


L2 読解力測定評価の基本的問題：構成概念を中心に

Essential Issues for L2 Reading Assessment with Special Reference to the Construct of Reading

中村 優治 Nakamura, Yuji

● 慶應義塾大学
Keio University

 **Keywords** L2 読解力, 測定, テクノロジー, 学習者の個人差
L2 reading, assessment, technology, learner differences

ABSTRACT

英語を外国語として教えたり、英語を第2言語として教えたりする際の、英語読解力測定、評価は非常に複雑で困難である。テストを作成する以前の段階で、英語読解力の構成概念、構成要素を十分に理解しておく必要がある。本稿ではL2英語読解力測定にあたり、そのテスト作成のために必要な基本的問題を考えようとするものである。全体的には4つの視点からとらえ、それぞれを深く掘り下げていく。一つ目は読解力の構成概念である。2つ目は読解力の下部構成要素、そして3つ目は読解力テスト方法、最後の4つ目に読解力に関連する問題（たとえば学習者個人差、また、コンピュータなどテクノロジーの活用など）を中心に検討をおこない、インターネットで容易に読む資料が手に入る時代の英語読解力のテスト、評価、測定について結論づける。

Assessing L2 reading ability is complicated as it involves not only the ability to read, but also the knowledge of and ability to use the second or foreign language (SFL). In order to assess or test a person's ability to read and understand texts in an SFL context, it is important to understand what comprehension is, and what one has to know and do in order to be able to read. Without knowledge of what reading involves, what is technically known as the construct of SFL reading, one cannot develop a valid test of reading. In other words, in order to test the ability to read, one needs to have a theory of reading. This paper deals with basic issues of L2 reading assessment as follows: 1) the construct of reading, 2) the subskills of reading ability, 3) test method of reading, and 4) related issues on L2 reading assessment. The reading construct is most likely under-represented by all well-known standardized reading assessment systems. The point is the

extent to which we can and should measure reading construct skills and subskills. Computers and new media are likely to alter the evolution of reading tasks and tests. We must recognize that the ability to read online texts is becoming an important part of the general construct of reading. As a result, more attention should be paid to this area in regards to reading assessment.

Introduction

It is impossible to directly observe the act of reading. In fact, there is ongoing discussion about how the component subskills and strategies involved in the process of reading actually work together so that a person creates meaning. However, there is consensus among researchers that reading is an essential language skill, perhaps the most paramount, since so much of our information comes to us in written form. Students need to be able to comprehend written information required for their language development, along with other skills such as listening and speaking (2012, Hubley).

In a literate society, it is not always noticed how complicated the act of reading is. Although we have to learn to read, most of us do so easily and with relatively little effort. And so, if the act of reading is second nature to most adolescents and adults, then what could be easier than testing the ability to read? Just take a text, ask a few questions about the text, and you have a valid test of reading. However, it is not that simple (Alderson, et.al, 2015).

We should take into account numerous components such as: What text should be selected? Will any text serve the purpose about any topic? How long should the text be? How familiar should the words be? What grammatical structures should the text contain? What genre should be used? What sort of questions should be asked? Should they be about details of the text or about metaphorical meanings? Should questions be asked that require the reader to detect irony? Should the reader be required to summarize the text? To criticize the

text? (Alderson et.al, 2015).

Assessing reading ability is a complicated process, and reading is far more complex than is generally recognized. As Alderson et. al (2015) maintains, assessing L2 reading ability is even more complicated as it involves not only the ability to read but also the knowledge of and ability to use the second or foreign language (SFL). In order to assess or test a person's ability to read and understand texts in an SFL context, it is important to understand what comprehension is, and what one has to know and do in order to be able to read. Without knowledge of what reading involves, what is technically known as the construct of SFL reading, one cannot develop a valid test of reading. In other words, in order to test the ability to read, one needs to have a theory of reading (Alderson et. al, 2015).

As stated previously, the issues that will be discussed will be: 1) the construct of reading, 2) the subskills of reading ability, 3) test method of reading, and 4) related issues on L2 reading assessment.

The discussion will mainly be based on the ideas of Grabe & Jiang (2014) who cover the fundamental reading assessment issues to be taken into consideration.

1. The construct of reading

One of the most important challenges for reading assessment stems from the complexity of the construct of reading ability itself. However, determining the construct provides rationales for various instructional and assessment practices (Grabe, 2009c).

The construct of reading comprehension abilities

in relation to reading assessment, and the description on how and why reading ability is assessed, have been discussed together. Historically speaking, as Grabe & Jiang (2014) state, “construct of reading” is a concept far behind the formal assessment of reading ability, with the subskills, knowledge base, and cognitive processes (“component skills”) left functionally unassessed until the 1990`s. It is also true that the construct of reading comprehension that is assessed and the processes engaged seem to change as a result of testing type and format (Grabe & Jiang, 2014). As Grabe (2009c) says, determining the construct provides rationales for various instructional and assessment practices.

2. The reading comprehension subskills

Since there are subskills under the theoretical construct, it is natural to pay attention to the following factors that strongly impact reading ability (Grabe & Jiang, 2014) in considering reading assessment.

- 1) Efficient word recognition
- 2) Vocabulary knowledge
- 3) Efficient grammatical skills
- 4) Ability to formulate the main ideas from a text
- 5) The ability to engage in a range of strategic processes
- 6) The ability to recognize discourse and genre patterns to support comprehension
- 7) The ability to use background knowledge
- 8) The ability to interpret text critically
- 9) Effective working memory ability
- 10) Reading fluency skills
- 11) Massive experience with L2 reading
- 12) Reading motivation

(Modified from Grabe and Jiang, 2014, p.188)

Although standardized reading assessments consider the construct of reading in multiple ways,

according to Grabe and Jiang (2014), there are some component abilities that are not yet well incorporated into L2 standardized reading assessments:

- 1) Reading rate and fluency
- 2) Rapid word recognition
- 3) Search processes
- 4) Morphological knowledge
- 5) Structure awareness and discourse organization
- 6) Strategic processing ability
- 7) Paraphrasing and summarization
- 8) Synthesis skills
- 9) Critical reading

The reading construct is most likely under-represented by all well-known standardized reading assessment systems. The point is the extent to which we can and should measure reading construct skills and subskills.

Classroom based assessment makes use of test tasks from standardized testing; however, assessment in classroom contexts is primarily used to measure a students’ learning. Classroom based assessment practices, unlike standardized tests, may include:

- 1) Having students read aloud in class
- 2) Keeping a record of students dialogue
- 3) Observing students reading motivation
- 4) Observing students’ audiotaped performance
- 5) Having students list words they want to know
- 6) Having students write book reports
- 7) Keeping charts of students’ progress
- 8) Having students read aloud for the teacher
- 9) Noting students’ uses of texts
- 10) Creating student portfolios

(Adapted from Grabe and Jiang 2014, p.192)

It is clear from assessment research that L2 vocabulary knowledge is a major component influencing L2 reading abilities. Similarly, and perhaps more surprisingly for some, L2 grammar

knowledge is a major component ability of L2 reading comprehension (Grabe, 2009c). While vocabulary knowledge seems to be more important in reading than grammatical knowledge, there are circumstances when grammatical knowledge may be more crucial. Recent thinking (cf. Purpura, 2014) suggests that the distinction between the two is probably a false dichotomy and it makes more sense to consider both aspects to lie on a lexico-grammatical continuum (Alderson et al., 2015).

According to Grabe (2009c), reading comprehension requires the following subskills and knowledge sources which emerged from research on L2 reading tests:

- 1) The ability to decode graphic forms for efficient word recognition
- 2) The ability to access the meaning of a large number of words automatically
- 3) The ability to draw meaning from phrase-and clause-level grammatical information.
- 4) The ability to combine clause-level meanings to build a larger network of meaning elations (comprehend the text).
- 5) The ability to recognize discourse-level relationships and use this information to build and support comprehension.
- 6) The ability to use reading strategies with more difficult text and from a range of academic reading tasks.
- 7) The ability to set goals for reading and adjust them as needed.
- 8) The ability to use inferences of various types and to monitor comprehension in line with reading goals.
- 9) The ability to draw on prior knowledge, as appropriate.
- 10) Abilities to evaluate, integrate, and synthesize information from a text to form a situation model of comprehension.
- 11) The ability to maintain these processes fluently for an extended period of time.
- 12) The motivation to persist in reading and to use the text information appropriately in line with reader goals.

This list is in line with the idea of subskills of reading by Urmston, Raquel, & Tsang, 2013):

- Identifying specific information
- Interpreting a word or phrase as used by the writer
- Understanding main ideas and supporting ideas
- Understanding information and making an inference
- Inferring the writer's reasoning
- Interpreting an attitude or intention of the writer
- Understanding grammatical relationships of words or phrases across text
- Identifying text type

If we go further into more practical testing of L2 reading, the following are recommended as good assessment practices by Grabe (2009).

- 1) Students should be tested on a range of relevant skills.
- 2) Students should be encouraged to read longer texts (for advanced assessment, 700-1200 words, assuming 120-150 wpm)
- 3) Background knowledge influences all comprehension and needs to be accounted for in a positive way (multiple topic, multiple tasks, general topics, limited interdependence of items within some subset of tasks.)
- 4) Group questions might be used to engage discussions of reader interpretations of texts.
- 5) Extensive reading should not be discouraged by assessment produces.
- 6) The importance of identification and fluency skills needs to be explored (reading word lists, oral reading for one minute, silent reading on computer, timed reading assessment of reading)
- 7) Tests might explore ways to assess synthesis skills, evaluation skills, strategies, metacognitive knowledge, and skills monitoring (text

monitoring while reading)

- 8) Reading might be tested within a content-focuses battery (but item interdependence has to be a concern.)
- 9) Tests might consider item types that take advantage of computer interfaces (e.g. allow a text to disappear after reading, use a few hypertext links in a test passage, combine information from multiple texts in order to complete a task)
- 10) Many skills might be measured usefully through informal assessment options in classroom contexts. What one loses in reliability and objective controls could be countered by any formal and informal assessments that can be made in the classroom. (However, informal assessment is not a substitute for more formal testing)

Hubley (2012) shows reading subskills in a more concrete, practical and detailed way as follows:

3. Test Method

The theory or theoretical ideas should be put into actual tests by using test methods or test formats. Test method now should be considered. The primary purpose of assessment is to collect information to make inferences about students' reading abilities; and the test method should play the role of this purpose by representing the operational definition or the theoretical construct of reading (cf. Grabe, 2009c).

- 1) Cloze formats
- 2) Gap-filling formats (a rational reason for selecting blanks)
- 3) Multiple-choice formats
- 4) Sentence completions
- 5) Matching, (and multiple matching) techniques
- 6) Classification into groups
- 7) Text segment ordering
- 8) Dichotomous items (t/f, y/n)
- 9) Editing formats
- 10) C-tests

Level	Subskill	Typical way to test
Whole passage	Comprehend the main idea or gist meaning	What is the best title for this passage? What is the reading mainly about?
	Recognize author's attitude and biases	With which of the following statements would the author mostly closely agree?
	Distinguish fact from opinion	Which of the following is not a fact?
Section	Understand logical organization	Where would be the best place for this sentence?
Paragraph	Identify main idea and supporting details	List main ideas and match with paragraphs
	Understand topic sentences	The most important idea in Paragraph 4 is...
Across sentence boundaries	Understand the function of discourse markers	In line 22, what does "in addition" mean?
	Identify pronoun reference	What does "its" in line 6 refer to?
Sentence level	Guess the meaning of unknown words in context	Which word is closest to the meaning of "nomadic" in line 40?
	Paraphrase wording	Which phrase means the same as "brothers and sisters"

Figure 1 Modified from Hubley's list of reading subskills (Hubley, 2012, pp 213-214)

A study has confirmed that different test formats measure different aspects of reading comprehension and that text organization can influence reading comprehension based on more complex reading tasks. Therefore, new tasks requiring information synthesis may assess something different from basic comprehension (cf. Grabe & Jiang, 2014).

As Alderson et. al (2015) indicated, in all cases, the test development process is similar to the processes of identifying learners' needs, developing test specifications and item writer guidelines, identifying suitable texts, drafting test items and tasks through a process of writing, editing, moderating, and revising tasks until they are considered to be ready for trialing.

Once sample test takers are identified and administration procedures are developed, the pilot tests are trialed, marked and analyzed statistically either with Classical Test Theory procedures, and/or with Item Response Theory methods. Suitable test tasks are selected or revised, and delivered (in practice) to target candidates. When the tests have been administered during the main trial, they are marked and analyzed again, typically including inferential statistics and expert judgments, in order to examine to what extent the intended constructs have been successfully operationalized. In addition, detailed feedback is given to the test takers, together with advice on what action, if any, they might be advised to take in order to correct or compensate for weaknesses identified (Alderson et. al, 2015).

4. Related Issues on Reading Assessment

1) The Idea of Content Language Integrated Learning (CLIL)

Different L2 reading tests likely measure students differently. This is not a new issue, but needs to be explored more explicitly and systematically in L2 reading contexts.

Although reading involves both linguistic and conceptual operations, far less attention has been given to conceptual skills in L2 reading assessment (Koda, 2013). Because the current practice focuses on linguistic knowledge as the primary source of reading difficulty, it mandates little provision for distinguishing reading ability and linguistic knowledge. L2 reading assessment tends to center on comprehension of conceptually neutral texts that requires understanding the meaning and construction of familiar content. Since the reading to learn capacity is rarely incorporated in L2 reading assessment, it would be useful to explore ways of measuring more conceptual manipulation skills necessary for learning from unfamiliar texts (Koda, 2013).

2) Learner Differences

L2 readers are not a homogeneous group and they bring different backgrounds when reading L2 texts. They vary in many ways such as cultural experience, topic interest, prior knowledge, and disciplinary knowledge. More attention should be paid to issues of individual variation in order to compensate for this, especially in classroom based assessment (Grabe & Jiang, 2014).

As Grabe (2009) claims, there is certainly a need for good, well-developed low-to-medium stakes reading tests that can be used in a variety of classroom contexts, are graded for multiple proficiency levels, are affordable, can support instructional decision making, and allow for the importance of learner differences.

Reading consists of at least two kinds of lower and higher level processes. Lower-level processes include recognizing the sounds of a language through print, word recognition, syntactic parsing, and encoding meaning as propositions. Higher-level processes involve skills and resources (background, topical, and cultural knowledge) strategies, inferences and monitoring of

comprehension. It is assumed that many weak readers rely too much on a situation model of comprehension, based on what little they manage to understand of the text, which can lead them astray when they try to build a text model.

Grabe (2010) points out that readers are not simply bottom-up or top-down readers. Readers have to be both bottom-up and top-down readers at the same time. We need to take it into account that all readers experience difficulty at some point.

3) The cooperation between assessment specialists and reading researchers

The possible need to develop a notion of reading construct varies with growing proficiency. It is clear that different proficiency levels require distinct types of reading assessments, and different texts, tasks, and task types are appropriate at different proficiency levels. In light of these findings, how should reading assessment task and task types change with growing L2 proficiency? Can systematic statements be made in this regard? (Grabe & Jiang, 2014).

Research demands valid and reliable information on an individual's reading ability derived from assessment, while assessment depends on research to construe the very construct of reading ability. While assessment aims at differentiating individuals on the basis of reading ability, research seeks principled ways to explain those differences. In order to benefit from the reciprocity, assessment specialists must keep pace with continually evolving models of reading, and reading researchers must stay informed about the specific task demands—both linguistic and cognitive—in diverse assessment procedures (Koda, 2013).

4) PC and Information Technology Communication (ITC)

As Grabe & Jiang (2014) state, computers and new media are likely to alter the evolution of

reading tasks and tests. While printed media is far from obsolete we must recognize that the ability to read online texts is becoming an important part of the general construct of reading. As a result, more attention should be paid to this area in regards to reading assessment. At the same time the use of computer based assessment will need to be researched and examined. The following six suggestions by Hubley (2012) should be taken into account when we consider the construct of reading.

- a. The Internet has changed the availability of written material, the scope of sources, and even the ways in which people read.
- b. Learning to recognize sites and sources which are useful has become an essential life tool, and as a result, skimming and scanning are no longer primarily classroom strategies.
- c. In addition, other reading subskills such as identifying key words and concepts and separating fact from opinion are now skills found in everyday life, not just the classroom.
- d. Written texts are now often accompanied by graphs, maps, photographs, and video. This requires students to be able to comprehend information across multiple media.
- e. As an international language, English is becoming highly non-standardized and readers now encounter a variety of abbreviated and other forms online.
- f. The challenge for instructors in the future is to assess relevant skills for reading today and their means of assessment.

Furthermore, eye-trackers—the analysis of the process of reading and responding to reading tasks recorded by eye-trackers, which would allow the provision of processing level feedback, will be available. Also, new computerized language examinations employ scoring algorithms that can analyze and evaluate a wide range of specific features in test takers' written and spoken

productions (Alderson, 2015).

Finally, teachers need to be trained more effectively to understand appropriate assessment. Assessment is an indispensable element of effective teaching and learning practices. Ultimately, as Green (2014) says, the value of assessments depends on the people involved in designing, building, administering and scoring them, and those who interpret the results. Good practice in language assessment and finding the most appropriate tools for each purpose, requires continuing commitment and highly professional expertise.

Acknowledgment:

This work was supported by the Grant-in-Aid for Scientific Research (KAKENHI) (C) (24520646). I am grateful to Dr. Adam Murray for his helpful comments and suggestions.

References

- Alderson, J.C., Haapakangas, E-L., Huhta, A., Nieminen, L., and Ullakonoja, R. (2015). *The Diagnosis of Reading in a Second or Foreign Language*. London and New York: Routledge.
- Chapelle, C.A. (2013). (ed.). *The Encyclopedia of Applied Linguistics*. UK: Wiley-Blackwell.
- Coombe, C., Davidson, P., O'Sullivan, B. and Stoyhoff, S.(Eds.). (2012). *The Cambridge Guide to Second Language Assessment*. Cambridge: Cambridge University Press.
- Denise, E. M. and Christison, M. (2011). *What English Language Teachers Need to Know Vol. II*. London and New York: Routledge.
- Grabe, W. (2009a). *Reading in a Second Language: Moving from Theory to Practice*. Cambridge, USA: Cambridge University Press.
- Grabe, W.(2009b). Teaching and Testing Reading. In M.H. Long & C.J. Doughty (Eds.), *The Handbook of Language Teaching* (pp.441-459). West Sussex, UK: John Wiley & Sons.
- Grabe, W. (2009c). Reading on teaching reading. *Annual Review of Applied Linguistics*, 2, 44-69. Doi:10.1017/S0267190504000030
- Grabe,W. and Jiang, X (2014). Assessing Reading. In Kunnan,J.A.(ed.) *The Companion to Language Assessment*. (185-200). UK: Wiley- Blackwell.
- Green, A. (2014). *Exploring Language Assessment and Testing*. London and New York: Routledge.
- Hubley, J. N. (2012). Assessing Reading. In Coombe, C., Davidson, P., O'Sullivan, B. and Stoyhoff, S. (Eds.), *The Cambridge Guide to Second Language Assessment*. Chapter 24. (pp.211-217) Cambridge: Cambridge University Press.
- Koda (2013): Assessment of Reading. In Chapelle, C. A. (ed.) *The Encyclopedia of Applied Linguistics*. (226-234). UK: Wiley-Blackwell.
- Kunnan,J.A.(ed.) (2014). *The Companion to Language Assessment*. UK: Wiley- Blackwell.
- Purpura, J.E. (2014). Assessing grammar. In A. Kunnan, (Ed.), *The companion to language assessment*. Boston, MA: John Wiley & Sons.
- Schmitt, N., Jiang, X., & Grabe, W. (2011). The percentage of words known in a text and reading comprehension. *The Modern Language Journal*, 95(1), 26-43. doi: 10.1111/j.1540-4781.2011.01146.x
- Urmston, A., Raquel, M., & Tsang, C. (2013). Diagnostic testing of Hong Kong tertiary students' English language proficiency: The development and validation of DELTA. *Hong Kong Journal of Applied Linguistics*, 14(2), 60-82.