

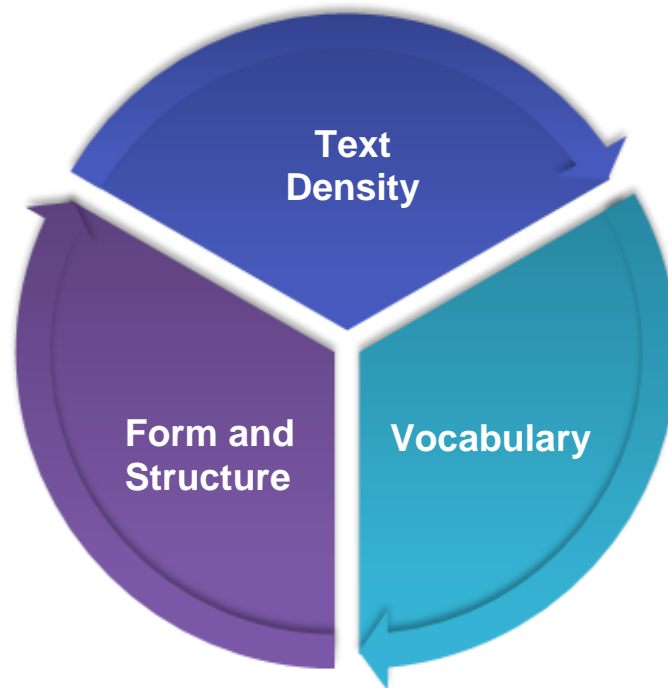
Tool to Evaluate Language Complexity of Test Items

H Gary Cook, Ph.D. & Rita MacDonald - Wisconsin Center for Education Research, WIDA

TEXT DENSITY—the average of **INFORMATION DENSITY** and **PASSAGE LENGTH**

Information Density—a ratio of the number of specific types of meaning-carrying words to the number of independent clauses

Passage Length—the number of paragraphs



FORM AND STRUCTURE—
a 4-point scale rating language forms from simple sentences to more complex forms employing adverbials, relative clauses, passive voice and reported speech

VOCABULARY—a 4-point scale rating lexical items from everyday vocabulary to highly technical words and metaphoric and idiomatic expressions

Tool to Evaluate Language Complexity of Test Items

H Gary Cook, Ph.D., Wisconsin Center for Education Research, WIDA
Rita MacDonald, Wisconsin Center for Education Research, WIDA

September 17, 2012

Background

The language complexity rubric, description and sample rating materials provided below comprise a method for evaluating the language complexity of academic content achievement test items. The aim for this tool is to inform educators, test developers, and test reviewers of the language demands required of students to respond to content assessment items. The specific focus of this tool is to better understand how students with different English language proficiency levels perform on assessment items and tests with varying levels of language complexity.

Three areas are evaluated on the language complexity rubric (shown below): text density, language structure, and vocabulary. Research has informed the selection of these areas, especially as they affect English language learners. Other aspects of language also affect what makes tasks, activities or interactions difficult for English language learners. For example, understanding how a particular interaction should occur in a particular socio-cultural context can be extremely complicated. Using an informal register to write a narrative in a literature class would be totally appropriate, but the use of an informal register to write a lab report for a science class would not be.

Areas of language like socio-cultural context, use of language in different power structures and relationships, or use of different genre have not been included because 1) substantial training would be required to appropriately identify specific features of these areas, 2) there is still professional debate about the significant components within these areas, or 3) in some cases there may be many ways to appropriately interact depending on the situation, context, interlocutors and desired outcome. Thus, they are not included in this tool.

What follows are brief descriptions of research associated with each of the evaluated areas.

Text Density is the average of the scores of 2 discrete elements: information density and passage length.

Information Density measures how tightly a text packs important information into clauses, based on the notions that the clause, as the central processing unit in grammar, constitutes an information unit, and that the greater the number of meaning-carrying words per clause, the greater the “processing load” of that clause for the reader or listener. The formula used here is an adaptation of the Lexical Density (LD) formula used in systemic functional linguistics (SFL), whereby the number of lexical items in a passage (all nouns, all main verbs except the verb “to be”, all adjectives, and only critical adverbs) is tallied and divided by the number of non-embedded (independent and dependent) clauses. The SFL calculation of LD was first described by Michael Halliday in 1994, and is used by SFL practitioners to analyze a text’s difficulty level.

“The more highly structured nature of school-based texts contributes to their lexical density. Through lexical density, academic registers pack more information into each clause, making it possible to build up the information in a text efficiently.” (Schleppegrell, 2004, p. 67)

“...in everyday spoken language, there are two to three content-carrying words per clause, but in written language, there are four to six. In science, the number can go up much higher, often exceeding ten.... Such a high density of information can create cognitive overload for students and slow down their processing of a text.” (Fang, Lamme, & Pringle, 2011, p. 28)

To reduce the level of linguistic analysis required by the rater, the formula used here for Informational Density is simplified by dividing simply by the number of independent clauses and by the elimination of adverbs from the count of lexical items.

Passage length has been included to capture the difficulty noted above by Fang, Lamme, & Pringle (2011). The greater the number of passages in general, the greater the processing load for the reader.

Language Structure is rated along a continuum of increasing difficulty, anchored at the low end by the prototypical language structures of everyday conversation, and progressing to include a high proportion of more syntactically complex structures, such as embedded relative clauses, passive voice, and adverbial phrases. As suggested by Pienemann’s (2008) theory of second language development, language is processed in hierarchical fashion, with simpler structures requiring less cognitive processing than more complex structures. The increased number of complicated syntactic structures in academic texts is recognized as a source of comprehension problems for students.

“...scientific language employs complex sentences with hierarchical structure using both subordinate clauses and embedded clauses. ...comprehension problems can arise when a sentence comprises multiple clauses where layers of semantic links and dependency relationships take time for students to sort through.” (Fang, Lamme, & Pringle, 2011, p. 71).

Vocabulary has been widely acknowledged as a critical element of academic language, and one that can be categorized for pedagogical purposes.

“...academic vocabulary, as one component of the broader academic language construct, comprises both a general academic lexicon...and specialized ones.” (Bailey, 2007, p. 12)

The importance of vocabulary knowledge to college readiness is underscored by Nation’s analysis of academic language at the college level (Nation, 2001), which found that approximately 9% of college text is comprised of general academic vocabulary, and an additional 5% is comprised of specialized or technical vocabulary. McLaughlin, August, & Snow (2000) report that 60-70% of English words carry multiple meanings, determined in relation to context, thus emphasizing the importance of the acquisition of multiple and specialized meanings for a significant portion of the English lexicon.

Our rating of vocabulary complexity is based on the frequency with which a reader is thought to encounter the specific meanings of the words in the passage, anchored at the low end by words encountered on a frequent, almost daily basis and progressing to specialized or technical words which are themselves encountered far less frequently, along with polysemous words whose particular meanings are encountered infrequently. The high end of the ratings scale also includes metaphoric meanings such as those encountered in poetry or satire.

References

- Bailey, A. (2007). *The language demands of school: Putting academic English to the test*. New Haven and London: Yale University Press.
- Fang, Z., Lamme, L.L., & Pringle, R. M. (2010). *Language and literacy in inquiry-based science classrooms*. Thousand Oaks, California: Corwin.
- Halliday, M.A. K. (1994). *An introduction to functional grammar* (2nd ed.). London: Edward Arnold.
- McLaughlin, B., August, D., & Snow, C.E. (2000). *Vocabulary knowledge and reading comprehension in English learners*. (Final Performance Report to OERI, Award No. R306F60077-97). Washington, DC: U.S. Department of Education.
- Nation, I.S.P. (2001). *Learning vocabulary in another language*. Cambridge; Cambridge University Press.
- Pienemann, M. (2008). A brief introduction to processability theory. In Kessler, J. (Ed.) *Processability approaches to second language development and second language learning*. Newcastle: Cambridge Scholars.
- Schleppegrell, M. J. (2004). *The language of schooling: A functional linguistics perspective*. Mahwah NJ: Lawrence Erlbaum.

Language Complexity Rubric

		Descriptors			
		1	2	3	4
TEXT DENSITY	Information Density	An average of one to three verbs, nouns or adjectives per sentence	An average of four to six verbs, nouns or adjectives per sentence	An average of seven to ten verbs, nouns or adjectives per sentence	An average of greater than 10 verbs, nouns or adjectives per sentence
	Passage Length	One to three paragraphs	Four to six paragraphs	Six to ten paragraphs	More than ten paragraphs
LANGUAGE FORM AND STRUCTURE	Language Forms	Simple sentences and simple grammatical forms	A few sentences and/or grammatical forms that are more sophisticated (e.g., compound S, prepositional phrases)	An even mix of sophisticated and simple sentences or grammatical forms	Over half of the passage consists of more complex sentences and/or grammatical forms (e.g., relative clauses, adverbials, passive voice, reported speech)
	Vocabulary	All high frequency, commonly used vocabulary	Few content-specific, metaphoric, uncommon meaning, or idiomatic words	A number of content-specific, technical, metaphoric, uncommon meaning and/or idiomatic words or expressions	A large number of content-specific, technical, metaphoric, uncommon meaning or idiomatic words or expressions

Description of the Language Complexity Rubric

Text Density

The language complexity rubric rates three areas of language: text density, language structure and vocabulary. Text density¹ refers to how difficult texts are. Two areas are rated here: information density and passage length. **Information density** is simply the average number of nouns, adjectives and verbs in a sentence. Here, every independent clause is considered a sentence. Here is an example passage to clarify how sentences (i.e., independent clauses), nouns, adjectives, and verbs are counted.

EXAMPLE 1: Language from a 4th grade reading passage²

The whistle blows,// and I hit him fast with a fireman's carry. // He's on his back in three seconds.// The ref's hand slaps the mat.// Pinned.// One match down.

David M. Simon's *Tough as Daisy*

The underlined words identify the nouns, adjectives and verbs to be counted and the double slashes represent clausal breaks. Notice that the verb “to be” in the second line (“He’s on....”) is not counted; no form of the verb “to be” should be counted. Helping verbs (e.g., will, may, could, should) are not counted either. The first sentence has two independent clauses: *the whistle blows* and *I hit him fast with a fireman’s carry*. Accordingly, it is counted as two sentences. Notice also that there are two sentences that are not independent clauses: *Pinned*. *One match down*. These sentences could be rewritten *He was pinned*, and *One match was down*. The dialogue makes clear that this is what was meant; thus, these incomplete sentences are counted as complete sentences. In total there are 15 selected words. There are a total of six clauses or an average of 2.5 nouns, adjectives, and verbs per sentence (clause). Using the rubric on the previous page the information density of this passage would be a 1.

Is it necessary to rate an entire passage? What if the passage is several pages long? Since introductory paragraphs sometimes differ in purpose and linguistic characteristics from those in the body of the text, our suggestion is to take two paragraphs in the middle of a passage and rate the information density there. If a passage is three or four paragraphs, the middle or last two paragraphs should be rated. Shorter passages should have all paragraphs rated. (Language Structures and Vocabulary should be assessed across the entire passage.)

Passage length is simply the number of paragraphs in a reading passage, item stimulus, or acceptable student response. This is generally straightforward, but it may require judgment in some areas. For example, here are two released NAEP mathematics items.

¹ Note that the term *text* does not just mean written text. It could also refer to spoken language as well. The discussion here focuses only on written text; however, it should be understood that the rubric could be for oral language as well.

² This reading passage was taken from a released 4th grade reading NAEP item.

4th grade, multiple-choice mathematics item characterized as “easy.”

7. Patty expects that each tomato plant in her garden will bear 24 tomatoes. If there are 6 tomato plants in her garden, how many tomatoes does she expect?

- A. 4
- B. 18
- C. 30
- D. 144

Did you use the calculator on this question?

8th grade, constructed response mathematics item characterized as “hard.”

18. Leroy has one quarter, one dime, one nickel, and one penny. Two of the coins are in his left pocket and the other two coins are in his right pocket. The coins have been randomly placed in the two pockets.

What is the probability that Leroy will be able to purchase a 30-cent candy bar with the two coins in his left pocket?

Using the coins, explain your reasoning.

Using the rubric’s passage length descriptors, how might these items be rated? One paragraph? Two paragraphs? Three paragraphs? Qualified and experienced English teachers might give different answers to this. For rating purposes, assume that a paragraph is the expression of a complete idea (a common definition of a paragraph). As a rule all elements of a test item (the stimulus, the correct response, and the distractors)) are considered part of the item stem. The item stem itself is an idea or paragraph. Consider statements like *Did you use a calculator on this question?* or *Using the coins, explain your reasoning* as part of the previous paragraph and not as separate paragraphs. With these rules, the 4th grade mathematics item would be one paragraph, and the 8th grade passage would be two paragraphs. Both of these items would be rated a 1 on the passage length section of the language complexity rubric.

Language Structure

The primary focus for this section is to rate the language forms (grammar) as being simple or complex. The goal is to identify the general pattern of language forms across the entire passage. Let us examine a reading passage by Margery Facklam and determine the complexity of language forms. This passage was used as a stimulus for a 4th grade item on NAEP. Note that sentence numbers were added for explanatory purposes.

(1) “What do bees do?” (2) Ask most people and they will say, “Bees make honey and they sting.” (3) They may even tell you that bees are fuzzy, black-and-yellow insects that live in hives. (4) But there are lots of kinds of bees, and they’re not all the same. (5) Some fly at night. (6) Some can’t sting. (7) Some live only a few months, and others live several years. (8) Every species of bee has its own story. (9) A species is one of the groups used by scientists to classify, or group, living things. (10) Animals of the same species can mate with each other. (11) And they give birth to young that can mate and give birth, or reproduce.

Declarative, interrogative, imperative or exclamatory sentences that just have a subject (if needed), verb and object with basic verb tenses are classified as simple sentences. In the passage above, the first, fifth, sixth, eighth, and tenth sentences are simple sentences.

- (1) "What do bees do?"
- (5) Some fly at night.
- (6) Some can't sting.
- (8) Every species of bee has its own story.
- (10) Animals of the same species can mate with each other.

This passage has several compound sentences: the second, fourth, and seventh sentences,

- (2) Ask most people and they will say, "Bees make honey and they sting."
- (4) But there are lots of kinds of bees, and they're not all the same.
- (7) Some live only a few months, and others live several years.

and it has several sentences with relative clauses: the third, ninth, and eleventh sentences.

- (3) They may even tell you that bees are fuzzy, black-and-yellow insects that live in hives.
- (9) A species is one of the groups used by scientists to classify, or group, living things.
- (11) And they give birth to young that can mate and give birth, or reproduce.

There are also more complex grammatical forms; see the ninth sentence, which has a reduced relative clause and is in the passive voice. This passage has a mix of sentence forms and would be rated a 3 in the language structure section of the Language Complexity rubric. The intent of this element in the rubric is not to classify every sentence but to get a general sense of the complexity of grammatical forms and sentences across a passage.

Vocabulary

This section of the rubric is meant to identify the type of vocabulary found in a reading passage, using the entire passage. That vocabulary could be everyday common words (e.g., walk, run, house), content-specific words (chart, graph, appendix), or specific technical vocabulary (secant, atom, simile). Complex vocabulary is not always related to technical terms. The way an author uses common words in uncommon ways or turns phrases can also be quite complex. Here is a quote from E.B. White.

"I have a lot of the cat in me and cats are not joiners."

The word cat is a common vocabulary word, but White uses it metaphorically to indicate his solitary, independent nature. Again, the goal in rating this section is to get a general sense of the types of vocabulary used in a passage.

Sample rating materials

Sample test items and passages taken from the National Assessment for Educational Progress (<http://nces.ed.gov/nationsreportcard/itmrlsx/landing.aspx>) released items website for reading, writing and mathematics are provided as examples to practice scoring using the rubric. Sample items are found in the following files:

R4_scr_medium.pdf – 4th grade short constructed response, reading item of medium difficulty

R4_scr_medium_key.pdf – 4th grade short constructed response, reading item key for correctly scoring the item

R4_scr_medium_resp.pdf – 4th grade short constructed response, reading item student responses

W12_ecr_easy.pdf – 12th grade, extended constructed response item of easy difficulty

W12_ecr_easy_key.pdf – 12th grade, extended constructed response item key for correctly scoring the item

W12_ecr_easy_resp.pdf – 12th grade, extended constructed response item student responses

M4_mc_easy.pdf – 4th grade multiple-choice, mathematics item of easy difficulty

M8_ecr_hard.pdf – 8th grade extended constructed response mathematics item of hard difficulty

M8_ecr_hard_key.pdf – 8th grade extended constructed response mathematics item key for correctly scoring the item

M8_ecr_hard_resp.pdf – 8th grade extended constructed response mathematics item student responses

Three test item formats are provided: multiple-choice, short constructed response, and extended constructed response. In total four items are used as examples: a 4th grade, multiple-choice reading items, a 12th grade, extended constructed response writing item, a 4th grade multiple-choice mathematics item, and an 8th grade extended constructed response, mathematics item. If a multiple-choice test item has an associated passage both the test item and the passage are rated. For constructed response items, the item stem, passage (if applicable), and acceptable student responses are rated. For rating purposes, a correct or sufficient answer is an acceptable rating student response in mathematics, full comprehension rating in reading, and sufficient rating in writing. The table below provides suggested language complexity ratings for these four items.

For scoring, an average of the information density (id) and passage length (pl) make up the text density score (TD). The sum of the text density (TD), language structure (LS) and vocabulary (V) sections make up the language complexity (LangCmplx) rating.

Table 1: Sample Language Complexity Ratings

Test Item	Language Complexity Ratings		
	Item stem	Student response*	Passage
<i>R4_scr_medium</i>	id: 2 pl: 1 TD: 1.5 LS: 2 V: 2 LangCmplx: 5.5	id: 1 pl: 1 TD: 1 LS: 2 V: 1 LangCmplx: 4	id: 1 pl: 4 TD: 2.5 LS: 2 V: 3 LangCmplx: 7.5
<i>W12_ecr_easy</i>	id: 3 pl: 2 TD: 2.5 LS: 4 V: 1 LangCmplx: 7.5	id: 2 pl: 3 TD: 2.5 LS: 3 V: 1 LangCmplx: 6.5	
<i>M4_mc_easy</i>	id: 2 pl: 1 TD: 1.5 LS: 3 V: 1 LangCmplx: 5.5		
<i>M8_ecr_hard</i>	id: 3 pl: 1 TD: 2 LS: 2 V: 2 LangCmplx: 6	id: 3 pl: 1 TD: 2 LS: 2 V: 2 LangCmplx: 6	

*If there is more than one student response, the rating is the average of student responses.