The testing of reading proficiency in vernacular languages

by Diane Wroge

Originally published as:

Wroge, Diane. 1998. "The testing of reading proficiency in vernacular languages." *Notes on Literacy* 24.1:9–33. See also: Wroge 1998 (in *Bibliography (Literacy)*)

[Topics: testing: literacy skills]

Diane Wroge has been a literacy specialist in Papua New Guinea (PNG) with the Summer Institute of Linguistics (SIL) since 1983. She is currently serving as SIL-PNG's Literacy Liaison Coordinator. She has a Master of Science degree from Florida State University in Multilingual/Multicultural Education.

1. Introduction

Developing vernacular literacy programs and the teaching of reading in vernacular languages found in developing nations around the world present additional challenges to educators and literacy specialists that are not found in industrialized nations. There are sociopolitical, educational, linguistic, and demographic challenges that pose a variety of interwoven factors that need to be addressed. They are, many times, even outside the realm of reading itself and beyond the influence of adult educators and literacy specialists. Then, there also is the challenge of the constant debate over which reading theory (bottom-up, top-down, interactive, or the newly emerging paradigm of social constructivism) more accurately represents the nature of the reading process. Along with all of these challenges, there are relative few articles written on how to test reading proficiency in vernacular languages found in developing nations. There is, however, a large amount of reading research available on first language (L1) and second language (L2) reading. Even though the L1 and L2 reading research studies the reading processes in languages with a long literate history, until educators and literacy specialists working in

vernacular literacy programs provide more research in how to test reading proficiency, perhaps some of the L1 and L2 testing instruments could be applied to the testing of vernacular language reading. However, the research on vernacular education and literacy development that is available is mostly written from anthropological or sociopolitical perspectives.

2. Purpose of article

The purpose of this article is to examine four types of reading tests from the L1 and L2 reading research and discuss their possible applications to the testing of reading proficiency in vernacular languages. The four types of reading tests can be subdivided into two sections. The first section is the identification of reading errors and miscues. The second section includes the testing of comprehension. Under the identification of reading errors or miscues section, the two types of tests examined are the Informal Reading Inventory (IRI) and Miscue Analysis. Under the testing of comprehension section, the two types of tests examined are the Sentence Verification Technique (SVT) and Immediate Recall Protocol. But first, because vernacular languages are never functioning in a society unaffected from other influences, a brief discussion of some of the sociopolitical, educational, linguistic, demographic, and reading challenges is included. These challenges are included to help situate the context and identify some of the complexity in testing reading proficiency in any language, much less in languages that only very recently have developed a writing system and have relatively new readers in the language. These new readers must develop their skills without the luxury of a wide variety of reading materials available in their language to help them maintain their vernacular reading ability.

3. Various challenges and influences on reading instruction

3.1. The challenges of the sociopolitical and historical context

First, there are the sociopolitical factors that exert tremendous pressure on any multilingual, multicultural country. How best can the country serve its people but also be part of the world's economy and international business community? What are the sociopolitical and historical factors that have led to a country's adoption of its language policy? Are culturally and linguistically diverse ethnic groups even recognized or given any official status? Of those who speak the language that is politically dominant, what is their attitude toward speakers of other languages? What status do vernacular language people assign to their language and to the language that is politically or socially dominate? The answers to some of these questions may help determine whether vernacular literacy efforts are even accepted much less whether successful or not (Trudell 1993a, citing a study by Anderson).

3.2. The challenges of the educational system

There are also educational challenges. What is the medium of instruction in the educational system? The medium of instruction usually is one of two types: (1) a language spoken in the country and probably taught as a second language (L2) or (2) a language of wider communication (LWC) that has limited acceptance beyond its educational, political, and economical spheres, because it is heard only in the

schools in a foreign language (FL) context. For example, Spanish is the medium of instruction in Latin American countries, so the Spanish formal school system is considered to be an L2 context. Yet, in contrast, English is the medium of instruction in both Papua New Guinea and anglophone African countries and French for francophone African countries, so the English or French schooling is considered to be in an FL context. Educational advancement is contingent upon how well students progress through their academic studies taught in the second language (L2) or the foreign language (FL). Therefore, developing nations do not have widespread and well-developed compulsory primary education systems leading to long histories as literate-producing and literacy dependent societies similar to what is present in industrialized nations with a strong societal pull to produce literates as demanded by the economics of the work situation.

Some countries have implemented bilingual education. Bilingual education programs are organized around two languages of instruction, the mother tongue and a second language. How much the mother tongue is used, in what contexts, and in which subjects, and when and how the second language is introduced is a function of the curriculum and the educational institution's goals. There are five types of bilingual education. They are transitional, immersion, submersion, maintenance, and vernacular (Trudell 1993a).

In transitional bilingual programs, the mother tongue of minority students is used only until the students have obtained enough fluency in the second (national) language to continue their education in it. In immersion bilingual programs, students receive instruction in the second language all the time, but the mother tongue is recognized with prestige. In submersion bilingual programs, all the instruction all the time is in the second language, and the mother tongue has such low prestige that it is not considered important enough to be part of the classrooms' instruction. In maintenance programs, there is instruction in both the mother tongue and second language throughout the primary education years of schooling. Though not really a bilingual education program, in vernacular programs, the mother tongue is used exclusively.

Vernacular languages are found in countries with at least one of these educational contexts. Some of these languages are found in countries that have a combination of educational programs to meet the educational needs of their country, for example, Peru's bilingual education system for many of the Amazonian language groups.

3.3. The challenges of the linguistic situation

Then, there are also the linguistic challenges found in the vernacular languages themselves and how best to develop a writing system from the linguistic analyses. The ideas listed are not meant to be exhaustive, but they are meant to be representative examples of some of the complex issues of not only linguistic analyses of vernacular languages but also related to the development of orthographic representations of a writing system in those languages. On what type of script (Roman or non-Roman) should the orthography be based? How should tone be represented (Kutsch Lojenga 1993 and Koffi 1994)? What about languages needing special characters; which diacritics are best to use? Should an orthography be more based on its phonological phonemic analysis and have a more sound-to-symbol correspondence to one of the dialects in the language area, or should it be more morphophonemically based so it is able to be used

by all the dialects (Weber, McConnel, Weber, andBryson 1994)? Then, there are highly agglutinating languages that have words, usually verbs, that can be very long, because a string of affixes attaches to the stem and root. The affixes and stem and roots can be their own morphemes (Matthews 1994). So, grammatically, what English considers as a sentence, in an agglutinating language, the same "sentence" would be one word with a number of morphemes attached to the root and stem.

3.4. The challenges of the demographic context of vernacular languages

Another one of the challenges not strictly limited to linguistic considerations is the size of the language groups. A vernacular literacy program that works in an area with small language groups may not be easily transferable to areas of the world with large language groups. For instance, in Papua New Guinea, 850 languages are divided among only 4,000,000 people, and no one language is significantly numerically prominent. The Enga language is the largest language group in Papua New Guinea and has only 200,000 speakers. In contrast, the African continent has numerically larger language groups than most parts of the world. The contrast between size of language group can even be present in a country. For example, Latin American countries tend to have smaller language groups in the jungle areas yet numerically larger groups in the mountain regions, especially the people of Inca or Mayan descent. Of course, the size of language group becomes complicated as it interfaces with other factors, such as, what language is socially and politically dominant. The dominant language may be one that was introduced from outsiders, and therefore, it is not linguistically related to any of the languages present in the country.

3.5. The challenges of the nature of the reading process

The nature of the reading process itself presents another set of challenges to the teaching of reading, and thus, the testing of reading proficiency, in vernacular languages. Reading theoreticians and researchers have debated over what are the necessary components to the reading process. Reading theories, generally, are characterized to be one of three types:

1. *Bottom-up* emphasizing text-driven,
2. *Top-down* emphasizing meaning-based or concept-driven and the contribution of the reader
3. *Interactive*, emphasizing simultaneous bottom-up and top-down processing

(See Appendix A: Characteristics of reading processes and approaches.) "In these characterizations it may be helpful to think of the *top* as the reader, and the *bottom* as the written text. The basic controversy among these theories concerns the location of the source of control in the reading act: Is it the text, or the reader, which controls the reading process, or both?" (McCormick 1988:1).

Phonics and basal readers, that is, the look-say method, stress reading *subskills.* Therefore, there is concerted effort to concentrate and practice the hierarchical, sequential, and analytical combination of the phonemes (that is, blending). Meaning will eventually result, especially once word recognition becomes automated, so short-term memory is free to concentrate on meaning. Therefore, decoding and word attack

skills are an emphasis in bottom-up methods of reading. The psycholinguistic nature of reading and whole language approaches to the teaching of reading stress reading comprehension *strategies* that are necessary for meaningful reading. Some of these strategies consist of the following steps:

- Hypothesis and prediction.
- Select and sample the hypothesis and prediction.
- Test the hypothesis.
- Confirm or revise the original hypothesis and prediction.

Reading, thus, is more than a decoding process; it is a holistic process that must be contextualized. The interactive theorists have stressed the importance of prior knowledge, the *schemata,* that the reader brings to the text during the bottom-up and top-down interactive reading of a text. New information is more likely to be remembered if instantiated into a cognitive network map of old and preexisting information. A fourth type of reading theory, which is a newly emerging paradigm, is the *social nature* of the reading process (Bernhardt 1987; Bloome 1986; and Palicsar and Brown 1988). In the *social nature* of the reading process, teachers and students co-construct the meaning of a text through their social interaction about the text. The *constructivist* view of reading has its theoretical underpinnings from the work of the Soviet developmental psychologist, L. S. Vygotsky.

The goals of reading teachers and literacy specialists are to produce readers who are:

1. *Fluent* (read at a natural speed and with accuracy)
2. Able to read with *comprehension* (understand the meaning of the text)
3. Also able to *decode* unfamiliar words (word attack skills)

3.6. The challenges of identifying reading proficiency levels

Language proficiency levels of any kind, of which reading is one type of skill, are difficult not only to define but also to measure. Language can be divided into two types of skills: productive (speaking and writing) or receptive (listening and reading). In general, there are four types of language tests. The types of tests are:

1. Achievement
2. Diagnostic
3. Placement
4. Proficiency (Underhill 1987)

For the purpose of this article, achievement tests will be contrasted with proficiency tests.

Achievement tests are concerned with measuring what has been learned during a particular reading program, syllabus, or reading primer. The tests themselves, therefore, are course or syllabus specific and, in the case of testing reading, should take into consideration the reading methodological considerations reflected in the primer design and how learning to read is taught. Achievement tests are "to sample appropriately from the syllabus content to enable the test interpreter to infer the student's mastery of the entire syllabus" (Byrnes and Canale 1987:77). Proficiency measurements, however, are "concerned with the results of the learning process but not with how that learning is acquired … and reflect the accumulation of knowledge by the student and indicates the lack of a specific syllabus" (ibid., 78).

Most proficiency levels are defined and designed based upon Indo-European languages. The ACTFL (American Council on the Teaching of Foreign Languages) proficiency guidelines initially focused upon establishing proficiency levels' guidelines in Spanish, German, and French. When the attempt was made to apply these guidelines to Chinese, Japanese, and Russian, an Indo-European bias emerged. Likewise, many of the reading proficiency levels are based upon English reading, and thus, English linguistic features necessary for reading in English. These designs, devised with the English language structure in mind, may not be transferable to vernacular languages that have no relation to the English language structure.

Davis (1994) realized the importance of this point in her study on the vernacular reading ability of the Machiguenga people of the Peruvian Amazon. Since the Machiguenga language has agglutinating verbs, some of its morphemes are only one letter long. Therefore, readers need to attend significantly more to individual letters than in English reading. Reading incorrectly, or ignoring a letter that is a morpheme in Machiguenga, is equivalent to missing a word in English and may affect comprehension as meaning could be misconstrued or missed all together. While working with Yale, an agglutinating language in Irian Jaya, Indonesia, Matthews (1994) found this to be true. One of the Yale-speaking readers read a passage in the wrong tense, because the tense marker affix in the Yale language is attached to the end of the string of affixes connected to the verb stem/root.

The "bottom-up" approaches to reading lend themselves to standardized tests. Webster and Braswell (1991) found that reading achievement test performance and norm-referenced test scores are susceptible to curriculum bias. Since the 1970s, and the development of the whole language approach based upon the psycholinguistic processes of reading, the overemphasized use of norm-referenced standardized tests is being questioned (Anderson, Heibert, Scott, and Wilkinson 1985; Royer, Carlos, Carlisle, and Furman 1991). Do standardized tests really measure what reading entails? Are they valid and reliable measures of the reading process? The emphasis is on the reading process and does not question standardized tests' validity and reliability.

In the area of reading methodology, including the testing of reading, there exists constant friction between production (bottom-up decoding skills) and reading ability competent (top-down meaningful understanding and use). Elley (1992) encountered this in coming up with a definition of literacy that became acceptable to the committee overseeing his comparative study on the variables that influence reading and its teaching in 32 countries. For the purpose of the comparative study, Elley defined reading literacy as "the ability to understand and use those written language forms required by society and/or valued by the individual" (Elley 1992:3).

> Such a definition was found to be general enough to accommodate the diversity of traditions and languages represented in the participating countries but specific enough to provide some guidance for test construction…. The emphasis on language forms required by society reflects the concept of functional literacy and coping with reading tasks frequently encountered in an organized society…. However, a broader definition was required to meet the needs of many of the National Committees who argued for the inclusion of higher-level thinking…. While the majority of NRCs favored an emphasis on both understanding and use, the constraints of mass testing, standardized conditions, traditional policies, and limited school time in many nations, made it imperative that the major stress be placed on understanding (Elley 1992:3–4, italics original).

Likewise, the same constraints Elley found in his study also apply to the testing of reading proficiency in vernacular languages. Lundberg and Linnakylä (1993) also examined the same 32 countries and the teaching of reading in them. They stated that the reading texts make comparisons difficult because of reading's "cognitive demands, processes of comprehension and individual strategies in approaching reading tasks may reflect unity as well as diversity in reading and the teaching of reading" (Lundberg and Linnakylä 1993:v).

The three considerations under the nature of receptive language proficiency include:

a. The "test taker's stage of development in receptive skills

b. The influence of (oral or written) text demands on comprehension strategies

c. The influence on comprehension strategies of the purposes for which a person deals with a text" (Canale 1984:349)

"That is, to assure an adequate representation of receptive language proficiency in a given test, it is suggested here that minimal consideration be given to developmental stages, text demands, situations and purposes of language use, and their differing influences on reading and listening skills" (ibid., 352). Canale states that proficiency testing should not only be valid, reliable, and practicable, but they should also be acceptable and provide feedback. Acceptability and feedback are two of the more ignored considerations in any evaluation design.

For the purpose of her study on the literacy acquisition, retention, and usage of reading in the Machiguenga language, Davis adopted Barr and Johnson's 1991 definition "that reading is a relationship between comprehension, prior knowledge, and skill with print. Skill with print must be developed for the process to function smoothly. Comprehension is the goal" (Davis 1994:11). Listed below are the minimum skill levels for Machiguenga readers that Davis used to access the reading ability of the Machiguenga. These skill levels can be applied to any material of any level of difficulty and be used as criteria for success. They are listed and explained here, because the findings provide some empirical support for adopting a reading evaluation method (in this case, the Informal Reading Inventory) designed for the reading in a first language (that is, English), to that of a non-European and agglutinating language (that is, Machiguenga). Before these levels can be accepted as valid reading measurements of proficiency

levels, they have to be applied to other kinds of reading measurements to make sure they are not only applicable to one measurement.

The four areas of the reading task she measured were accuracy, rate, fluency, and comprehension. Her minimum levels for each area were

- *accuracy* (92 percent)
- *rate* (80 syllables per minute, but may have to be adjusted as a function of linguistic structure)
- *fluency* (two on a scale of one to five), and
- *comprehension* (three on a scale of one to five).

Accuracy will be defined by the type of analysis and is dependent on the evaluation instrument used. For example, the percentage of accurate reading can include the number of words or syllables in a passage divided by the reading errors or the reading miscues. One of the major differences between the two is in interpretation; errors reveal what aspects of reading ability are lacking, and miscues reveal insights into the psycholinguistic reading developmental process.

The rate of reading is a function of time. The five-point fluency scale included:

1. Struggling, syllable-by-syllable
2. Reading word-by-word
3. Reading phrase-by-phrase
4. Observance of punctuation
5. Natural, communicative expression

The comprehension of the Machiguenga readers was categorized on a five-point scale. The comprehension scale focused on the literal reporting of facts, since those are demonstrable and the most easily identified aspects of comprehension. The progression of comprehension includes:

None: No demonstrated recall.

Poor: Fragments of the passage, but not the main topic.

Fair: Identified the topic and evidenced literal understanding of the most important facts.

Good: Identified the topic and elaborated on some facts, but did not cover all the support ideas contained in the passage.

Excellent: Identified the topic and also reported most of the supporting information given in the passage.

4. The Summer Institute of Linguistics: Opportunities to address the challenges

[Topics: evaluation: methods]

The linguists and literacy specialists with the Summer Institute of Linguistics (SIL) study the world's cultures and vernacular languages. The goals of SIL include anthropological studies, linguistic analysis, translation, and vernacular literacy promotion. Therefore, SIL and other international organizations, like UNESCO, have unique opportunities to provide the necessary and additional reading research to support whether the Informal Reading Inventory (IRI), Miscue Analysis, Sentence Verification Technique (SVT), and Immediate Recall Protocol are applicable to non-European and vernacular languages found throughout the world.

4.1. The Informal Reading Inventory

Davis (1994) chose to use the Informal Reading Inventory as the basis of testing the Machiguenga reading, because it has been used for the last 50 years. The Informal Reading Inventory was devised by Betts in the 1940s and 1950s and utilizes oral reading tests. Davis devised a set of tests, and each contained

a. a set of ten syllables

b. one story to read aloud, and

c. one story to read silently and retell.

In the set of stories were a variety of genres and some of them were used for readers of different skill levels. The three reading skills levels included basic, intermediate, and advanced. The genres in the reading texts differed according to reading level. For the basic skills assessment, stories were familiar to all the readers. Narratives and procedural genres were used. In the procedural texts, however, incorrect information was embedded in the familiar stories. For the intermediate skills assessment, traditional folk stories and letters, either from the school or written by the native speaker research assistant, were used. Since the Machiguenga New Testament is the only advanced reading material available, it was used for the advanced skills assessment. The New Testament provided reading assessment of the most difficult kind of reading material, because it is translated material, set in another culture, and written in a more formal style. Each of the stories was controlled for length, and a syllable count provided the basis of uniform calculations. There also was a Spanish reading section that included excerpts from reports or business letters written by local authorities.

The formal data collection took place during house-to-house interviews before the reading tests were administered. The procedures for administering some of the reading tests were for the subjects to read the passage silently first, and then they read aloud. The subjects' oral reading was tape recorded. For other reading tests, after silent reading, the passage was returned to the researcher and the subjects retold what they read. The subjects' oral retelling was also tape recorded. For the advanced readers of Spanish, the Machiguenga research assistant dictated a couple of sentences for the subjects to write down on a piece of paper, and their writing samples were clipped with the reading tests.

The interviews were tape recordings that were transcribed. Individual Reading Score Charts were designed to include what passage was read and scores for accuracy, rate, fluency, and comprehension for

the Machiguenga reading, and if appropriate, the Spanish reading. The following procedure was practiced for scoring the reading tests. Errors were coded and defined as "any observed response which deviated from the expected response to the printed text" (Davis 1994:34). The coding system included the following:

Circle	=	Omission	Overline + A	=	Attempted correction
Caret	=	Insertion	Overline + C	=	Successful correction
Overwrite	=	Substitution	Parenthesis	=	Aid received
Overline	=	Repetition			

An error counted as one point. Missed segments (words, phrases, line or lines) also counted as one point. Within a word, each incorrect syllable, which in Machiguenga could be a morpheme, was counted as an error. Also, a sequence of incorrect syllables was counted as another within-word error. Each substitution, dropped ending, regression, and unsuccessful error correction was counted as one error.

Davis concludes that the Informal Reading Inventory method of testing proved very helpful in the testing of Machiguenga reading. She also notes that it is a method that can be used locally, applied to nonstandard situations, applied to any language, and used with different genres of material.

4.2. Miscue analysis

Miscue analysis was developed by Kenneth Goodman as a way to identify insights into the psycholinguistic processes of reading. Miscues in oral reading give clues into the cognitive processes that happen during reading, and, therefore, there is not an emphasis on erroneous reading.

Goodman (1985), in his 1965 study on linguistic cues and miscues in reading, performed a descriptive study of the children's oral reading. He defined reading as an active reconstruction of a message from written language (involving comprehension) and as a psycholinguistic process. "The child learning to read his native language has already internalized these cue systems to the point where he is responding to them without being consciously aware of the process" (Goodman 1985:129). The subjects were asked to orally read the story. Their reading behavior was noted on the worksheets the children read. Finally, each subject was asked to retell the story without looking back at it. The oral reading and retelling were tape recorded. In the 1965 study, he also had the children read words from a word list. Those same words also appeared in the stories read. This study is the seminal work for the justification of using miscue analysis for testing reading rather than discrete point or word identification tests that are common on standardized tests.

Under the heading, "Average words missed in word list and in story," Goodman recorded list average, the average and percentage both under also missed in story, and the ratio. Another table recorded the "Ability to read words in context which were missed on list" and included the column headings and percentages of: less than ½, more than ½, more than 2/3, more than 2/4, more than 4/5, followed by N. The "Total errors and substitution errors on lists" included the column headings: list errors average, the average and percentage under included substitutions, and the ratio. The "One-time substitutions for known words in stories" table included the column headings: average substitutions, average lines read, and substitutions per line read. The "Regressions in reading" table, on the other hand, had two rows of subheadings. The first was of word only and under word only: to correct word, to correct intonation on word, and total. Another row's subheading was the phrase and rows and under the phrase were: to correct word by repeating phrase, to rephrase, to change intonation, and total. These rows included calculations under the columns. The column headings under each of the three grades were per child and per line read. The tables' descriptions were written out to provide examples in case they have relevance to the testing of reading in vernacular languages.

The purpose of the study by Parker, Hasbrouck, and Tindal (1992) was to compare "the criterion validity of traditional oral reading frequency (ORF) with the criterion validity of two new types of oral reading assessment procedures modified to include Goodman's miscues." The traditional ORF is the number of words read correctly per minute. It is considered a well researched, reliable, and valid reading measure and is especially used in remedial and special education. The first modified miscue assessment procedure included modifications of the traditional ORF to include: meaning change miscues (ORF-M), severe meaning change miscues (ORF-S), or uncorrected miscues (ORF-U). Only these miscues were counted, not all errors. The second modified procedure, on the other hand, included modifications of traditional oral reading accuracy (ORA) or percentage of words read correctly with no consideration of reading time. These modifications included: meaning change miscues (ORA-M), severe meaning change miscues (ORA-S), or uncorrected miscues (ORA-U) in counting errors. Again, only these miscues were counted, not all errors.

During the timed oral readings, these six errors were counted:

- Mispronunciations and substitutions
- Insertions
- Reversals of words within a phrase
- Omissions
- Repetitions
- Hesitations of at least three seconds (after which students were told the word)

These errors were subtracted from the total number of words read to obtain the words read correctly.

Matthews (1994) identified the miscues of Yale semiliterate readers and used the results to help determine what the miscues reveal about the Yales's reading ability in their native language and what areas of reading instruction need more attention. The miscue analysis of the individual's readings was used "to

determine the patterns of response and how they approached the reading task" (Matthews 1994:2). The author recognized that "one's interpretation of a pattern of reading miscues depends, obviously, on the model of reading adopted" (ibid.). The model Matthews follows is one promoted by Marilyn Jager Adams. In the summary prepared by Steven Stahl, Jean Osborn, and Fran Lehr (Adams 1990) of Adams's *Beginning to read: Thinking and learning about print*, Adam proposes that reading encompasses how the parts of four processors (orthographic, phonological, meaning, and context) work together.

Since semiliterates spend so much time and energy on decoding, there is little energy and memory available for comprehension of the text. Therefore, semiliterate readers are characterized by such tendencies as word-by-word reading, phonic errors, and lack of attention to phrasing. The seven subjects read a passage of about 140 words written by a Yale author, the subject of which was a familiar incident in the village. The subjects read the story silently first, then the second reading was tape recorded, then they retold the story. The tape recordings were replayed and miscues marked. The miscues marked were

- repeats
- omissions or additions of words
- errors on the suffixes or suffix omissions, and
- other phonic miscues.

4.3. Sentence Verification technique

Royer, Carlos, Carlisle, and Furman's (1991) study was designed to test the validity of a new assessment procedure, the Sentence Verification Technique (SVT), to measure the progress in transitional bilingual education programs. They also stated that standardized tests cannot meet the needs of bilingual assessment. Therefore, the solution of reliable and valid tests should be developed at the local level. These tests should include three characteristics. First, the local curriculum goals and materials should be the focus of the tests. Second, the gains in linguistic skills and subject matter content that characterize the progress of transitional bilingual education students need to be incorporated into the tests. Finally, the content of the tests should be written to included the experiences and values of the culture. The SVT technique measures reading and listening comprehension and has proven to be reliable, valid, and easy to administer in a variety of situations and for a variety of purposes.

The SVT tests are developed from local reading materials and designed by writing one of four types of test sentences representing each sentence in a reading passage. The test sentences include:

1. *Originals*, replicas within of a sentence in the passage
2. *Paraphrases*, changing words in a passage without changing meaning
3. *Meaning changes*, changing one or two words that alter the meaning of a sentence
4. *Distractors*, different in wording and meaning but similar in theme, syntax, and vocabulary

After reading a passage, the subjects responded "Yes" or "No" to the SVT test consisting of a balance between the test sentence types. "Yes" meant the sentence had the same meaning (original or paraphrase) as the passage read, and "No" meant the sentence had a different meaning (meaning changes and distractors).

The authors claim that SVT tests possess a number of other properties that make them attractive to be used to measure reading and listening progress. First, they can be constructed from any text based on materials relevant to the school curriculum and reading methodology. Second, the test development takes minimal training and can be developed by local school personnel. Third, performance interpretation is straightforward and does not require the development of norms of performance. Finally, the greatest attraction is that they can be developed in any language by a native speaker of that language.

The stories should be based upon narrative passages selected from the reading series commonly used in mainstream classes of the school system or the vernacular literacy classes. The stories' sentences can vary in length according to the reading level of the subjects, and they should have a story-like structure (beginning, middle, and ending).

4.4. Immediate recall protocol

Recall is a common measure for the assessment of comprehension in L1 research. Considered to be the "most straightforward assessment of the result of the text-reader interaction" (Johnston, quoted in Bernhardt 1991:200), recall reveals "something about the organization of stored information, about some of the retrieval strategies used by readers, and reveals the method of reconstruction which [the reader] employs to encode information in a text" (Bernhardt 1991:200, brackets in original). Bernhardt states that there is a rich source of information in recall data, because the data explains and reflects how subjects analyze information and how they think about it.

Immediate recall has been shown to be an effective and valid integrative measure of comprehension. Recall protocols permit the analysis in which learners reconstruct messages from text as well as the observation of the textual components that interfere with understanding. It provides a relatively direct view of the process of comprehension (Bernhardt 1988).

Recall accrues several advantages. It can show where a lack of grammar is interfering with the communication between text and reader while not focusing attention on linguistic elements in texts. Generating recall data does not influence a reader's understanding of a text. "In other words, a free recall measure provides a purer measure of comprehension, uncomplicated by linguistic performance and test interference" (Bernhardt 1991:200).

The Immediate Recall Protocol uses Meyer's recall protocol scoring system and is based on Grices's case grammar. "It identifies the structural characteristics as well as lexical units of a passage which helps to assess the relationship between passage type and level of performance. Moreover, because the system arranges the idea units hierarchically, its use allows the experimenter to see not only which lexical and relational units are remembered, but also from which portion of the structure those units are recalled" (Bernhardt 1991:201).

The central part of a recall protocol scoring template is the three-prong text's structural relationships. The development of structural relationships are determined first on the top level structure (for example, comparison-contrast), second on the macrostructural relationships from words like "because" and "therefore," and third, on the lexical predicates and role arguments or "referential indices" on the microstructural level.

One of the disadvantages in Meyer's recall protocol scoring system is that it takes too much time to analyze a text's hierarchical relationship. Bernhardt estimates that it takes between 25 to 50 hours per 250-word text. In addition, the amount of time training in the use of the scoring system is too time extensive. Therefore, it may not be very practical nor applicable for vernacular languages, since, more than likely, there will not be the luxury of more than one researcher who is accompanied by a native speaker assistant. "There is a need, therefore, of developing alternative methods of scoring recall protocols that, nevertheless, provide the same quality data without the concomitant time and training commitment" (Bernhardt 1991:203).

During oral recall, the researcher could list a set of ideas, words, or concepts from the text and merely check an item that is mentioned. Alternatively, the tape recording of the recalls can be qualitatively described to reveal other aspects of the comprehension process on how readers construct meaning from the text that may not be able to be measured quantitatively.

5. Conclusion

These four types of reading tests were chosen, because:

1. They have been developed in response to different kinds of reading situations.
2. They have been applied to different languages.
3. In some cases, they have been developed in response to dissatisfaction of other standardized reading tests' results for linguistically and culturally diverse minorities in contrast to the mainstream culture.

How culturally and linguistically adaptable are the Informal Reading Inventory, Miscue Analysis, Sentence Verification Technique, and Immediate Recall Protocol? Until these reading testing techniques are applied to other vernacular language readings, the question will remain.

Davis (1994:38) warns that "in cross-cultural studies, however, miscue analysis must be used with great caution since miscues are easily misinterpreted when either the test administrator or the test taker are working in a second language." She also recognized that some of the errors classified in the Informal Reading Inventory of Machiguenga readers could have been analyzed as miscues, even though many attempted corrections were undoubtedly miscues, though in her study they were classified as errors. In addition, she warns that, since comprehension assessment is considered to be imprecise, teachers and researchers should exercise caution when drawing conclusions from comprehension scores. Yet, the teachers and researchers need to be aware that fluent oral reading does not automatically include comprehension, and therefore, oral reading may not be the best indicator of people's reading ability. Also, some clues in the text may activate the wrong schemata. If that happens, readers may be lead "down the wrong garden path" in their understanding of the meaning of a passage read (Wroge 1994:11).

In conclusion, this article raises a number of questions that need consideration and perhaps testing. How can testing reading proficiency be designed that is fair to both poles of the bottom-up and top-down reading methodologies continuum? By nature of the test design, would results be skewed to one pole or the other? Is it even feasible for this comparison to be done? The vernacular reading teachers and supervisors, in partnership with linguists and literacy workers of the Summer Institute of Linguistics, have developed teaching of reading materials found within the full spectrum of the reading methodology continuum (see Appendix A: Characteristics of the reading processes and approaches). Therefore, they are in the position to best tackle these questions and answer whether it is feasible to design a reading proficiency test useable by any or all the vernacular languages in the world.

Appendix A

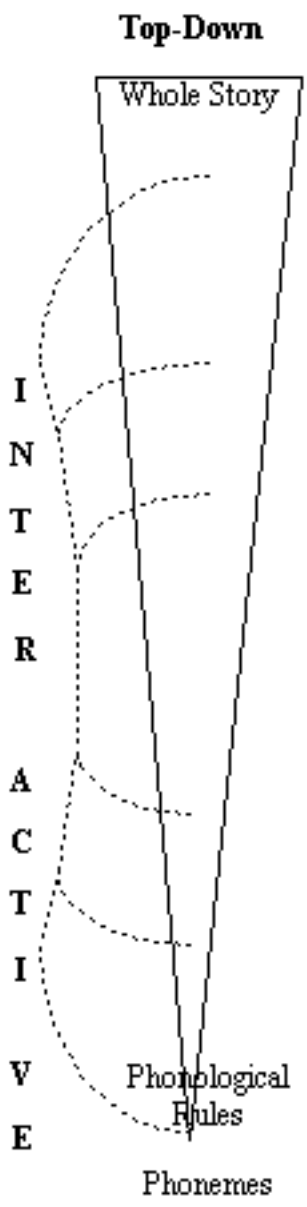
Reading Processes	Reading Approaches	Characteristics
	Whole Language (Psycholinguistics)	Develops comprehension and fluency, uses whole stories, natural and authentic texts that are repetitive and predicable, contextualized
	Language Experience Approach (LEA)	Generate story from student's experience, write down story student(s) say, read story
	Sentence	Story charts, sentence strips
	The Multi-Strategy Method by Stringer (SIL) and Faraclas (UPNG)	Two **separate** processes per lesson taught within one day: whole language and syllable primer, processes not explicitly mixed, new readers put together on their own
	The Gudschinsky Method (SIL) by Sarah Gudschinsky	Eclectic within same primer lesson—picturable key word and its syllables analyzed, synthesized, contrasted and built, connected story material with limited vocabulary taught in previous lessons
	Whole Word or Syllable	Basal readers, sight words on word cards, syllables and words decoded through phonics
	Phonics	Develops word attack (decoding) skills, sound-symbol correspondence, blending, phonological rules, decontextualized

Chart 1: Characteristics of reading processes and approaches

References

(1998). *Notes on Literacy*, *24*(1).

Adams, Marilyn Jager. 1990. Beginning to read*:* Thinking and learning about print. Summary of Adams, M. 1990b. Urbana-Champaign, IL: *University of Illinois, Reading Research and Education Center*.
See also Adams, M. 1990a Bibliography (Literacy) Citations

Alderson, J. Charles, Karl J. Krahnke, and Charles W. Stansfield (editors). 1987. *Reviews of English language proficiency tests.* Washington DC: TESOL.

Anderson, Richard C., Elfreda H. Hiebert, Judith A. Scott, and Ian A. G. Wilkinson. 1985. Becoming a nation of readers. Washington, DC: *US Department of Education, The National Institute of Education*.
See also Anderson, R. C., Hiebert, Scott, and Wilkinson 1985 Bibliography (Literacy) Citations

Anderson, Ron (editor). [1994]. "Cumulative Index 1966–1993." *Notes on Literacy*19.4.

Bernhardt, Elizabeth B. 1988. "Merging research and assessment perspectives in foreign language reading." *Technical report submitted to the 1988–1989 Institute of Advanced Studies.* Mellon Fellowship Program: National Foreign Language Center.
Citations

Bernhardt, Elizabeth B. 1991. *Reading development in a second language: Theoretical, empirical, and classroom perspectives.* Norwood, NJ: Ablex.
Citations

Bernhardt, Elizabeth B. 1987. "The text as a participant in instruction." *Theory Into Practice* 26:32–37.
Citations

Bloome, David. 1986. "Building literacy and the classroom community." *Theory Into Practice* 25:71–76.
Citations

Byrnes, Heidi, and Michael Canale (editors). 1987. *Defining and developing proficiency: Guidelines, implementations and concepts.* Lincolnwood, IL: National Textbook. Citations

Canale, Michael. 1984. "Considerations in the testing of reading and listening proficiency." *Foreign Language Annals* 17.4:349–357. Citations

Carrell, Patricia, Joanne Devine, and David Eskey (editors). 1988. *Interactive approaches to second language reading.* New York, NY: Cambridge University Press.

Cziko, Gary A. 1978. "Differences in first- and second-language reading: The use of syntactic, semantic, and discourse constraints." *Canadian Modern Language Review* 34:473–489.

Davis, Patricia M. 1991. Cognition and learning*:* A review of the literature with reference to ethnolinguistic minorities. Dallas, TX: *Summer Institute of Linguistics*. View document See alsoDavis, P. 1991 Bibliography (Literacy)

Davis, Patricia M. 1994. Literacy acquisition, retention, and usage*:* A case study of the Machiguenga of the Peruvian Amazon. Ph.D. dissertation. University of Texas at Austin: *University Microfilms, Ann Arbor, MI*. See also Davis, P. 1994 Bibliography (Literacy) Citations

Elley, Warwick B. 1992. "How in the world do students read?" *IEA study of reading literacy.*The Hague: IEA (The International Association for the Evaluation of Educational Achievement). Citations

Fajonyomi, Ayodele A., and F. B. O. Ala-Adeyemi. 1993. "Comparative effectiveness of three languages of instruction among adult literacy learners." *Journal of Reading* 37.3:174–182.

Froese, Victor. 1991. “Aspects of reading in ESL and L1 students.” Paper presented at the 41st Annual Meeting of the National Reading Conference, Palm Springs, CA. December 4, 1991.

Goodman, Kenneth. 1985. “A linguistic study of cues and miscues in reading.” In *Theoretical models and processes of reading, third edition,* edited by Harry Singer and Robert Ruddell. 1985:129–134.
Citations

Gregerson, Marilyn J. 1994. Methodological considerations in teaching reading to ethnic minority children. Notes on Literacy 20.3:43–50.
View document See alsoGregerson 1994 Bibliography (Literacy)

Hedgcock, John, and Dwight Atkinson. “Differing reading-writing relationships in L1 and L2 literacy development.” *TESOL Quarterly*27.2:329–333.

Hughes, Arthur. 1989. *Testing for language teachers.* Cambridge, England: Cambridge University Press.

Hulstijn, Jan H. 1991. “How is reading in a second language related to reading in a first language?” *AILA Review: Reading in two languages*8:5–13.

Jernudd, Bjorn H. 1993. “Planning English language acquisition in ESL and EFL societies: Development and maintenance of languages and cultures.” *Journal of Multilingual and Multicultural Development Special Issue: Bilingualism and national development,* edited by Gary M. Jones and A. Conrad K. Ozog. 14:135–149.

Koffi, Ettien N. 1994. The representation of tones in the orthography. Notes on Literacy 20.3:51–59. (Edited excerpt from Dr. Ettien’s dissertation “The interface between phonology and morpho(phono)logy in the standardization of Anyi orthography” (97–109.)
View document See alsoKoffi 1994 Bibliography (Literacy) Citations

Kutsch Lojenga, Constance. 1993. "The writing and reading of tone in Bantu languages." *Notes on Literacy* 19.1:1–19..
View document See alsoKutsch Lojenga 1993 Bibliography (Literacy) Citations

Lundberg, Ingvar, and Pirjo Linnakylä. 1993. *Teaching reading around the world.* The Hague: IEA (The International Association for the Evaluation of Educational Achievement).
Citations

Matthews, Delle P. 1994. A miscue analysis of Yale reading ability. Notes on Literacy 20.2:1–9.
View document See alsoMatthews, D. 1994 Bibliography (Literacy) Citations

McCormick, Thomas W. 1988. Theories of reading in dialogue*:* An interdisciplinary study. New York: *University Press of America*.
See also McCormick, T. 1988 Bibliography (Literacy) Citations

Palincsar, Annemarie, and Ann L. Brown. 1988. "Teaching and practicing thinking skills to promote comprehension in the context of group problem solving." *RASE* 9:53–59.
Citations

Parker, Richard, Jan Hasbrouck, and Gerald Tindal. 1992. "Greater validity for oral reading fluency: Can miscues help?" *The Journal of Special Education* 25.4:492–503.
Citations

Pikkert, Joost. 1993. Schema theory, reading, shell books, and curriculum development. Notes on Literacy 19.3:26–33.
View document See alsoPikkert 1993 Bibliography (Literacy)

Royer, James M., Maria S. Carlos, Joanne Carlisle, and Gail Furman. 1991a. "A new procedure for assessing progress in transitional bilingual education programs." *The Bilingual Review* 16.1:3–14.
Citations

Royer, James M., and Maria S. Carlos. 1991b. "Assessing the language acquisition progress of limited English proficient students: Problems and a new alternative." *Applied Measurement in Education* 4.2:85–113.

Singer, Harry, and Robert B. Ruddell. 1985. Theoretical models and the processes of reading. 3rd edition. Newark, DE: *International Reading Association*. 960 pages.
See also Singer and Ruddell 1985 Bibliography (Literacy)

Spangenberg-Urbschat, Karen Pritchard, and Robert Pritchard (editors). 1994. *Kids come in all languages: Reading instruction for ESL students.*Newark, DE: International Reading Association.

Trudell, Barbara. 1993a. *Beyond the bilingual classroom: Literacy acquisition among Peruvian Amazon communities.* Dallas, TX: Summer Institute of Linguistics; University of Texas at Arlington.
Citations

Trudell, Barbara L. 1993. The dilemma of biliteracy in Peruvian Amazonia. Notes on Literacy 19.3:16–25.
View document See alsoTrudell, B. 1993b Bibliography (Literacy)

Underhill, Nic. 1987. *Testing spoken language: A handbook of oral testing techniques.* Cambridge, England: Cambridge University Press.
Citations

Weber, David J., Stephen R. McConnell, Diana Weber, and Beth J. Bryson. 1994. Primer*:* A tool for developing early reading materials. Occasional Publications in Academic Computing 18. Dallas: *Summer Institute of Linguistics*.
View document See alsoWeber, McConnel, Weber, and Bryson 1994 Bibliography (Literacy)
Citations

Wroge, Diane. 1994. "The influence of first and second languages' orthographies on the oral reading and retelling of a Pidgin English children's story." Paper presented at the Southeast

Regional TESOL Conference, New Orleans, LA, September 31, 1994. Manuscript.
Citations