0.000*	.5678 p = 0.002
Kindergarten	.6249 p = 0.001	.6013 p = 0.001
First Grade	.6177 p = 0.001	.4405 p = 0.016
All Subjects	.6795 p = 0.000	.6461 p = 0.000
<u>BLST</u>		
Preschool	.8175 p = 0.000	
Kindergarten	.5045 p = 0.006	
First Grade	.4423 p = 0.015	
All Subjects	.7113 p = 0.000	

* This correlation is statistically significant at or beyond the .05 level.

Ten of the fourteen overlapping test items revealed a strong association between tests. Ninety to one hundred percent of the children responded consistently across test instruments for identical items. The exception included the following four items: "bigger," "smaller," "best," and "children." For the item "bigger," the TEEM and BLST

TABLE X
NUMBER OF CORRECT RESPONSES TO OVERLAPPING ITEMS BY AGE LEVEL

Item	TEEM	BLST	TOLD-P
<u>Reading</u>			
Preschool	24	24	
Kindergarten	24	22	
Grade 1	24	24	
Total	<u>72</u>	<u>70</u>	
<u>Swimming</u>			
Preschool		24	24
Kindergarten		24	22
Grade 1		24	24
Total		<u>72</u>	<u>70</u>
<u>Leaves</u>			
Preschool	3		4
Kindergarten	4		4
Grade 1	7		8
Total	<u>14</u>		<u>16</u>
<u>Children</u>			
Preschool	3	3	
Kindergarten	6	1	
Grade 1	9	9	
Total	<u>18</u>	<u>13</u>	
<u>Women</u>			
Preschool	3		5
Kindergarten	8		6
Grade 1	14		13
Total	<u>25</u>		<u>24</u>
<u>Men</u>			
Preschool		11	12
Kindergarten		11	11
Grade 1		17	20
Total		<u>39</u>	<u>43</u>
<u>Drives</u>			
Preschool	23		24
Kindergarten	24		23
Grade 1	24		24
Total	<u>71</u>		<u>71</u>

TABLE X (continued)

Item	TEEM	BLST	TOLD-P
<u>Eats</u>			
Preschool	24		21
Kindergarten	24		24
Grade 1	24		24
Total	<u>72</u>		<u>69</u>
<u>Bigger</u>			
Preschool	21	20	14
Kindergarten	19	17	14
Grade 1	<u>23</u>	<u>22</u>	<u>21</u>
Total	<u>63</u>	<u>59</u>	<u>49</u>
<u>Biggest</u>			
Preschool	23	23	
Kindergarten	23	23	
Grade 1	<u>23</u>	<u>23</u>	
Total	<u>69</u>	<u>69</u>	
<u>Smaller</u>			
Preschool	20		7
Kindergarten	21		4
Grade 1	<u>22</u>		<u>10</u>
Total	<u>63</u>		<u>21</u>
<u>Smallest</u>			
Preschool	20		21
Kindergarten	20		16
Grade 1	<u>22</u>		<u>20</u>
Total	<u>62</u>		<u>57</u>
<u>Better</u>			
Preschool	4	4	
Kindergarten	6	4	
Grade 1	<u>10</u>	<u>9</u>	
Total	<u>20</u>	<u>17</u>	
<u>Best</u>			
Preschool	6	9	5
Kindergarten	9	7	0
Grade 1	<u>18</u>	<u>15</u>	<u>9</u>
Total	<u>33</u>	<u>31</u>	<u>14</u>

revealed similar numbers of correct responses, but the TOLD-P revealed a correct number of responses substantially below the other two tests. Similarly, for the item "best," the correct number of responses strongly agreed between the TEEM and the BLST, but the correct number of responses on the TOLD-P was far below those of the other two instruments. For the item "smaller," an almost inverse relationship was depicted between the correct number of responses on the TEEM and the correct number of responses on the TOLD-P, with 63 correct responses on the TEEM and only 21 correct responses on the TOLD-P. For the item "children," the only difference in the correct number of responses was at the kindergarten level.

DISCUSSION

Results of the study appear to substantially support the construct validity of the TEEM, with high-moderate correlations between the TEEM and two additional tests of morphology. At first glance, however, the two additional instruments appear to have a stronger association with each other than with the TEEM. Analysis of the data collectively revealed a high-moderate correlation between the TEEM and the TOLD-P, a high-moderate correlation between the TEEM and the BLST, and a high correlation between the BLST and the TOLD-P.

Results of the mean scores obtained on the TEEM by age group revealed a consistent increase of the mean with an increase in age. Nonetheless, the standard deviation for each group revealed a consistent overlap and an indication that the TEEM did not discriminate particularly well between ages.

Like the TEEM, the mean scores obtained on the BLST increased as age increased. The standard deviation score for the preschool group overlapped with the kindergarten group, and the kindergarten score overlapped with both of the other groups, supporting the contention that the BLST also did not discriminate well between these two age groups. On the other hand, the standard deviation score for the first-grade group did not suggest an overlap between first-grade scores and kindergarten scores. The results of the BLST appeared to adequately discriminate the oldest age category from the other ages.

In contrast to the first two instruments, the TOLD-P mean scores did not increase as a function of age. The mean score for the kindergarten group was slightly less than the mean score for the preschool group. In addition, the standard deviation scores for the preschool and kindergarten groups did not support a clear difference between age groups. However, the mean score for the first-grade group was greater than both of the other scores, and

the standard deviation supported a clear differentiation of first-grade scores from kindergarten scores. An overlap between preschool scores and first-grade scores was noted.

A possible explanation for the inconsistent increase of mean scores on the TOLD-P is revealed when a comparison is made with standardization data described in the manual. At similar age levels, approximately 350-450 children were tested in each group, compared to the 24 children examined per group in this study. The mean scores by age were approximated and reported as 14, 17, and 22, with standard deviations of 7, 7, and 6, respectively. The mean scores obtained in this investigation are within one point of the manual mean scores at the kindergarten and first-grade levels. However, there is a difference of five between the mean scores of the preschool children in this investigation and those in the standardization study. Perhaps a smaller N and an above-average group of preschool children are responsible for a high mean score at this level. It is critical to note, nonetheless, that the mean score obtained in this study is still within one standard deviation of the mean reported in the TOLD-P manual.

Further evidence to support a high-scoring group of preschool children is found when the TEEM and BLST standardization mean scores are compared to this study. Again, the mean scores of this study are within one point of

the TEEM manual scores at the kindergarten and first-grade levels. Yet, the preschool mean score is four points higher, but still within one standard deviation as reported by the TEEM authors. Similarly, the mean scores are within two items of the BLST manual scores at the kindergarten and first-grade levels. The preschool mean is three points higher, but still within one standard deviation, as reported by the BLST author.

On the whole, it appears that the mean scores of the kindergarten and first-grade groups in this study were similar to those reported in standardization studies across all three tests. The mean scores of the preschool group were consistently higher across tests than the reported manual mean scores, but were still within one standard deviation.

Further examination of the correlation coefficients obtained by age group per instrument also supports the construct validity of the TEEM. The overall correlation between the TEEM and the TOLD-P was a positive .6795. The correlations obtained between the TEEM and the TOLD-P for the preschool, kindergarten, and first-grade groups were .6318, .6249, and .6177, respectively. These scores support a high-moderate association between the TEEM and the TOLD-P for all age groups examined.

The construct validity of the TEEM is also supported by

its association with the BLST. The overall correlation between the TEEM and the BLST was a positive .6461, slightly lower than the score for the TEEM and the TOLD-P. The correlations obtained between the BLST and TEEM for the preschool, kindergarten and first-grade groups were .5678, .6013, and .4405, respectively. The scores support a moderate association between the TEEM and BLST for all age groups examined.

The high positive correlation of .7113 obtained between the TOLD-P and the BLST is somewhat misleading upon further analysis of the correlation scores obtained by age group. The correlations obtained for the preschool, kindergarten, and first-grade groups were .8175, .5045, and .4423. While there appears to be an unquestionably high association between the BLST and the TOLD-P for the preschool group, there is only a moderate association apparent between the tests for the other ages. The association between the TOLD-P and the BLST was variable across ages and did not support a stable, high correlation between the instruments for all age groups tested.

A discussion of the results of overlapping test items is also an important consideration. Twenty-two percent of the TEEM, 30 percent of the BLST, and 33 percent of the TOLD-P overlap with items on the other two instruments. Scores obtained on the overlapping test items revealed that

ten of the fourteen items, or 71 percent, had a strong association across tests. In other words, 90-100 percent of the children responded consistently across test instruments for ten identical items. For the remaining four items, results were inconsistent across age groups and across tests. With one item, a correct score on the TEEM almost always yielded an incorrect score on the TOLD-P. In general, results of overlapping test items also support the construct validity of the TEEM, with the exception of four items. However, since the discrepancy between four overlapping test items was greatest between the TEEM and the TOLD-P and the correlation between these instruments was the most stable across age groups, results of these four overlapping test items did not appear to significantly affect the overall correlations obtained.

An additional point with respect to item analysis is the fact that approximately 33 percent of the BLST consists of free morphemes and has no relationship to items on either of the other two instruments. Furthermore, 10 percent of the items on the TOLD-P test a category of morphemes not examined by the other instruments (derivational nouns). The differences between morpheme categories tested is a factor which possibly weakened the association among the instruments.

Another point to examine is that the TOLD-P does not

provide training items nor allow repetition of instructions at the beginning of the test. This is a factor which possibly may have affected the outcome of the study. After careful analysis, however, the examiner concluded that the lack of training items on the TOLD-P was not significant because only four of the 72 subjects missed one item on the first five items administered.

Another important consideration which may have influenced the outcome was the fact that the entire morphology subtest of the TOLD-P was not given to every subject. The test was administered per manual instructions, which suggested a cut-off of five incorrect consecutive items. Consequently, not every test item of the TOLD-P was administered to every child. Yet, all items of the TEEM and the morphology section of the BLST were administered. This difference in procedures between instruments is a factor which may have influenced the results of this study.

It is critical to note that results on only one of the fourteen overlapping test items ("best") may have been influenced by the TOLD-P cut-off criteria. The remaining thirteen overlapping items were administered to every subject. Nonetheless, since only three subjects were not given an opportunity to respond to "best," the cut-off procedure does not effectively account for the discrepancy in results obtained for "smaller," "bigger," "children," or

"best."

Examiner variability from prescribed test manual procedures is an issue for discussion. The examiner took some liberty in accepting responses that may have been questionable by strict manual rules. For example, if the desired response was "smaller" and the child responded with "littler," the response was counted as correct. As long as the child's answer contained the bound morpheme in question and a reasonable or logical substitution for an item was produced, the examiner responded favorably. Since only four children produced alternative responses, examiner variability is a slight consideration which may have influenced the results of the study.

Finally, test format is a factor which may have affected the association between the TEEM and the other two tests. The TEEM utilized a distinct visual format accompanied by gestures and a sentence completion task which appeared to attract the children. For the most part, subjects responded with ease to the TEEM and seemed to enjoy the task. While the BLST also utilized a sentence completion format and pictures, some children seemed to have more difficulty understanding the training items on the BLST as compared to the TEEM. The input for the TOLD-P was purely auditory, and some subjects appeared restless without pictures to accompany the test. Furthermore, the sentences

CHAPTER V

SUMMARY AND IMPLICATIONS

SUMMARY

The administration of standardized tests is an important method used by speech-language pathologists in the diagnosis of speech and language problems. Test validity is an important consideration in selecting a measurement tool. The construct validity, or trait measurement, of a test necessitates the accumulation of information from many sources. Correlation with other instruments is one important procedure used to establish construct validity.

The purpose of this investigation was to examine the construct validity of a new test, which purports to measure morphology, entitled Test for Examining Expressive Morphology (TEEM) (Shipley, Stone and Sue, 1983). Additional tests of expressive morphology, the Bankson Language Screening Test (BLST) (Bankson, 1977) and the Test of Language Development-Primary (TOLD-P) (Newcomer and Hammill, 1982) were utilized to determine the association of the TEEM with two highly-researched instruments.

Seventy-two preschool, kindergarten, and first-grade children comprised the subjects of this study. All children

were chosen randomly from the Portland metropolitan area public schools and preschools. Each subject demonstrated hearing within normal limits and exhibited normal voice quality, fluency, and overall intelligibility. The TEEM, the morphology section of the BLST, and Subtest V of the TOLD-P (Grammatical Completion) were administered to all subjects in one session. Results were recorded as raw scores for all three instruments.

The Pearson product-moment correlation coefficient was used to analyze the raw scores collectively and by age group. Analysis of the data collectively revealed high-moderate correlations between the TEEM and the TOLD-P and between the TEEM and the BLST. With two high-moderate correlations, the construct validity of the TEEM was supported. In addition, an overall high correlation was demonstrated by the BLST and the TOLD-P.

Mean scores and standard deviations obtained on each instrument by age group revealed a considerable overlap between scores of all ages on the TEEM. Similarly, a considerable overlap between scores of preschool and kindergarten children was shown by both the BLST and the TOLD-P. Scores from the first-grade group did not overlap with younger groups on the BLST, but did overlap with the preschool group on the TOLD-P.

Correlation coefficients by age group per instrument

supported the construct validity of the TEEM. A stable, high-moderate association between the TEEM and the TOLD-P was shown for all age groups examined. Moderate correlations were also obtained between the TEEM and the BLST across ages. There appeared to be an undoubtedly high correlation between the BLST and the TOLD-P at the preschool level, but only moderate correlations were exhibited by these tests for the other ages.

Although the results obtained on overlapping test items did not appear to significantly influence the overall correlations obtained, differences between morpheme categories tested may have been a factor affecting the outcome. Three additional considerations which may have affected the outcome are: 1) fewer items administered to some subjects due to the ceiling suggested by the TOLD-P instruction manual; 2) examiner variability, and 3) test format variables, specifically the absence of visual input to accompany auditory stimuli on the TOLD-P.

IMPLICATIONS

Clinical

The results of this study lend support to the construct validity of the TEEM as an accurate instrument for measuring expressive morphology. Use of the TEEM by speech-language pathologists as part of a diagnostic battery in a school or

clinical setting is, therefore, appropriate. Furthermore, the ease of administration and scoring, short administration time, stable validity across age groups, and appealing format suggest that the TEEM may be a more suitable test than other tests of expressive morphology which are buried in lengthier tests.

The examiner preferred the TEEM over the BLST or the TOLD-P. The TEEM manual was straightforward, and administration and scoring were easy to learn. The children seemed to respond with more enthusiasm to the TEEM than to the other instruments and did not appear confused by test instructions. Attention to the task did not seem to be a problem with the TEEM, nor did the children complain about the length of the test. Pragmatically, the TEEM picture stimuli and test format are appropriate for young children. With almost twice as many items on the TEEM compared with those on the other two instruments, a broader sampling of bound morphemes gave the examiner confidence in the results. All factors considered, the examiner supports use of the TEEM over the BLST and the TOLD-P.

Research

Further research involving the TEEM is important to establish its acceptability and use by speech-language pathologists over time. For example, a replication study may yield more information regarding the overlap between age

groups, mean scores, and the correlation between various ages. A comparison of TEEM scores between normal children and language-delayed or language-disordered children would yield further information.

One of the factors which may have significantly altered the outcome of this study was the ceiling established by the TOLD-P. In this investigation, all of the items of the TEEM and the BLST were administered, yet a ceiling of five incorrect responses was used with the TOLD-P. Perhaps more useful and more accurate information might be obtained if all items were given on all three instruments and the raw scores were correlated.

Another aspect of the TEEM for further study would be a detailed item analysis for significant items which may differentiate age groups. According to the primary author, Kenneth Shipley, such a study has not been considered (Shipley, 1985).

Finally, as noted previously, the construct validity of an instrument requires the gradual accumulation of information from a number of sources. A comparison of the TEEM with tests of expressive morphology other than the BLST and/or TOLD-P is an area worthy of further research.

BIBLIOGRAPHY

- BAKER, H., and LELAND, B. (1967). Detroit Tests of Learning Aptitude. Indianapolis, Indiana: Test Division of Bobbs-Merrill.
- BANKSON, N. (1977). Bankson Language Screening Test. Baltimore, Maryland: University Park Press.
- BATES, E. (1976). Pragmatics and sociolinguistics in child language. In M. Morehead and A. Morehead (Eds.), Language deficiency in children: Selected readings. Baltimore, Maryland: University Park Press.
- BERKO, J. (1958). The child's learning of English morphology. WORD, 14, 150-177.
- BERRY, M. (1966). The Berry-Talbott Language Tests: Comprehension of Grammar. Rockford, Illinois: M. Berry.
- BLAXLEY, L., CLINKER, M., and WARR-LEEPER, G. (1983). Two language screening tests compared with developmental sentence scoring. Language, Speech, and Hearing Services in Schools, 14, 38-46.
- BOEHM, A. (1971). Boehm Test of Basic Concepts. New York, New York: The Psychological Corporation.
- BROWN, R. (1973). A first language: The early stages. Cambridge, Massachusetts: Harvard University Press.
- CARROW, E. (1973). Test for Auditory Comprehension of Language. Austin, Texas: Learning Concepts.
- CARROW-WOOLFOLK, E., and LYNCH, J. (1982). An integrative approach to language disorders in children. New York, New York: Grune and Stratton, Inc.
- DALE, P. (1976). Language development: Structure and function (2nd ed.) New York, New York: Holt, Rinehart and Winston.

- DEVER, R. (1972). A comparison of the results of a revised version of Berko's Test of Morphology with the free speech of mentally retarded children. Journal of Speech and Hearing Research, 15, 169-178.
- de VILLIERS, J., and de VILLIERS, P. (1973). A cross-sectional study of the acquisition of grammatical morphemes in child speech. Journal of Psycholinguistic Research, 2, 267-278.
- DUBLINSKE, S., and HEALEY, W. (1978). Special reports: PL 94-142: questions and answers for the speech-language pathologist and audiologist. ASHA, 20, 188-205.
- DUNN, L. (1965). The Peabody Picture Vocabulary Test. Circle Pines, Minnesota: American Guidance Service.
- FLUHARTY, N. (1978). Fluharty Preschool Speech and Language Screening Test. Boston, Massachusetts: Teaching Resources Corporation.
- KIRK, S., MCCARTHY, J., and KIRK, W. (1968). Illinois Test of Psycholinguistic Abilities. Urbana, Illinois: The University of Illinois Press.
- LEE, L. (1971). Northwestern Syntax Screening Test. Evanston, Illinois: Northwestern University Press.
- LEE, L. (1974). Developmental sentence analysis. Evanston, Illinois: Northwestern University Press.
- MCLOUGHLIN, C., and GULLO, D. (1984). Comparison of three formal methods of preschool language assessment. Language, Speech, and Hearing Services in Schools, 15, 146-153.
- NEIDECKER, E. (1980). School programs in speech-language: Organization and management. Englewood Cliffs, New Jersey: Prentice-Hall, Inc.
- NEWCOMER, P., and HAMMILL, D. (1977). The Test of Language Development. Austin, Texas: Empire Press.
- NEWCOMER, P., and HAMMILL, D. (1982). Test of Language Development: Primary (2nd ed.). Austin, Texas: Pro-Ed.

- NEWFIELD, M., and SCHLANGER, B. (1968). The acquisition of English morphology by normal and educable mentally retarded children. Journal of Speech and Hearing Research, 11, 693-706.
- PETERSON, H., and MARQUARDT, T. (1981). Appraisal and diagnosis of speech and language disorders. Englewood Cliffs, New Jersey: Prentice-Hall, Inc.
- SHIPLEY, K., STONE, T., and SUE, M. (1983). Test for Examining Expressive Morphology. Tucson, Arizona: Communication Skill Builders, Inc.
- SHIPLEY, K. (1985). Personal communication, January 31.
- STONE, C. (1950). Progress in primary reading. St. Louis, Missouri: Webster Publishing Company.
- STONE, T. (1980). A method of examining expressive morphology. Unpublished master's thesis, University of Nevada-Reno.
- SUE, M. (1981). Standardization of the Test for Examining Expressive Morphology. Unpublished master's thesis, California State University.
- TEMPLIN, M., and DARLEY, F. (1980). The Templin-Darley Tests of Articulation. Iowa City, Iowa: Bureau of Educational Research and Service, University of Iowa.
- VAN RIPER, C. (1978). Speech correction: Principles and methods (6th ed.). Englewood Cliffs, New Jersey: Prentice-Hall, Inc.
- VOGEL, S. (1977). Morphological ability in normal and dyslexic children. Journal of Learning Disabilities, 10, 35-43.
- WEBER, R. (1982). A closer look at three measures of English morphology. Journal of Learning Disabilities, 15, 86-89.
- WECHSLER, D. (1974). Wechsler Intelligence Scale for Children-Revised. New York, New York: Psychological Corporation.
- WEISS, C., and LILLYWHITE, H. (1976). Communicative disorders: Prevention of communicative disorders. St. Louis, Missouri: C. V. Mosby Co.

- WEISS, C., LILLYWHITE, H., and GORDON, M. (1980). Clinical management of articulation disorders. St. Louis, Missouri: C. V. Mosby Co.
- WEPMAN, J. (1968). Auditory Discrimination Test. Chicago, Illinois: Language Research Associates.
- WIENER, F., LEWNAU, L., and ERWAY, E. (1983). Measuring language competency in speakers of Black American English. Journal of Speech and Hearing Disorders, 48, 76-84.
- WIIG, E., and SEMEL, E. (1984). Language assessment and intervention for the learning disabled (2nd ed.). Columbus, Ohio: Bell and Howell Company.

APPENDIX A

PERMISSION LETTER

Dear Parent:

I am a graduate student at Portland State University, and I am conducting a study regarding language development in children. I am attempting to find out the accuracy of a new test available to speech-language pathologists. To do this, I need children who are developing language normally.

This study can be accomplished by administration of a brief hearing test, followed by three short evaluation instruments. The evaluation will take about 20 minutes of your son/daughter's time, and your child will be asked to look at pictures and answer some questions.

This evaluation will be done by myself, and only appropriate school personnel will have access to any results. Your son/daughter's name will not be used in reporting the results of this study. He/she may be excused from participation at any time.

Please indicate your approval by signing below, and return with your child to school tomorrow.

It is important I schedule the children as soon as possible. Your cooperation is greatly appreciated.

Kathy Zuehlsdorff
 Graduate Student
 Speech and Hearing Sciences
 Portland State University

Child's Name: _____

Child's Birthdate: _____

Does your child have a history of ear infections? _____

If yes, for how long? _____

Parent Signature: _____

Date: _____

APPENDIX B

TEEM SCORE FORM

Items		75%*	90%*	Items		75%*	90%*
1. blocks _____	/s/	3-6	3-6	28. best _____	(irr)	7-12+	7-12+
2. brushes _____	/əz/	4-12	4-12	29. dropped _____	/t/	4-12	5-6
3. witch's _____	/əz/	4-12	5-6	30. shorter _____	/ər/	5-6	7-6
4. feet _____	(irr)	7-12	7-12+	31. shortest _____	/əst/	7-6	7-12+
5. cakes _____	/s/	3-6	3-6	32. pushes _____	/əz/	5-6	6-6
6. eats _____	/s/	3-12	4-6	33. crawling _____	/əŋ/	3-6	3-6
7. climbs _____	/z/	4-6	5-6	34. roped _____	/t/	3-6	5-6
8. rabbit's _____	/s/	3-12	4-6	35. drank _____	(irr)	7-12+	7-12+
9. nurse's _____	/əz/	4-12	6-12	36. drives _____	/z/	5-6	5-12
10. leaves _____	/vz/	7-12+	7-12+	37. watches _____	/əz/	3-12	4-6
11. combed _____	/d/	4-6	5-12	38. reading _____	/əŋ/	3-6	3-6
12. melted _____	/əd/	6-12	7-12	39. cowboy's _____	/z/	3-12	5-6
13. drawn _____	(irr)	7-12+	7-12+	40. flowers _____	/z/	3-6	3-6
14. smaller _____	/ər/	5-6	7-6	41. mouse's _____	/əz/	4-6	6-6
15. smallest _____	/əst/	7-12+	7-12+	42. planted _____	/əd/	6-12	6-12
16. cat's _____	/s/	3-6	3-12	43. caught _____	(irr)	7-12+	7-12+
17. emptied _____	/d/	5-12	6-6	44. teeth _____	(irr)	6-12	7-6
18. cut _____	(irr)	7-12+	7-12+	45. bigger _____	/ər/	5-6	7-6
19. writing _____	/əŋ/	3-6	4-6	46. biggest _____	/əst/	5-6	7-12
20. dogs _____	/z/	3-6	3-6	47. walks _____	/s/	4-6	5-12
21. monkey's _____	/z/	3-6	3-6	48. calves _____	/vz/	7-12+	7-12+
22. washes _____	/əz/	4-6	5-12	49. fed _____	(irr)	7-12	7-12+
23. houses _____	/əz/	3-12	4-6	50. counted _____	/əd/	5-6	7-6
24. children _____	(irr)	7-12+	7-12+	51. swinging _____	/əŋ/	3-6	3-12
25. longer _____	/ər/	6-6	7-6	52. dances _____	/əz/	5-6	6-12
26. longest _____	/əst/	6-6	7-12	53. knives _____	/vz/	7-12+	7-12+
27. better _____	(irr)	7-12+	7-12+	54. women _____	(irr)	7-12	7-12+

*The age by which 75% and 90% of children respond correctly to the test item.

NOTE: 7-12+ is used because no children were tested who were eight years or older.

APPENDIX C

BLST SCORE FORM

Part Two: MORPHOLOGICAL RULES

I. Pronouns Plate 8

73. her _____
74. them; both of them _____
75. him _____
76. she _____
77. they _____
78. he _____
79. hers _____
81. his _____

J. Verb Tenses Plate 9

82. is running _____
83. is reading _____
84. is swimming _____
85. runs _____
86. reads _____
87. swims _____
88. smiled _____
89. climbed, or picked
the flowers _____
90. will fall _____

K. Plurals/Comparatives/Superlatives
Plates 12, 13, 14, & 15

91. books _____
92. pennies _____
93. boxes _____
94. children _____
95. men _____
96. bigger _____
97. biggest _____
98. better _____
99. best _____

APPENDIX D

TOLD-P SCORE FORM

GRAMMATIC COMPLETION	Discontinue after 5 consecutive failures	Score 1 or 0
1. Mary has a dress and Joan has a dress. They have two (dresses). _____		1. _____
2. Joey likes to play. Right now he is (playing). _____		2. _____
3. The shoes belong to the boy. Whose shoes are they? They are the (boy's). _____		3. _____
4. Betty likes to swim everyday. Today she is (swimming). _____		4. _____
5. A lady likes to drive. Everyday she (drives). _____		5. _____
6. A boy likes to ride his bicycle everyday. Today he is (riding). _____		6. _____
7. The toys belong to the children. Whose toys are they? They are the (children's). _____		7. _____
8. A girl plays the piano everyday. Yesterday she (played). _____		8. _____
9. The hat belongs to mother. Whose hat is it? It is (mother's). _____		9. _____
10. The dress belongs to the woman. Whose dress is it? It is the (woman's). _____		10. _____
11. A person who sings is a (singer). _____		11. _____
12. Betty likes to eat cookies. Everyday she (eats). _____		12. _____
13. John likes to cook everyday. Yesterday he (cooked). _____		13. _____
14. Jane likes to jump. Now she is (jumping). _____		14. _____
15. A cake might be small, but a cupcake is (smaller). _____		15. _____
16. A person who paints fences is a (painter). _____		16. _____
17. A dog can be big, but a horse is (bigger). _____		17. _____
18. A person who plays a drum is a (drummer). _____		18. _____
19. Joe had a gumdrop, and Sue had a handful of gumdrops; but Tom had a bagfull so he had the (most). _____		19. _____
20. Bob is a man. Bill is a man. Bob and Bill are two (men). _____		20. _____
21. A cake might be small, and a cupcake smaller, but a cookie is the (smallest). _____		21. _____
22. John likes to throw the ball everyday. Yesterday he (threw). _____		22. _____
23. Today I found a leaf. Yesterday I found two (leaves). _____		23. _____
24. A boy likes to ride his bicycle everyday. Yesterday he (rode). _____		24. _____
25. A spoonful of ice cream is good, two spoonfuls are better, and a dishful is (best). _____		25. _____
26. Joe had one gumdrop. Sue had a handful of gumdrops, so she had (more). _____		26. _____
27. Mary is a woman. Joan is a woman. Mary and Joan are two (women). _____		27. _____
28. Betty likes to draw everyday. Yesterday she (drew). _____		28. _____
29. I have a mouse. She has a mouse. We have two (mice). _____		29. _____
30. Jeff ate the candy quickly; and when Bill came, it had all been (eaten). _____		30. _____

No. of 1s _____
 No. of 0s _____
 Total _____ (30) _____

APPENDIX E

RAW TEST SCORES

	PRESCHOOL				KINDERGARTEN				GRADE 1			
	Age in Months	TEEM	BLST	TOLD-P	Age in Months	TEEM	BLST	TOLD-P	Age in Months	TEEM	BLST	TOLD-P
1.	61	43	22	26	73	44	23	23	87	51	23	24
2.	63	46	23	25	75	45	18	20	87	45	24	18
3.	61	34	13	12	70	42	23	21	81	47	27	26
4.	61	39	20	19	76	38	23	19	85	40	22	17
5.	64	42	19	19	71	45	25	15	81	44	25	22
6.	64	42	21	20	70	42	23	18	83	43	22	25
7.	61	41	22	18	72	48	23	26	83	47	24	28
8.	58	43	23	22	75	45	20	18	82	40	22	18
9.	63	42	20	18	71	44	22	19	85	46	22	24
10.	64	43	24	22	70	46	20	19	82	45	25	24
11.	61	42	21	21	70	45	22	21	84	47	22	24
12.	58	43	22	23	71	41	20	19	81	45	22	21
13.	64	40	17	11	71	43	23	21	84	49	25	25
14.	62	43	21	18	73	40	20	16	86	48	26	24
15.	59	36	20	17	70	30	15	13	84	45	22	24
16.	64	41	19	18	70	41	21	18	82	50	24	26
17.	63	40	20	18	74	42	19	17	87	51	26	24
18.	58	32	21	17	75	45	24	20	81	42	22	21
19.	63	43	23	20	73	48	22	22	87	51	23	23
20.	60	48	25	23	70	43	23	20	82	45	25	26
21.	58	38	19	18	70	46	24	19	82	46	26	23
22.	62	43	23	19	70	48	24	21	87	44	22	21
23.	63	43	15	15	76	48	22	17	83	47	26	25
24.	58	44	25	25	76	38	19	13	87	44	23	20