Dissertations and Theses @ UNI   Student Work

1985

The Development and Initial Validation of a Test of Students' Use of Morphemic Analysis in Deriving Word Meaning

Maelou Kelly Baxter

Let us know how access to this document benefits you

THE DEVELOPMENT AND INITIAL VALIDATION OF A TEST

OF STUDENTS' USE OF MORPHEMIC ANALYSIS

IN DERIVING WORD MEANING

An Abstract of a Thesis

Submitted

In Partial Fulfillment

of the Requirements for the Degree

Master of Arts in Education

Maelou Kelly Baxter

University of Northern Iowa

August 1985

ABSTRACT

Baxter, Maelou Kelly. M.A. in Education, University of Northern Iowa, August 1985. THE DEVELOPMENT AND INITIAL VALIDATION OF A TEST OF STUDENTS' USE OF MORPHEMIC ANALYSIS IN DERIVING WORD MEANING.

The purpose of this study was to develop and initially validate a test of students' use of morphemic analysis in deriving word meaning. Eight steps in test construction were followed: (a) definition of test objectives, (b) definition of target population, (c) review of related measures, (d) development of an item pool, (e) preparation of a prototype, (f) evaluation of the prototype, (g) revision of the measure, and (h) collection of data on test validity and reliability.

The field testing was conducted with 15 fifth-grade and 15 sixth-grade readers who lived in a midwestern town with a population of about 35,000. The subjects were average readers according to their performance on a standardized reading assessment.

The morphemic analysis test consisted of two subtests of 12 words each that were matched according to frequency, imagery, length, number of syllables, and pronounceability. The subtests varied according to morphemic composition. One subtest contained monomorphemic words (e.g., avalanche), and the other contained suffixed words (e.g., betrayal). Each word was placed in a sentence context. A 3-point system was used for scoring subjects' free-response definitions of the target words. Analyses of the

data revealed that while some aspects of the test were satisfactory, the test as a whole could not be recommended for use in its present form. Suggestions for modifications were presented.

THE DEVELOPMENT AND INITIAL VALIDATION OF A TEST

OF STUDENTS' USE OF MORPHEMIC ANALYSIS

IN DERIVING WORD MEANING

A Thesis

Submitted

In Partial Fulfillment

of the Requirements for the Degree

Master of Arts in Education

Maelou Kelly Baxter

University of Northern Iowa

August 1985

This Study by: Maelou Kelly Baxter

Entitled: THE DEVELOPMENT AND INITIAL VALIDATION OF A TEST OF STUDENTS'

USE OF MORPHEMIC ANALYSIS IN DERIVING WORD MEANING

has been approved as meeting the thesis requirement for the

Degree of Master of Arts in Education

David W. Moore

7/8/85

Date

Chairman, Thesis Committee

Ned Ratekin

7-8-85

Date

Member, Thesis Committee

Marlene Strathe

7-8-85

Date

Member, Thesis Committee

John C. Downey

7-25-85

Date

Dean of the Graduate College

ACKNOWLEDGEMENTS

TABLE OF CONTENTS

LIST OF TABLES

CHAPTER 1

THE PROBLEM

Introduction

Estimates of the number of words in the English language vary substantially. This variation is caused largely by differences in the definition of the term _word_. One source (McWhirter, 1983) gives English credit for having the largest vocabulary of any language, about 490,000 general words plus 300,000 technical terms. Of course, no one person has in his or her personal lexicon 790,000 words. In fact, even if the quantity is reduced to the number of words in printed school English, which is estimated by Nagy and Anderson (1984) to be about 88,500, the task of acquiring an adequate vocabulary still seems insurmountable.

How an adequate vocabulary is acquired has been the subject of countless studies, books, chapters, and articles. Teaching students all the words they need to know is an impossible task. There are too many words; there is too little time. Consequently, strategies need to be identified which will increase students' ability to learn words on their own.

Predominant strategies for vocabulary instruction may be divided into three general categories. These categories are direct teaching of meanings from word lists, use of context, and morphemic analysis (Petty, 1968).

Many strategies consist of some variation of the direct study of words and their definitions and usually include dictionary usage. No matter how creatively these are handled, they are still inefficient ways of acquiring meaning vocabulary. For each item studied, only one is accessed. Chances are good that very few of the words studied will be remembered and that only a small percentage of students will develop the dictionary habit.

A second broad category of vocabulary instruction is use of context clues. Using syntactic and semantic context clues has been demonstrated to aid students in determining word meaning during the act of reading (Amoriell & Hofler, 1984). Use of context, however, is more useful in improving short-term reading comprehension than in acquiring long-term meaning vocabulary. Nagy and Anderson (1984) report that "Experimental studies have often indicated that children do not seem to learn word meanings very well from context" (p. 327).

A third widely accepted way of helping students to acquire meaning vocabulary is through morphemic analysis, or inferring meaning from word parts. The term structural analysis is sometimes incorrectly used as a synonym for morphemic analysis. Structural analysis, however, is a technique for breaking words into pronunciation units (Harris & Hodges, 1981). These units would include syllables which have no meaning when they stand alone. In morphemic analysis, students learn common roots and affixes which help them to deduce the meanings of words containing

these components. Educators have traditionally accepted the value of teaching morphemic analysis, but there has been little research in this area (Moore, 1980).

In 1940, Carroll noted that the correlation of morphemic analysis with vocabulary development had not been adequately investigated. Fifteen years later Otterman (1955) called research in the area of teaching prefixes and word roots "scanty" and the studies "not consistent in their findings" (p. 611). And yet, 44 years after Carroll's lament, the subject still has not been adequately investigated. "There is basically no experimental literature that confirms the success of any of these [motivation, morphology, context] in facilitating children's learning of words on their own" (Nagy & Anderson, 1984, p. 325).

Nevertheless, there appears to be good reason for helping students attend to the morphemic composition of words. Nagy and Anderson (1984) reported that there are an estimated 139,020 words whose meanings can be directly inferred when there is knowledge of the root. For example, a person who knows the meaning of _agree_ will probably be able to infer the meaning of _disagree_ even though he or she has never seen it before. The sheer volume of such words makes a strong case for the teaching of any method which will aid students in learning them.

Despite the traditional acceptance by educators of the fact that morphemic analysis is valuable, despite the number of words whose meanings can be inferred through its application, and

despite its wide-spread use as a strategy for teaching meaning
vocabulary, tests which assess students' use of morphemic analysis
in deriving word meaning are inadequate. Two of the three reading
tests which specifically use the term morphology involve the use
of isolated word parts. The Morpheme Recognition Test, developed
by Carroll in 1940, assessed students' knowledge of word elements.
The Morpheme Knowledge Test (Shepherd, 1973) was designed as a
test of knowledge of Latin roots and prefixes. Both of these
tests are somewhat artificial and measure knowledge of morphemes
in relative isolation. A third test, designed by Moore (1980),
assessed the effect of morphemic composition on the ability of
students to pronounce words. Although this test is more natural
than the others because the words are presented in sentences, it
yields no information on students' use of morphemes to obtain
meaning. No test with acceptable validity has been discovered
which measures students' use of morphemic composition in deriving
word meaning. Therefore, the development of such a test was the
primary purpose of this study.

Statement of the Problem

There seems to be a general consensus that students should be
taught how to derive word meaning from morphemes. However, there
is no test with acceptable validity which measures this
proficiency.

In order to develop such a diagnostic instrument, the eight-
stage process described by Borg and Gall (1983) was followed.

This process involved (a) definition of objectives, (b) definition of the target population, (c) review of related measures, (d) development of an item pool, (e) preparation of a prototype, (f) evaluation of the prototype, (g) revision of the measure, and (h) collection of data on test validity and reliability.

In the development and initial validation of the test, the following questions were addressed:

1. To what extent is this test a valid measure of students' use of morphemic analysis in deriving word meaning?

2. To what extent is the test a reliable measure of students' use of morphemic analysis in deriving word meaning?

3. What is the normative performance of a group of students with the test?

Importance of the Study

Students are faced with the enormous task of acquiring a sufficiently large meaning vocabulary to enable them to meet the reading demands of their schools and their society. Since it is clearly impossible for teachers to teach enough words directly or for students to memorize enough words and definitions to meet these demands, more efficient methods must be used. One such method seems to be morphemic analysis, a system which can be used to discover the meanings of large numbers of unfamiliar words which contain familiar word parts. At present, validated procedures to assess a student's strategy for applying morphological knowledge when inferring word meanings are not

available. The information from this study may be useful in validating such a procedure.

Assumptions

It is assumed that the scoring system used in this study, which is an adaptation of that used for the WISC-R vocabulary subtest (Wechsler, 1974), assesses word meaning.

Limitations of the Study

One limitation of this study is that the subjects were drawn from one school system. This may reduce the generalizability of the study in terms of the target population. Another limitation is that only nominalizations were used to assess students' use of morphemic analysis. This may reduce the generalizability of the study in terms of the reading process being assessed.

Definition of Terms

For purposes of this study, the following definitions will be used:

Base Word: A word to which affixes are added to create related words (Harris & Hodges, 1981). Base words are frequently referred to as root words or stems. All base words used in this study are free morphemes, or morphemes which can stand alone as independent words.

Monomorphemic Words: Words consisting of a single morpheme (Moore, 1980).

Morpheme: A meaningful linguistic unit which cannot be divided into smaller meaningful units (Harris & Hodges, 1981).

Morphemic Analysis: The process of determining the meaning of an unfamiliar printed word by examining its meaningful parts or morphemes (Johnson & Pearson, 1978). In this study morphemic analysis refers specifically to deriving meaning of nominalizations.

Suffixed Word: A derived word formation obtained by the addition at the end of a word of any affix other than an inflection (Moore, 1980).

CHAPTER 2

REVIEW OF RELATED LITERATURE

This review of related literature is important to this study

primarily for the purpose of reporting what was not found in the

literature. Research reveals that morphemic analysis is generally

considered to be one of the primary methods of teaching

vocabulary. However, research supporting this consideration is

sparse, and methods of testing students' application of such

instruction are inadequate.

This chapter first reviews the literature which is concerned

with the theoretical role of morphemic analysis in determining

word meaning. Next, the literature relative to data-based

research on the role of morphemic analysis in determining word

meaning is reviewed. The final section deals with tests of

students' use of morphemic analysis to determine word meaning.

The Theoretical Role of Morphemic Analysis

in Deriving Word Meaning

There has long been interest in the teaching of morphemic

analysis as an aid to increasing students' meaning vocabularies.

Many reading educators have recommended this strategy. This

section reviews theoretical literature which has presented the

role of morphemic analysis in obtaining word meaning.

A few limitations regarding the value of teaching word

morphology have been presented. Dolch (1955) advocated letting

children discover affixes for themselves and cautioned the teacher

not to teach those which might be beyond the students'

understanding. Indeed, since three out of four words used in

elementary school do not have prefixes, Dolch advised against

depending on them too much. Schell (1967) stated that prefixes

have meaning to the elementary school child only if the base word

is known. Both writers, however, favored the study of morphemic

analysis as long as their cautions were heeded.

A major point in favor of the study of morphemic analysis as

opposed to the memorizing of lists of individual words is that it

gives students a tool which they can use on their own to learn new

words as they encounter them. Nagy and Anderson (1984) studied a

number of words in printed school English. Their results

convinced them that the staggering number of words encountered by

students indicated that vocabulary instruction strategies that

prepare students to learn words on their own are vital. They

suggested inferring meaning from word parts (morphemic analysis)

as an important way of accomplishing this end. They also believed

that the study of morphology has the added advantage of calling

students' attention to word formation processes so that they will

be more likely to take advantage of such processes when they are

working independently.

Amoriell and Hofler (1984) believed that the developing, as

well as the mature, reader needs to be able to derive the meaning

of new words and thus be released from total reliance on the

dictionary. They recommend the study of morphological units for this purpose.

The writers of textbooks on the teaching of vocabulary seem to be in agreement on the value of teaching morphemic analysis. In their classic text on the teaching of vocabulary, Dale and O'Rourke (1971) called morphemic analysis one of the key techniques offered in their book. They devoted four complete chapters to various aspects of the subject and incorporated it in several other chapters.

In another text, Deighton (1959) stated that the "prime goal of vocabulary development" (p. 56) is to equip students to get the meaning of unfamiliar words as they encounter them. He suggested that training students to get meaning from word parts is the only method which allows them to determine meaning of unfamiliar words both in context and in isolation.

In their text devoted to the teaching of reading vocabulary, Johnson and Pearson (1978) stated:

The beauty and advantage of morphemic analysis lies in the relatively high degree of consistency between letters, sounds, and meanings when words are analyzed at the morphemic level. . . . Hence, analysis of words at the level of the morpheme is probably more rational and intuitively sensible than at the level of relating letters to sounds. (p. 83)

Dale and Milligan (1970) added yet another dimension to the argument for teaching morphemic analysis. They agreed that a

developed awareness of morphological units equips students with skills that may enable them to determine meanings of new words, but that the most important advantage of such an awareness is that it allows students to discover that the study of words can be exciting.

Some writers have been concerned with slightly more specialized applications of morphemic analysis. Students' study of roots and affixes has the effect of expanding and elaborating their vocabulary and spelling knowledge in what Templeton (1978) called the "spelling/meaning connection." Writing about vocabulary development in the content areas, Burmeister (1976) asserted that the area of morphology is of extreme importance and that its rewards in word power "multiply in geometric progression" (p. 482).

In conclusion, it should be noted that there are many other sources which could be cited in support of the value of teaching morphemic analysis. Morphemic analysis is included as a strategy for vocabulary study in the majority of reading and language arts textbooks and vocabulary texts and study books. The primary rationale given by writers for the teaching of morphemic analysis is that it provides a powerful tool for students to use on their own in determining meaning for unfamiliar words.

Data-based Research on the Role of Morphemic

Analysis in Deriving Word Meaning

Despite the general acceptance by educators of the value of teaching morphemic analysis, very little data-based research in this area is available. What research there is, however, is mostly favorable.

Stauffer (1942) conducted a study to discover which ones of the prefixes listed in Thorndike's word book (Thorndike, 1921) should be taught in elementary schools. He based his conclusions on which prefixes occurred most frequently and were therefore most useful. He found that 15 prefixes accounted for 82% of the total number of prefixes which occur in Thorndike's list. In addition to compiling the list, Stauffer reached the conclusion that "A mastery over prefixes is essential for meeting the language needs of the students" (p. 456).

Shepherd (1975), basing his study in part on Stauffer's work, surveyed the literature to consider the validity of the assumption that the meanings of morphemes taught in schools were related to the present-day meaning of English words. He concluded that knowledge of Latin roots and the affixes which combined with them was strongly related to knowledge of their derivatives, and that the same was true of English word bases and their affixes. At about the same time, Groff (1973) surveyed 40 leading writers on reading and spelling and found that the majority endorsed teaching the meaning of affixes.

This review of literature revealed only two studies, widely spaced in years, which indicated serious doubts about the value of teaching morphemic analysis. Barnes (1942) conducted a study with 75 college freshmen in which the subjects were presented with a difficult word and the meanings of its parts and asked to write a sentence defining the word. The results led him to believe that a knowledge of meanings of word parts alone was not of great value in determining meanings of English words. Otterman (1955) conducted a much-cited study with seventh-grade students in which she found that the systematic study of prefixes and word roots only helped students know isolated parts. She reported that such study resulted in no significant vocabulary improvement.

In contrast, other research has affirmed the traditionally held importance of morphemic analysis. Tate (1936) indicated that instruction in morphemic analysis resulted in tremendous gains in reading vocabulary for sixth- and eighth-grade pupils, and that teacher observations indicated that pupils were keenly interested in root words and meanings. Hunt (1953) investigated the relationship among vocabulary, morphemic analysis, and reading at the college level. He stated that though there had been few research studies concerned with the study of morphemic analysis and word elements, he believed that such study was an important vocabulary development method. In his research, he found that the ability to use morphemic analysis was more than a matter of general intellectual ability and tended to be related to reading

and vocabulary even when the effect of intelligence was controlled. Also working at the college level, Thompson (1958) reported on his use of Brown's "Master Word" approach to teaching vocabulary. Brown's method relied on the teaching of 14 master words which contained some of the most frequently used roots and affixes. Thompson reported a 20% gain between pretest and posttest scores in the portion of the tests devoted to word definition.

In conclusion, data-based research on the role of morphemic analysis in determining word meaning is scarce and the results are somewhat contradictory. However, it appears that most of the little research which has been done is favorable to morphemic analysis.

<u>Tests of Students' Use of Morphemic</u>

<u>Analysis in Deriving Word Meaning</u>

Tests which purport to measure students' use of morphemic analysis fall into two groups. These are vocabulary subtests of general reading tests and tests which pertain only to morphemes.

Although they do not use the terms "morphemic analysis" or "morphology," some reading tests have subtests which measure some aspect of the subject. Two of these are the Stanford Diagnostic Reading Test (Karlsen, 1976) and the McCulloch Word Analysis Tests (McCulloch, 1963).

The Stanford Diagnostic Reading Test states that Part A of its subtest requires pupils to identify the first syllable in

two-syllable words (Karlsen, 1976). A sample item lists the word

sometimes with the choices so, some, and somet, and instructs

students to select the first syllable of the key word. Part B

assesses students' ability to divide words into meaningful

elements and to blend word elements together. A sample given in

this section lists the words air, one, and plane and instructs

students to choose two word parts which make a word, noting that

they are in correct order form left to right.

A McCulloch Word Analysis Tests subtest tells students "This

is a test of your ability to see the prefix or suffix of a word as

separate from its root" (McCullough, 1960, p. 8). An example

lists the words thankful, darkness, and unbend. The students are

instructed to draw a circle around the prefix or suffix in each

word.

Neither of these two tests is concerned with assessing

students' use of morphemic analysis in deriving word meaning.

They involve purely mechanical division and blending of word parts

without any consideration of the meaning of the whole words. The

same is true of three tests which pertain only to morphology.

Carroll's Morpheme Recognition Test (1940) is the earliest of

the tests which are concerned solely with morphology. In that

test, students are asked to mark words which have a common element

of meaning and then to choose the meaning for that element. A

sample item includes the words ready, read, regression, region,

repeat, return, and rectangle. The students are to choose the

correct meaning for the common element (re-) from this list: writing, back or again, true, and very. Carroll's test is designed to measure knowledge of the meaning of individual morphemes, not of words.

Shepherd's Morpheme Knowledge Test (1973) was designed as a test of knowledge of Latin roots and prefixes. Students are presented with a morpheme and three words containing that morpheme. For example, one item on Shepherd's test instructs students to select the best meaning for "re-." Three words using "re" are listed, REceive, REduce, and REject. The meanings provided for student selection are (a) back or again, (b) overly or above, (c) not or opposite, and (d) across or beyond. Shepherd's test, like Carroll's, measures only knowledge of the meaning of individual morphemes.

Moore's (1980) test was designed to assess the ability of students to pronounce words. In it, words classified as either monomorphemic or suffixed were presented to students in sentence contexts, and the students were asked to pronounce the words. No attempt was made to assess students' knowledge of the meanings of the words.

At present, there appear to be very few tests which assess any aspect of word morphology. In an examination of test review sources, Buros (1975) was found to list no tests relating to morphemes, and Fagan (1975) listed only Shepherd's Morpheme Knowledge Test. There appear to be none with acceptable validity

that measure students' use of morphemic analysis in deriving word meaning.

Summary

This chapter reviewed the professional and research literature dealing with the effect of morphemic composition of words on students' knowledge of word meaning. Also reviewed was the literature dealing with tests of morphemic knowledge.

The review revealed that many professional sources support the teaching of morphemic analysis as an aid to the acquisition of meaning vocabulary. Data-based research on this subject is sparse. Most researchers have concluded that the study of morphemic analysis is valuable, although much of the research is flawed when evaluated with contemporary standards.

It was found that some general reading tests have subtests which measure some aspect of morphology, and three tests were located which dealt solely with morphology. However, there appears to be no diagnostic instrument which provides a valid test of students' use of morphemic analysis in deriving word meaning.

CHAPTER 3

DESIGN OF THE STUDY

This study was an extension of Moore's (1980) study in which

he assessed the effect of morphemic composition on students' word

identification. In that study, students' word identification

performance was compared between suffixed and monomorphemic words.

In the current study, a diagnostic test was developed and

initially validated which measured students' use of morphemic

analysis in deriving word meaning. The construction and initial

validation followed the eight steps in test construction outlined

by Borg and Gall (1983). These steps are (a) definition of test

objectives, (b) definition of target population, (c) review of

related measures, (d) development of an item pool, (e) preparation

of a prototype, (f) evaluation of the prototype, (g) revision of

the measure, and (h) collection of data on test validity and

reliability.

Definition of Test Objectives

The objective of this test was to determine students' use of

morphemic analysis in deriving word meaning. A student who uses

morphemic analysis generates or retains acceptable meanings for

derived words (e.g., betrayal) by attending to base words and

suffixes.

Definition of Target Population

Because extensive instruction in morphemic analysis most

commonly begins about third grade, it was determined that the most

useful level for a test of students' use of morphemic analysis in
determining word meaning would be in the upper elementary grades.
Therefore, the test was designed to be used primarily in grades
four through six. The sample used in the pilot test consisted of
5 fourth-grade and 5 sixth-grade students. The field testing was
done with 15 fifth- and 15 sixth-grade students. All students
were of average reading ability from an elementary school in a
midwestern urban school district. For the purpose of this study,
average reading ability was defined as scoring in the 30th
percentile to the 70th percentile on the Iowa Tests of Basic
Skills (ITBS) (1978). The average percentile ranks on the ITBS
were 57.13 for the fifth graders, 56.00 for the sixth graders, and
56.57 for the group as a whole.

Review of Related Measures

The review of related measures was reported in Chapter 2.
Only a few tests were located which purported to measure any
aspect of morphology. To date, no instrument has been discovered
which satisfies the objectives of the test which was developed as
the purpose of this study.

Development of an Item Pool

The items used for this test were adapted from those used in
the study by Moore (1980). Moore selected words from a list of
925 nouns rated for imagery by Paivio, Yuille, and Madigan (1968).
In making his selection, Moore first identified all monomorphemic
and suffixed words. Suffixed words were included only if they

were derived from base words which could stand independently as English words.

Following the initial screening, the final list of words was selected by identifying words which were similar in terms of six key properties. In addition to form class, these properties were imagery, frequency, word length, number of syllables, and pronounceability.

The first of these properties, imagery, is defined as the capacity to arouse a sensory experience such as a mental picture or a sound (Moore, 1980). It has been found to affect word learning and processing (Paivio & O'Neill, 1970). The Standard Frequency Index (SFI) reported by Carroll, Davies, and Richman (1971) was the measure of word frequency used by Moore. Word length was determined by the number of letters in the word. The words were also grouped by the number of syllables they contained according to the American Heritage Dictionary of the English Language (Morris, 1969). Because no measure of pronounceability could be located, this property was assessed for each word by a jury of five graduate students in Reading Education. An average pronounceability rating for each word was then computed.

Once pronounceability was determined, the words were grouped into three morphemic categories. These were monomorphemic words, suffixed words with no spelling change in the stem, and suffixed words with a spelling change in the stem. Twelve words in each category were then selected. The current study made use of only

the first two categories, monomorphemic words and suffixed words
with no spelling change in the stem.   The 24 words and their word
property measures are shown in Appendix A. As Appendix A shows,
the two sets of words were matched in terms of the properties
listed above. Thus, the two sets of words would seem to be of
approximately equal difficulty.

After the final selection of word groups, Moore placed the
words in sentence context, and it was in this setting that they
were presented to the subjects. Before use, the sentence contexts
with the key words omitted were administered to a panel to assure
that the contexts provided sufficient contextual information for
each word so that readers could produce syntactically acceptable
words but not the key word. Necessary changes were made until no
exact words were produced for the sentence contexts finally
selected.

Preparation of a Prototype

The test prototype consisted of two parts. The first part
was composed of the 24 sentences to be used in determining
students' knowledge of the meaning of the key words. The second
part was composed of the set of 12 suffixed words and was not a
part of the test itself, but was used for validation purposes.
These words were used to evaluate students' ability to recognize
base words.

Knowledge of Word Meaning

The first part was scored according to a 3-point system. The scoring procedure was adapted from the WISC-R vocabulary subtest (Wechsler, 1974). In this system, responses are scored 2, 1, or 0 depending on their quality. Examples and model answers were written for the 24 words included in this study. Scoring criteria developed for this test are found in Appendix B.

The list of words was compiled by Moore. However, because the current test was intended to measure knowledge of word meaning instead of word identification, the sentence contexts were checked for clues which might lead students to a portion of meaning. After revision, the sentence contexts with the key words omitted were presented to a panel of graduate students to assure that they did not offer undue textual information.

Once the final selection of sentence contexts had been made (see Appendix C), the student form of this part of the test instrument was prepared. It consisted of laminated sentence strips containing the 24 test sentences which were administered in random order to determine students' knowledge of the meaning of the key words.

Knowledge of Base Words

The second part was composed of a sheet containing the 12 suffixed words. There were three forms so that these could also be administered in random order. These sheets were used for the base word circling task which was used to validate the test as a

measure of sensitivity to morphemes. A low moderate relationship

was expected between the students' ability to circle the base word

and their ability to define the derived word. Such a relationship

was expected because identifying base words was thought to be only

one aspect of using morphemic analysis in deriving word meaning

and was not thought to represent the entire process.

In order to standardize testing conditions as much as

possible, instructions to be given to the students by the examiner

were written out (see Appendix D). A copy of each of the

protocols used in this study is presented in Appendix E.

Pilot Test

The final step in preparing the prototype was the pilot

testing of the items, the scoring system, and the procedures with

5 fourth-grade and 5 sixth-grade students. During the pilot test

it became apparent that the words were too difficult for fourth

graders and the decision was made to use fifth and sixth graders

in the field test. Following the pilot testing, other necessary

adjustments were made in the sentence contexts, the scoring

system, and the testing procedures.

Evaluation of the Prototype

Following the pilot testing of the prototype and the making

of necessary revisions in the instrument and the procedures, the

prototype was field tested. Upon completion of the field testing,

the instrument was evaluated for future revision.

For the field test, 30 subjects, 15 fifth graders and 15 sixth graders, were selected from elementary classrooms in the same midwestern school district. The students were pulled from their regular classrooms one at a time. The instructions for taking the test were read to the students by the examiner. Students were instructed to read each sentence either silently or orally as they chose, and then to tell the examiner what the underlined word meant.

When the 24 sentences had been administered in random order, the students were asked to do the circling task. The 12 suffixed words were presented to the students in random order, and they were asked to circle the base word in each one of the key words.

Item Analysis

Students' performance with the individual items of the test were examined. An item-to-test correlation and a difficulty index were computed for each item. The difficulty index was computed by dividing the number of subjects who answered the item correctly by the number of subjects tested.

Revision of the Instrument

Elimination of items would have necessitated repeating the exhaustive process by which Moore selected the words for his study. Such a process was not within the scope of this study.

Collection of Data on Test Validity and Reliability

Two indications of reliability were computed. In addition, two areas of validity for the test were considered, validity of

the instrument as a measure of word meaning and validity as a

measure of sensitivity to morphemic composition.

Reliability of the Test

Two statistical indications of reliability were calculated.

They were reliability coefficients, Cronbach's alpha and KR-21,

and the standard error of measurement.

Validity of the Test as a Measure of Word Meaning

Evidence of the validity of the test as a measure of word

meaning was provided by the use of an analysis of variance to

compare fifth graders as a group and sixth graders as a group in

their overall scores. It was expected that the sixth graders

would score significantly higher than the fifth graders, and thus

demonstrate that the test does in fact measure word meaning. This

was based on the assumption that the sixth graders as a group knew

the meanings of more words than the fifth graders as a group by

virtue of their time spent in school. The ANOVA also compared

performances on the two word types, monomorphemic and suffixed.

Validity of the Test as a Measure of Sensitivity to Morphemic Composition

Evidence of the validity of the test as a measure of

sensitivity to morphemic composition was provided in two areas.

First, the characteristics of the words themselves were carefully

controlled by Moore as discussed in the section on development of

an item pool. As a consequence, test results reflected the

behavior of interest, sensitivity to morphemes, rather than some

other characteristic of the words themselves. Toward the same end, the sentence contexts were made as free of contextual information as possible.

Furthermore, to demonstrate that the test did indeed measure sensitivity to morphemes, a chi-square analysis, followed by the calculation of a phi coefficient, was performed. Students' performance with deriving meaning from each suffixed word (e.g., betrayal) was compared with their performance with identifying the base of each suffixed word (e.g., betray). A significant, positive, low moderate relationship was expected for each word.

Summary

A diagnostic test of students' use of morphemic analysis in determining word meanings was developed and initially validated. This was done by following an eight-stage process described by Borg and Gall (1983).

To test this effect, 12 monomorphemic words and 12 suffixed words were used which had been selected by Moore (1980), who matched the words on six key properties. The words were presented in sentence contexts to 15 fifth-grade and 15 sixth-grade average readers who were asked to give the meanings of the 24 words. Definitions generated by the students were evaluated according to a 3-point system. In the second part of the test, students were given the 12 suffixed words and asked to choose the base word in each one. This task was evaluated according to the the students' ability to circle the base words. Following the collection of the

data, questions of reliability and validity of the test instrument

were considered. Statistical analyses of the findings are

presented in Chapter 4.

CHAPTER 4

FINDINGS

The purpose of this study was to construct and initially
validate a test of students' use of morphemic analysis in deriving
word meaning. This chapter presents the findings from the
analyses of the data.

The analysis of variance computation reported in this chapter
was performed using the Statistical Analysis System (SAS) (Barr,
Goodnight, Sall, & Helwig, 1982). All other computations were
performed with the Statistical Package for the Social Sciences
(SPSS), Version H (Nie & Hall, 1980). The findings are presented
in four categories: (a) descriptive statistics, (b) item
analysis, (c) reliability, and (d) validity.

Descriptive Statistics

Table 1 shows the means and standard deviations of the test
scores. Scores are reported for the total test and are broken
down by grade level and word type. As can be seen, the mean
scores for the 12-word subtests were 10.76 ($\underline{SD}$ = 3.53) for the
monomorphemic words and 8.90 ($\underline{SD}$ = 3.50) for the suffixed words.
Mean scores for the total test by grade level were 17.60 ($\underline{SD}$ =
5.62) for fifth-grade students and 21.73 ($\underline{SD}$ = 6.63) for sixth-
grade students. Data for all students are found in Appendix F.

Table 1

Descriptive Statistics for Test Scores

Grade Level	Word Type					
	Monomorphemic		Suffixed		Total	
(n = 15)	mean[a]	(SD)	mean[a]	(SD)	mean[b]	(SD)
5	9.80	(3.55)	7.80	(2.65)	17.60	(5.62)
6	11.73	(3.35)	10.00	(3.96)	21.73	(6.63)
Total	10.76	(3.53)	8.90	(3.50)		

[a]Maximum score = 24.   [b]Maximum score = 48.

Item Analyses

Item-to-test Correlations

Item-to-test correlations indicate the degree to which each individual item is correlated with overall test scores. These analyses are shown in Table 2. They ranged from a high of .84 for avalanche and tragedy to a low of -.37 for prosperity. The mean item-to-test correlation was a low moderate .40.

Table 2

Item-to-Test Correlations

armadillo	.51
avalanche	.84
derelict	.79
enterprise	.73
episode	.61
fantasy	.50
hurricane	.37
intellect	.00
panorama	.00
semester	.31
tragedy	.84
velocity	.51
betrayal	.51
deduction	.60
engagement	.60
formation	.19
heroism	.70
instructor	.00
mastery	.60
prisoner	.00
prosperity	-.37
settlement	.00
socialist	.00
tendency	.61

Difficulty Index

A difficulty index shows the difficulty of the individual items and can range from .00, which indicates no correct responses, to 1.00 which indicates that all responses were correct. This index is obtained by dividing the number of subjects who answered an item correctly by the total number of subjects tested. As seen in Table 3, item difficulty indexes ranged from .00 for panorama to 1.00 for instructor. These two items do not appear to be of appropriate difficulty for inclusion in future test forms. The mean difficulties were .47 for all words, .47 for the monomorphemic words, and .48 for the suffixed words. The findings indicate that the average difficulty index for this test is satisfactory.

Reliability

Two aspects of the reliability of the instrument, subjects' consistency and scorer's consistency, were assessed. Subjects' consistency was determined by computing Cronbach's alpha, KR-21, and standard errors of measurement. Scorer's consistency was determined by computing the percentage of scoring agreement between the researcher and a second rater.

Subjects' Consistency

Cronbach's alpha, a reliability coefficient, was calculated at a moderate .77 for the entire 24-word test. The KR-21 coefficient for the suffixed words was also moderate, .72, but for the monomorphemic words it was a very high .99.

Table 3

Item Difficulty Indexes

armadillo	.80
avalanche	.60
derelict	.23
enterprise	.70
episode	.76
fantasy	.83
hurricane	.96
intellect	.23
panorama	.00
semester	.73
tragedy	.63
velocity	.10
betrayal	.43
deduction	.43
engagement	.33
formation	.76
heroism	.43
instructor	1.00
mastery	.36
prisoner	.86
prosperity	.06
settlement	.76
socialist	.06
tendency	.20

The standard error of measurement (SEM) was recommended by Williams and Zimmerman (1984) as being relatively independent of the group tested or of the testing situation and as being conceptually easy to understand. The SEM for the entire 24-item test was calculated at 3.09. For the 12-point subtests it was 1.85 for the suffixed words, but a much lower .40 for the monomorphemic words. The findings indicate that the reliability of the total test and the suffixed word subtest is moderate, whereas the reliability of the monomorphemic subtest is very good.

Scorer's Consistency

To establish the reliability of the scoring system, the responses of all subjects were tape recorded. Each subject's protocol then was scored by a second rater. The interrater agreement was 85%; all discrepancies were resolved through consultation.

Validity

Validity of the Test as a Measure of Word Meaning

A two-way analysis of variance with grade level as a between-subjects factor and word type as a within-subjects factor was computed to assess the validity of the test as a measure of word meaning. Sixth-grade students' scores were expected to be substantially higher than fifth-grade students' scores. Results of this analysis are reported in Table 4. The .05 level of significance was used. The F value of 3.40 which resulted from

Table 4

ANOVA Summary Table

Source	df	SS	F
Grade Level (GL)	1	64.07	3.40
Error 1	28	528.27	
Word Type (WT)	1	52.26	11.85*
GL x WT	1	.27	.06
Error 2	28	123.47	

*$p < .01$

comparing the overall score of fifth and sixth graders was not
significant at the .05 level.

Validity of the Test as a Measure of Sensitivity to Morphemic
Composition

Chi-square analyses and phi coefficients were calculated in
order to establish the validity of the test as a measure of
sensitivity to morphemes. For this analysis, the .01 level of
significance was used to protect against Type I error. As Table 5
shows, the two variables, knowledge of word meaning and ability to
circle the base word, were not found to be significantly related.
No chi-square analysis reached statistical significance.

Table 5

Chi-Square Summary Table

Word	Corrected Chi-Square	Phi	p
betrayal	2.17	0.34	0.14
deduction	0.00	0.01	1.00
engagement	0.00	0.00	1.00
formation	0.32	0.19	0.57
heroism	0.29	0.23	0.59
instructor	---	---	---
mastery	0.00	0.05	1.00
prisoner	0.00	0.05	1.00
prosperity	0.00	0.07	1.00
settlement	0.00	0.10	1.00
socialist	0.00	0.05	1.00
tendency	0.00	0.07	1.00

Additionally, the phi coefficients ranged from a low of .00 for engagement to a high of only .34 for betrayal. The findings from this analysis do not support the test as a measure of sensitivity to morphemic composition.

Summary

This chapter reported the findings from analyses of the data collected for this study. The next chapter will summarize the study, discuss the results, and offer conclusions.

CHAPTER 5

SUMMARY, DISCUSSION, AND CONCLUSIONS

This chapter summarizes the first four chapters of this report, presents a discussion of the findings, and offers conclusions about the resolutions of the problem that this study addressed. All generalizations based on this study should be considered within the limitations described in Chapter 1.

Summary

The purpose of this study was to develop and initially validate a test of students' use of morphemic analysis in deriving word meaning. The eight steps in test construction described by Borg and Gall (1983) were followed.

Definition of Test Objective

The objective of the test was to develop a test of students' use of morphemic analysis in deriving word meaning. The test was meant to be a diagnostic instrument that would assess students' sensitivity to morphemic composition of words when producing or retaining the meaning of words.

Definition of Target Population

The population for which this test was intended was fifth-through sixth-grade children. The field testing was restricted to average readers who lived in a midwestern town with a population of about 35,000.

Review of Related Measures

The primary reason for choosing the stated objective was that a review of literature revealed the existence of no such test. Other existing tests related to morphemic analysis measure either knowlege of the meaning of individual morphemes or the ability to separate the morphemes in a word. The validity of these tests is questionable. Current morphemic analysis tests do not assess students' knowledge of the meaning of derived words.

Development of an Item Pool

The test items were taken from an earlier study by Moore (1980). Two sets of words were matched according to frequency, imagery, length, number of syllables, and pronounceability. The sets varied according to morphemic composition--one set consisted of monomorphemic words and the other consisted of suffixed words.

Preparation of a Prototype

The prototype consisted of two subtests, one consisting of 12 monomorphemic words and the other consisting of 12 suffixed words. Each word was placed in a sentence context. The items were presented in random order to the subjects. A 3-point system patterned after the WISC-R Vocabulary subtest was used for scoring subjects' free response definitions of the target words.

Evaluation of a Prototype

Following a pilot study, necessary revisions were made in the prototype and in the testing procedures. The instrument then was

field-tested with 15 fifth-grade and 15 sixth-grade readers whose ITBS Reading subtest percentile ranks ranged from 30 to 70.

Characteristics of the individual test items were examined. First, item-to-test correlations were calculated for each item. One item, prosperity, was found to be negatively correlated with total test scores. The average correlation was in the low moderate range. The difficulty index calculated for each item indicated that two words, instructor and panorama, were inappropriate for the test. The average difficulty index was satisfactory.

Revision of the Prototype

Revision of the prototype was not within the scope of this study. Recommendations for future revision are offered later in this chapter.

Collection of Data on Test Validity and Reliability

Statistical indications of reliability were calculated for the total test and for the monomorphemic- and suffixed-word subtests. Reliability coefficients for the total test and for the suffixed words were moderate, but the reliability coefficient for the monomorphemic word subtest was very high. The standard errors of measurement followed the same pattern.

To initially validate the test as a measure of word meaning, an analysis of variance was calculated to compare the scores of the fifth graders with the scores of the sixth graders. As expected, the sixth graders scored higher, but the difference was

not statistically significant. Analysis of subjects' performance on the two types of words, monomorphemic and suffixed, also was compared in the analysis of variance. The students scored significantly higher on the monomorphemic words than on the suffixed words (F = 11.85, p = .0018).

Another step in the validation process concerned the validity of the test as a measure of sensitivity to morphemic composition. In this step, the students were asked to circle the root word in each of the suffixed words. A chi-square analysis and phi coefficient calculation were used to analyze the relationship between two variables, knowledge of word meaning and identity of the root, for each word. These variables were not found to be statistically significantly related for any of the words.

Discussion

Morphemic analysis is often recommended as a method of helping students acquire meaning vocabulary, and some research evidence is available to support this recommendation. However, no test with acceptable reliability and validity has been developed that assesses students' use of morphemic analysis in deriving word meaning. The purpose of this study was to begin development of such a test. This section presents a discussion of the steps that were taken toward that end.

Definition of Test Objectives

The objective defined for this test was to determine students' use of morphemic analysis in deriving word meaning.

This objective remains an important one. No test that satisfies
this objective was discovered during the course of this study.

Definition of Target Population

Because instruction in morphemic analysis commonly begins
about third grade, this test was initially designed for use with
fourth- through sixth-grade students. Results of the pilot study
indicated that this test was too difficult for fourth-grade
students who were average readers, so fifth-grade students were
substituted for the field test. Additionally, the low mean scores
of the fifth graders, who earned 37% of the total score, and the
mean scores of the sixth graders, who earned 45%, indicate that
the test might be of sufficient difficulty for use with older
students.

In order to extend the target population for whom this test
is appropriate, it should be field-tested with low- and high-
achieving students in grades 5 and 6 and with students in the full
range of reading abilities in grades 7 through 9. In addition,
information should be gathered on students from other geographic
areas. Socioeconomic status might be another variable to include
when defining the target population.

Development of an Item Pool

A problem arose with multiple-meaning words. For example,
the sentence for engagement was Their next engagement will be an
important one. For those students who gave the meaning of
engagement as going to be married, the procedure was to probe for

another meaning. If the students were unable to give another, they were given 1 point instead of a possible 2. In other words, the students were penalized even though they had given a correct meaning for the word. Other multiple-meaning words that presented this problem were deduction and settlement. Future development of this test should consider how to score multiple-meaning words.

It is possible that the measure of imagery, which was one property Moore (1980) used to balance the two sets of words, was not appropriate for this study's target population. The imagery ratings were from a study by Paivio, Yuille, and Madigan (1968). In that study, the subjects were college students, ages 17 to 46, with a median age between 19 and 20. Current research does not reveal whether imagery ratings vary with age, but at least three of the words on the list of monomorphemic words seemed to elicit especially high imagery from students in this study. Armadillo, avalanche, and hurricane might have a higher imagery rating for students of the ages tested in this study when compared with the college students used in establishing the ratings. Thus, the two sets of words might not be balanced with respect to all properties except morphemic composition.

Preparation of a Prototype

The sentence contexts that were used approximated real reading. Even though the sentences had been revised to limit their semantic constraints, the sentences should be revised even more. It was difficult to determine in many cases whether the

students were deriving meaning from context or whether the students actually knew the meanings of the target words. For instance, the sentence for tragedy is The terrible tragedy was on T.V. The word terrible seemed to lead many students to a correct definition for tragedy. Because words are rarely read in isolation, the sentence context format should be retained, but the contexts should be made less semantically constraining.

The testing procedure used in this study was satisfactory. The students were shown individual sentences on separate laminated strips and asked to give the meaning of the underlined words. The sentence strips were somewhat unwieldy, but they had two important advantages. The strips allowed the students to see only one sentence, which limited possible distractions, and the strips could be administered in random order. Since the test is designed to be administered on an individual basis, this system appears to be acceptable.

The scoring system, while somewhat time-consuming and difficult to use, is superior to a multiple-choice or true-false format because it seems to give a complete, rich indication of a student's understanding of words. However, further refinements could be made in the system so that it would give an even better indication of the quality of an answer. For example, one of the two-point categories for the vocabulary answers on the WISC-R is the class to which something belongs. Therefore, a student defining armadillo as an animal would receive two points.

Clearly, a student who knows that an armadillo has armor, lives in the desert, and eats ants has a greater understanding of the word than one who only knows that an armadillo is an animal. Consideration should be given to the development of a system which would include three-point answers and consequently allow for greater discrimination in assessing the quality of an answer.

Evaluation of the Prototype

Two indications of the characteristics of the individual items were calculated, item-to-test correlations and difficulty indexes. Tinkelman (1971) suggested that .70 is an acceptable item-to-test correlation for vocabulary items, although in other areas such as social studies .40 or .50 is acceptable. The average correlation obtained here falls short of .70, but it is still in the low moderate range. For future revisions, a first step toward improving this correlation should be the elimination of the word prosperity, which showed a negative correlation.

A possible explanation for the fact that the mean item-to-test correlation was not higher is that item-to-test correlations usually are calculated on items that are either right or wrong. However, the scoring system used in this study was not dichotomous; there were two degrees of correctness. Subjects could have received 1 or 2 points for correct responses. This scoring system was masked by the way item-to-test correlations were calculated (i.e., both 2-point and 1-point answers were simply scored as correct, and 0-point answers were scored as

incorrect). Further refinement of the calculations to reflect subjects' degrees of correctness might alter the correlations.

A difficulty index also was calculated for each item. Ideally, the mean for such indexes should be about .50 (Borg & Gall, 1983). The mean for the items in the current instrument was .47, which compares favorably with the ideal. Eliminating instructor, which all subjects answered correctly, and panorama, which no subject answered correctly, and deleting prosperity, which was found to be negatively correlated with the total test, would bring the mean difficulty of the test to .50, which is the ideal. Of course words substituted for these three would need to be evaluated.

Since this test is designed to measure use of morphemic analysis in deriving word meaning, the roots of the suffixed words should offer a clear clue to the meaning of their derivatives. The word socialist does not fit the specification and should be eliminated. The unsuitability of the item is also evidenced by its low item-to-test correlation.

Collection of Data on Reliability and Validity

Reliability and validity are prime considerations in test development. The reliability coefficient that was calculated for the total test, Cronbach's alpha, resulted in a coefficient of .77. This is close to the .80 level accepted by most authorities for determining whether a test is reliable or not (Anastasi, 1976). This .77 coefficient also compares favorably with the .84

coefficient listed for the WISC-R Vocabulary subtest for the 10 1/2-year-old age group. The coefficient for the suffixed words was slightly lower, .72, but the one for the monomorphemic words was a very high .99.

A second statistical indication of reliability is the standard error of measurement (SEM), a measure of the discrepancies between the obtained scores and the "true" scores. The SEM for the total 24-word instrument was 3.09. In comparison, the SEM for the 20-word vocabulary subtest of the WISC-R is 1.19. The similarity of the SEM for this test with the one for the WISC-R vocabulary subtest and the fact that the SEM is derived from this test's .77 reliability coefficient indicates that it is marginal. In addition, the SEM for the suffixed word subtest, 1.85, was comparable to that of the total test, but the SEM for the monomorphemic word subtest, .40, was exceptionally good. These SEMs and reliability coefficients indicate that some caution must be taken when interpreting scores from this test, but that the reliability is acceptable.

The scoring system itself was found to be reliable. The initial interrater agreement of 85% was satisfactory for the somewhat subjective scoring system.

In order to validate the test as a measure of word meaning, an analysis of variance was used to compare total test scores by grade level. A distinct, though nonsignificant, difference was found between the scores of the two groups. Thus, this study did

not support the test as a measure of word meaning. One explanation for this finding is that a greater difference might have been obtained if more than one grade level separated the groups. The initial plan was to use fourth and sixth graders; clear differences were expected between the two groups because older students are known to have larger vocabularies than younger students. Because the results of the pilot study showed the test to be too difficult for fourth graders, the decision was made to use fifth graders. This apparently did not allow for sufficient differences in students' reading levels.

One outcome of this study is the normative data that are displayed in Table 1. In the future, subjects' performance on this test could be interpreted by comparing the total test scores of individual students with the group's total scores, and by comparing individual students' discrepancies between scores on the two subtests with the results of the current study.

Chi-square analyses were calculated for subjects' scores on knowledge of the meaning of each suffixed word and subjects' scores on ability to circle the base of each suffixed word. These analyses were an attempt to establish the validity of the test as a measure of sensitivity to morphemes. It was expected that there would be a low moderate relationship between these two measures. However, subjects' performance with these measures was not found to be related. The researcher is unable to discern any variables

which would explain this apparent lack of relationship between these tasks.

Additionally, subjects tended to define more monomorphemic words than suffixed words, and this difference was statistically significant. Three possible explanations for this difference are presented here.

First, study subjects might not be sensitive to morphemes as an aid to meaning. The students in this sample might have been taught to look for little words within big words as a spelling strategy (e.g., son in prisoner) with no attention to meaning. The students might also have been taught to divide root words and affixes as a decoding strategy, again with little or no attention given to meaning. In either case, the students might take a purely mechanical approach to breaking words into parts, and the approach might give them no help in determining the meaning of derived words. Indeed, students might have received no instruction at all in structural or morphemic analysis, so they might not analyze word parts at all. If the subjects in this study genuinely were not sensitive to the morphemic composition of words in deriving word meaning, then they would not be expected to produce evidence of such sensitivity on this test.

Second, an explanation for the subjects' better performance with monomorphemic words than suffixed words is the nature of the monomorphemic words. As pointed out earlier, the monomorphemic subtest included a few words, armadillo, avalanche,

and hurricane, that seemed to generate stronger images than any suffixed words. Thus, the two subtests might not have been of approximately equal difficulty.

Finally, the perspective that this researcher favors on the subjects' relatively high scores on monomorphemic words is that this is to be expected. That is, this study begins a collection of normative data on middle school students' performance with monomorphemic and suffixed words. Perhaps average students at this stage can be expected to produce the discrepancy found here. If this is the case, then future individuals who do not produce similar discrepancies would be considered unusual.

Conclusions

The primary purpose of this study was to begin the development of a meaning vocabulary test which would assess students' use of morphemic analysis. This concluding section consists of two parts. The first part summarizes the noteworthy outcomes of this study, and the second part contains specific recommendations for revising the test.

The following outcomes of this study deserve attention:

1. Some of the individual words are questionable. One had a negative item-to-test correlation and two were of inappropriate difficulty. The overall item-to-test correlation was only in the low moderate range, and needs improvement. The average difficulty index was quite good. Finally, for this study's age group the imagery level for three monomorphemic words, armadillo, avalanche,

and hurricane, seemed quite high in relation to any suffixed words.

2. The sentence contexts need further revision. Ideally, they would give syntactic information but would offer very limited semantic clues.

3. The procedure is satisfactory for use in individual testing. The scoring system seems superior to any multiple-choice system, but further refinement might make it even more sensitive to the quality of subjects' responses.

4. The test has face validity, but empirically derived indicators of validity were not obtained. Although the sixth graders did score higher than the fifth graders, the difference was not statistically significant. Therefore, the validity of the test as a measure of word meaning has not been established empirically. The validity of the test as a measure of sensitivity to morphemes also is questionable. The two variables, knowledge of the meaning of the suffixed word and ability to circle its root, were not found to be statistically significantly related.

5. The information gathered in this study represents only the performance of a particular group of 5th- and 6th-grade students from a particular school district, who have been taught in a certain way, and who are all of average reading ability. Their performance may or may not be representative of students of different abilities or age groups. Normative data need to be

gathered for many groups in order to present expected, normal performance with this test.

6. The information gathered in this study represents only the performance of readers with a particular type of word (i.e., nominalizations). This performance may or may not be representative of readers' performance with other words containing more than one morpheme (i.e., prefixed words, contracions, compound words, and suffixed words that are not nominalizations).

This test should not be used in its present form. The following steps are recommended for its improvement:

1. Delete the word prosperity, which had a negative item-to-test correlation.

2. Delete the two words, instructor and panorama, which were not of appropriate difficulty.

3. Delete the word socialist because the root word does not give a clear clue to the meaning of the suffixed word and because its item-to-test corrrelation is unsatisfactory.

4. Revise the sentence contexts to make them less semantically constraining.

5. Develop a more sophisticated scoring system. Consider using a 4-point system instead of the 3-point system now used. Within this scoring system, develop a method of dealing with multiple-meaning words.

6. Investigate the possiblity that imagery ratings might be different for this age group. If imagery does change with age,

then select new words in order to balance the two sets on all properties but morphemic composition.

7. As an alternative to balancing the sets of words as suggested above, simply select monomorphemic and suffixed words without stressing their balance on key word properties. Then collect normative data in order to assess normal variation between subjects' performance with the two word sets.

8. As further normative data are collected, set standards for what constitutes a noteworthy discrepancy between performance on the two sets of words.

REFERENCES

Amoriell, W. J., & Hofler, D. B. (1984). Vocabulary development --Contextual and morphemic cuing. Reading World, 23, 383-385.

Barnes, M. W. (1942). Root words and current meanings. School and Society, 56, 610-611.

Barr, A. J., Goodnight, J. H., Sall, J. P., & Helwig, J. T. (1982). Statistical Analysis System (SAS). Raleigh, NC: SAS Institute, Inc.

Borg, W. R., & Gall, M. D. (1983). Educational research: An introduction. New York: Longman.

Burmeister, L. (1976). Vocabulary development in content areas through morphemes. Journal of Reading, 19, 481-487.

Buros, O. K. (1968). Reading tests and reviews. Highland Park, NJ: Gryphon Press.

Carroll, J. B. (1940). Knowledge of English roots and affixes as related to vocabulary and Latin study. Journal of Educational Research, 34, 102-111.

Carroll, J. B., Davies, P., & Richman, R. (1971). The American Heritage word frequency book. Boston: Houghton Mifflin.

Dale, E. L., & Milligan, J. L. (1970). Techniques in vocabulary development. Reading Improvement, 7, 1-5.

Dale, E., & O'Rourke, J. (1971). Techniques of teaching vocabulary. Palo Alto, CA: Field Educational Publications, Inc.

Deighton, L. C. (1959). Vocabulary development in the classroom. New York: Columbia University Press.

Dolch, E. W. (1955). Recognition of long words. Education, 75, 604-608.

Fagan, W. T., Cooper, C., & Jensen, J. (1975). Measures for research and evaluation in the English langauge arts. Urbana, IL: ERIC Clearinghouse on Reading and Communication Skills.

Groff, P. (1972). Teaching affixes for reading improvement. Reading Improvement, 9, 28-31.

Harris, T. L., & Hodges, R. E. (1981). A dictionary of reading and related terms. Newark, DE: International Reading Association.

Hieronymus, A. N., Lindquist, E. F., & Hoover, H. D. (1978). Iowa Tests of Basic Skills, Form 8. Boston: Houghton Mifflin.

Hull, C. H., & Nie, N. H. (Eds.). (1980). Statistical Package for Social Sciences (SPSS), Version H. New York: Mcgraw-Hill.

Hunt, J. T. (1953). The relation among vocabulary, structural analysis and reading. Journal of Educational Psychology, 44, 193-202.

Johnson, D. D., & Pearson, P. D. (1978). Teaching reading vocabulary. New York: Holt, Rinehart & Winston.

Karlsen, B., Madden, R., & Gardner, E. F. (1976). Stanford diagnostic reading test. New York: Harcourt Brace Jovanovich.

McCullough, C. M. (1963). McCullough Word Analysis Tests. Lexington, MA: Ginn & Co. Personnel Press, Inc.

McWhirter, N. (1983). Guiness 1983 book of world records. New York: Bantam.

Moore, D. W. (1980). The effect of morphemic composition of words and grade level of students on word identification. Unpublished doctoral dissertation, University of Georgia.

Morris, W. (Ed.). (1969). The American Heritage dictionary of the English langauge. Boston: Houghton Mifflin.

Nagy, W. E., & Anderson, R. C. (1983). How many words are there in printed school English? Reading Research Quarterly, 19, 304-330.

Otterman, L. M. (1955). The value of teaching prefixes and word-roots. Journal of Educational Research, 48, 611-616.

Paivio, A., Yuille, A. C., & Madigan, S. A. (1968). Concreteness, imagery and meaningfulness values for 925 nouns. Journal of Experimental Psychology Monograph, 76 (1, Pt. 2).

Petty, W. T., Herold, C. P., & Stoll, E. (1967). The state of the knowledge about the teaching of vocabulary. Sacramento, CA: Sacramento State College. (ERIC Document Reproduction Service No. ED 012 395)

Schell, L. M. (1967). Teaching structural analysis. The Reading Teacher, 21, 133-137.

Shepherd, J. F. (1973). The relation between knowledge of word parts and knowledge of derivations among college freshmen. Unpublished doctoral dissertation, New York University.

Shepherd, J. F. (1975). Research on the relationships between the meanings of morphemes and the meanings of derivations. In P. L. Nacke (Ed.), Interaction research and practice in college-adult reading (Twenty-third Yearbook). The National Reading Conference, Inc.

Stauffer, R. G. (1942). A study of prefixes in the Thorndike list to establish a list of prefixes that should be taught in the elementary school. Journal of Educational Research, 35, 453-458.

Tate, H. L. (1936). Two experiments in reading-vocabulary building. Modern Language Journal, 23, 214-218.

Templeton, S. (1983). Using the spelling/meaning connection to develop word knowledge in older students. Journal of Reading, 27, 8-14.

Thompson, E. (1958). The master-word approach to vocabulary training. Journal of Developmental Reading, 2, 62-66.

Thorndike, E. L. (1921). The teachers' word book of 10,000 words. New York: Teachers College Press.

Tinkelman, S. N. (1971). Planning the objective test. In R. L. Thorndike (Ed.), Educational measurement (2nd ed.). Washington, DC: American Council on Education.

Wechsler, D. (1974). Wechsler Intelligence Scale for Children-- Revised. New York: The Psychological Corporation.

Williams, R. H., & Zimmerman, D. W. (1984). On the virtues and vices of the standard error of measurement. Journal of Experimental Education, 52, 231-233.

APPENDICES

Appendix A

Word Property Measures of Monomorphemic Words

Test Word	Imagery Rating	Frequency Index	Pronounceability Rating	Number of Syllables	Word Length
armadillo	3.77	42.0	3.0	4	9
avalanche	6.27	39.6	5.8	3	9
derelict	4.60	31.0	3.8	3	9
enterprise	3.07	46.8	3.0	31	10
episode	3.37	44.4	3.6	3	7
fantasy	3.70	43.9	3.0	3	7
hurricane	6.33	50.3	2.4	3	9
intellect	2.93	40.9	3.2	3	9
panorama	4.47	40.9	4.6	4	8
semester	3.53	42.2	2.6	3	8
tragedy	4.70	46.4	4.0	3	7
velocity	3.97	48.9	4.6	4	8
M	4.23	43.11	3.63	3.25	8.25
SD	1.12	5.08	.99	.45	.97
SE	.32	1.47	.29	.13	.28

Word Property Measures of Suffixed Words

Test Word	Imagery Rating	Frequency Index	Pronounceability Rating	Number of Syllables	Word Length
betrayal	3.57	34.0	2.6	3	8
deduction	2.67	34.0	2.8	3	9
engagement	4.70	44.2	5.0	3	10
formation	3.87	50.2	1.8	3	9
heroism	4.17	38.5	5.6	3	7
instructor	5.70	42.8	2.8	3	10
mastery	2.77	43.9	2.6	3	7
prisoner	6.23	48.7	3.0	3	8
prosperity	4.47	46.1	3.8	4	10
settlement	4.60	52.9	2.4	3	10
socialist	3.50	36.0	5.0	3	9
tendency	2.20	48.2	3.0	3	8
$\underline{M}$	4.01	43.29	3.37	3.08	8.75
$\underline{SD}$	1.22	6.41	1.21	.29	1.13
$\underline{SE}$	.35	1.85	.35	.08	.33

Appendix B

Scoring Criteria

Each word is scored 2, 1, or 0. (See the general scoring
principles and model answers.)

If a subject gives an answer which is not contradictory to
the correct meaning and which leads the examiner to believe he may
know more, the examiner should ask him to clarify his answer,
being careful not to reveal any additional information. Only
questions such as, "Can you tell me more?" or "Can you explain
that?" are permitted.

General Scoring Principles

2 Points

1. A good synonym ("a panorama is a view," "an instructor is
a teacher," "a derelict is a bum").

2. One or more identifying or important features (" a
prisoner is somebody in jail," "a fantasy is an imaginary story,"
"heroism is being brave").

3. A general classification to which the word belongs ("an
armadillo is an animal," "a hurricane is a storm," "a semester is
part of a school year").

4. Several less important but correct features which
together indicate understanding of the word ("a hurricane has big
winds and rain and tears things up," "velocity is going fast or
going slow").

1 Point

1. A vague or less important synonym ("a <u>hurricane</u> is a wind").

2. A characteristic which is correct but not important or definitive ("an <u>armadillo</u> has four legs," "an <u>avalanche</u> is in the mountains").

3. An example using the word itself, not elaborated ("a <u>panorama</u> of colors," "like an <u>engagement</u> ring").

4. A correct definition of a related form of the word (defining <u>settle</u> instead of <u>settlement</u>, <u>prosper</u> instead of <u>prosperity</u>, <u>form</u> instead of <u>formation</u>).

0 Points

1. Wrong answers.

2. Verbalisms (<u>Fantasy</u> Island).

<u>Sample Responses</u>

1. armadillo

2 points: an animal; an animal with armor, etc.

1 point: an armadillo was in the desert; it has four legs.

0 points: a ship.

2. avalanche

2 points: lots of rocks or snow fall down a mountain; wall of snow falling like a mudslide; a rock slide.

1 point: it can kill you; he was in an avalanche.

0 points: they're in the movies.

3. derelict

   2 points: a bum; he wears old clothes and stays on the

   street; somebody who is down and out.

   1 point: a poor person.

   0 points: a disease.

4. enterprise

   2 points: a business; a purposeful activity; venture.

   1 point: he was very enterprising; job; invention.

   0 points: it's downtown.

5. episode

   2 points: one story on a TV show; isolated section in an

   ongoing story; one show in a series.

   1 point: it just tells a little about the story.

   0 point: a party.

6. fantasy

   2 points: an imaginary story; dream, not reality.

   1 point: I like to read fantasy.

   0 points: Fantasy Island.

7. hurricane

   2 points: a big storm; a big storm and things get torn up,

   etc.; large windstorm generated over a body of

   water.

   1 point: lots of rain; wind.

   0 points: they're in Florida.

8. intellect

    2 points: the ability to think and reason; smartness.

    1 point: a real smart person.

    0 points: work.

9. panorama

    2 points: a big view; a wide view.

    1 point: like a sunset.

    0 points: picture.

10. semester

    2 points: half a school year; frame or segment of time; half

               of something.

    1 point: like a semester test; 9 weeks of school.

    0 points: year.

11. tragedy

    2 points: something terrible that happens; catastrophe.

    1 point: tragedy is when your dog dies; accident.

    0 points: movie.

12. velocity

    2 points: how fast something goes; speed; speed and

               direction.

    1 point: cars have velocity.

    0 points: fuel.

13. betrayal

    2 points: you think somebody is on your side and he goes

               against you; not being loyal.

1 point: he betrayed his friend; hurt someone's feelings.

0 points: cheating.

14. deduction

2 points: something that is subtracted; something that is

taken away; reduction.

1 point: it's on Dad's check; subtract.

0 points: bill.

15. engagement

2 points: like a date; an appointment; meeting.

1 point: bands have them; say they're going to get married;

speech; going together.

0 points: talking.

16. formation

2 points: the way something is shaped; physical organization

of people or objects into a pattern; the pattern

or shape of something.

1 point: like in marching; single file; in a row.

0 points: farming.

17. heroism

2 points: being brave or courageous; courage; bravery.

1 point: you get a medal for it; saving somebody.

0 points: leadership.

18. instructor

2 points: a teacher; person who guides you in learning a

subject.

1 point: Mr. X is an instructor.

0 points: a knife.

19. mastery

2 points: knowing all of something; great expertise; skill.

1 point: like a mastery test.

0 points: a secret.

20. prisoner

2 points: somebody who's locked up in jail; one who has no

freedom; an inmate; jailbird.

1 point: like a prisoner of war.

0 points: he wears stripes.

21. prosperity

2 points: wealth and material well-being; doing well in

life; having things.

1 point: money

0 points: inclination.

22. settlement

2 points: a town or loosely organized small community; place

where people come and live.

1 point: Dike is the next settlement.

0 points: like an argument.

23. socialist

2 points: somebody who thinks everybody should have the same

money and stuff; person who believes the purpose

of society is to benefit everyone equally.

1 point: a Russian.

0 points: a welfare worker.

24. tendency

2 points: it likes to do something; it leans toward it;

leaning or an attitude.

1 point: she tends to be bossy; direction; a need; towards.

0 points: shyness.

Appendix C

Test Sentences

Monomorphemic Words

A large armadillo is quite unusual.

The sudden avalanche destroyed the town.

The old derelict sat on the corner.

Jim's new enterprise earns much money.

Next week's episode should be exciting.

My brother's fantasy made me laugh.

Last year's hurricane was the worst.

Our group intellect will solve this.

The wide panorama left Sharon speechless.

This next semester will be easy.

The terrible tragedy was on T.V.

The rocket's velocity worried the men.

Suffixed Words

His friend's betrayal deeply hurt him.

One more deduction will ruin me.

Their next engagement is in Atlanta.

The marching formation was single file.

The captain's heroism insired his men.

Our English instructor is here.

His coaching mastery won many games.

That one prisoner looks really mean.

This country's prosperity is in danger.

The next <u>settlement</u> is named Parker.

The young <u>socialist</u> lives in Russia.

A plant's <u>tendency</u> is toward sunlight.

Appendix D

Instructions for Testing

Begin by saying to the student, "Hi, my name is Mrs. Baxter. I'm a student at UNI, and I'm doing this for a class I'm taking. What's your name? I have some sentences on these strips and in each one of them one word is underlined. I would like you to read the sentence either out loud or to yourself, it doesn't matter which, and then I want to tell me what the underlined word means. Do you understand? (Explain if necessary.) If it's OK with you, I'd like to tape your answers so I can go back and listen to them if I need to. Is that OK? OK, here's the first sentence...."

If the child gives an answer which makes you believe he/she knows what the word means, it is permissible to say, "Explain what you mean," or "Tell me more about it," or some other form of question that gives no information.

When this is finished, give the student the sheet containing the suffixed words and ask him/her to circle the base in each word.

Appendix E

Scoring Protocol

Name_____Teacher_____Date_____

P.R._____   Grade_____ School_____

Score 1-12 _____

Score 13-24_____   Score

_____

  1. armadillo-

_____

  2. avalanche-

_____

  3. derelict-

_____

  4. enterprise-

_____

  5. episode-

_____

  6. fantasy-

_____

  7. hurricane-

_____

  8. intellect-

_____

  9. panorama-

_____

10. semester-

_____

11. tragedy-

_____

12. velocity-

_____

13. betrayal-

_____

14. deduction-

_____

15. engagement-

_____

16. formation-

_____

17. heroism-

_____

18. instructor-

_____

19. mastery-

_____

20. prisoner-

_____

21. prosperity-

_____

22. settlement-

_____

23. socialist-

_____

24. tendency-

_____

<u>Preliminary Sentence Contexts</u>

Please fill in the blanks with any words that make sense to you.

Return to Maelou Baxter, 144 Ed. Center, 273-2698, by Jan. 31.

Thank you.

A large _____is quite unusual.

The sudden _____destroyed the town.

The old _____sat on the corner.

Jim's new _____earns much money.

Next week's _____should be exciting.

My brother's _____made me laugh.

Last year's _____was the worst.

Our group _____will solve this.

The wide _____left Sharon speechless.

This next _____will be easy.

The terrible _____was on T.V.

The rocket's _____worried the men.

His friend's _____deeply hurt him.

One more _____will ruin me.

Their next _____is in Atlanta.

The marching _____was single file.

The captain's _____inspired his men.

Our English _____is here.

His coaching _____won many games.

That one _____looks really mean.

This country's _____is in danger.

The next _____is named Parker.

The young _____lives in Russia.

A plant's _____is toward sunlight.

Root Word Circling Task

Name _____

settlement

prisoner

socialist

betrayal

tendency

mastery

deduction

heroism

prosperity

instructor

engagement

formation

Appendix F

Data for All Subjects

			Scores		
Subject		Percentile			
ID Number	Sex	Rank	Monomorphemic	Suffixed	Total
01	F	46	11	7	18
02	F	63	9	6	15
03	M	55	10	7	17
04	M	37	6	7	13
05	M	53	11	5	16
06	M	70	10	7	17
07	M	44	15	14	29
08	M	70	10	6	16
09	F	53	4	8	12
10	M	70	6	4	10
11	M	70	9	11	20
12	M	55	17	14	31
13	F	63	10	8	18
14	M	68	6	8	14
15	M	40	14	8	22
16	F	70	11	7	18
17	F	38	5	9	14
18	F	70	14	12	26

Scores

Subject ID Number	Sex	Percentile Rank	Monomorphemic	Suffixed	Total
19	F	62	9	7	16
20	F	56	10	5	15
21	M	60	17	18	35
22	M	56	11	10	21
23	M	68	14	15	29
24	M	60	12	15	27
25	M	40	11	10	21
26	M	44	17	11	28
27	M	56	14	11	25
28	F	48	9	6	15
29	F	50	8	8	16
30	M	62	13	5	18