



Graded reader readability: Some overlooked aspects

AMANDA GILLIS-FURUTAKA
Kyoto Sangyo University, Japan

Most extensive reading programs rely on graded readers as the main source of reading material, especially for learners at lower proficiency levels. There is considerable variation among publishers with regard to the way graded readers are categorized into levels of difficulty. The most common means of categorization is the number of headwords. This paper reports on some initial findings of a 3-year study of Japanese learners of English as a foreign language. Interviews and think aloud protocols carried out with junior and senior high school and university students (N = 83) show that additional factors need to be taken into consideration when assessing the level of difficulty of graded readers. The findings suggest that authors and editors need to pay closer attention to the likely age range of the target readers, cultural issues, use of idiomatic and figurative language, literary devices, illustrations, and plot structure when determining the readability of graded readers.

Gillis-Furutaka, A. (2015). Graded reader readability: Some overlooked aspects. *Journal of Extensive Reading* 3(1). <http://jalt-publications.org/jer/>

One of the main benefits of extensive reading (ER) programs is increased exposure to the target language. It has been demonstrated that learners can improve their reading ability through ER (Robb & Kano, 2013; Beglar, Hunt & Kite, 2012; Al-Homoud & Schmitt, 2009; Grabe & Stoller, 1997; Robb & Susser, 1989). The reading practice that an ER program provides allows second language (L2) readers to build up automaticity (rapid, unconscious processing of the target language), which reduces the burden on their working memory and affords greater capacity

for processing meaning (Grabe, 2011). It is commonly assumed that when EFL learners are reading a graded text that is at their lexical comprehension level (i.e. they can understand 98%-100% of the vocabulary), they are processing it directly in English. Research by the author (Gillis-Furutaka, 2012) revealed that this is not always the case. In fact, Japanese university students reported the need to switch into the first language (L1) frequently to analyze or make sense of graded readers deemed to be at their lexical comprehension level. This finding suggests that the texts were

not readable enough for direct L2 processing to take place and that there may be factors, other than gaps in lexical or syntactic knowledge, which are reducing the readability of graded reading material and requiring the learners to interact with the text in their L1.

The role of automaticity in L2 reading

Second or foreign language (FL) learning can be viewed as information processing. Human beings are considered to have limited information-processing ability and in order to deal with tasks that put a heavy load on our information-handling capacity, we develop organizing strategies (McLeod & McLaughlin, 1986). Learners “master complex cognitive skills by concentrating processing energy on to-be-mastered subtasks, thereby freeing up the system to work on the mastery of other subtasks” (p. 110). When first learning to read, for example, learners need to devote considerable effort to decode letters to comprehend individual words in a written text. Through repeated encounters with the same word, learners develop automatic recognition of the word and when automatic recognition of individual words is achieved, the learner has more processing energy available to devote to other subtasks, such as attending to syntax and semantic information (Grabe, 2009).

For these reasons, automatic word recognition is considered an indispensable component of fluent reading, not only in an L1, but also in an L2 or a foreign language (Akamatsu, 2008). Fluent readers are able to process the orthographical, phonological, syntactic and semantic information of a printed word while processing a stream of words for comprehension, and at the same time, hold the semantic and syntactic information of previous sentences in working memory. This is referred to as lower-level processing (Grabe, 2009). Because higher-level processing (e.g. iden-

tifying main ideas and connecting them with supporting ideas, drawing inferences, and evaluating information in relation to background knowledge) involves more complex multiple activities, readers need to allocate as many cognitive resources as possible to this in order to read with good comprehension. In other words, it is important for L1 and L2/FL readers to reach a stage where the moment they fix on a word, all the corresponding mental associations are automatically retrieved, allowing higher-level processing to be carried out within the capacity of the remaining limited cognitive resources.

Extensive reading is considered to be an effective means of providing a great amount of repeated exposure to high frequency words and syntactic structures, thereby enabling L2/FL learners to develop automaticity (Day, 2011). A vital tenet of ER is that learners should read books that are at their linguistic level, or a little lower, in order to provide these essential opportunities for repeated encounters with frequent basic words and structures. If the lexical and grammatical content of reading material is closely controlled, both lower- and higher-level processing should proceed smoothly. Given that this is not always the case (Gillis-Furutaka, 2012), a closer examination of the ways in which the language of graded readers is controlled may be needed.

Readability and how it has been measured

Readability is generally thought of as the degree of comprehensibility of a text. With regard to the readability of an L2/FL text, it seems reasonable to say that readability can also be characterized as the ease with which learners are able to process the meaning of a text directly in the L2/FL. This is of special importance in graded readers used in ER programs in order to build automaticity, as has been explained

above. It is therefore important to investigate how the readability of graded readers is measured.

Crossley, Allen and McNamara (2011) explained that when materials developers want to simplify texts to provide more comprehensible input to L2/FL learners, they generally use two approaches: a structural and an intuitive approach. A structural approach depends on the use of grammatical structure and word lists that are defined by level and these are typically referred to when writing graded readers. Another form of structural approach is the use of traditional readability formulas such as the Flesch Reading Ease (Flesch, 1948) or the Flesch-Kincaid Grade Level (Kincaid, Fishburne, Rogers & Chissom, 1975). These readability formulas indicate text readability based on factors such as word and sentence lengths and the number of embedded clauses found in the text.

A number of researchers have expressed doubts about the reliability of such formulas. Carrell (1987) explained that text-based factors related to syntactic complexity, rhetorical organization, and propositional density are omitted from readability formulas. In addition, such formulas are based on the assumption that 'we know what comprehension is, that it is a unified phenomenon susceptible to measurement via text/style variables. "Comprehension" is a complex term which not only means different things to different people, but it is a complex concept which covers multiple behavioral and cognitive factors' (p. 27). Carrell went on to point out that:

The real problems with readability formulas lie not only in the textual features they ignore (syntactic complexity, textual cohesion, propositional density, rhetorical structure) but in the fact that they ignore the reader and reader variables such as background knowledge. They ignore the interactive nature of the read-

ing process—the interaction of the reader with the text. (p. 32)

In 1989, the Lexile Framework for Reading was developed. This system aims to match a reader's level with reading texts of a suitable level. The reader's level is measured through reading comprehension tests. The score is reported as a Lexile measure. The readability of a text is measured through software that analyzes characteristics such as word frequency and sentence length and assigns a numerical value, or Lexile measure. Thousands of websites and articles as well as fiction and non-fiction books have been assigned a Lexile measure and the system is widely used in the United States education system. There have been criticisms of this widespread adoption of the Lexile measure. Krashen (2001) argues that such a system is unnecessary. Readers can quickly determine whether a text is comprehensible and interesting by spending a few minutes sampling it for themselves. Moreover, a strict application of the Lexile Framework can lead to the limitation of reading choice for learners. Hiebert (2009) pointed out that this new readability formula, like its predecessors, fails to take into account the ways in which differences in text structures and features of genres, as well as the background knowledge of the readers, can influence text difficulty and readers' comprehension. She also demonstrated how the overall Lexile measure for a textbook or novel does not reflect the variability that can occur across individual parts of a text.

Within a single chapter of *Pride and Prejudice*, for example, 125-word excerpts of text (the unit of assessments used to obtain students' Lexile levels) that were pulled from every 1,000 words had Lexiles that ranged from 670 to 1310, with an average of 952. The range of 640 on the [Lexile Scale] represents the span from third grade to college". (p. 8)

Another unreliable aspect of Lexile measures is that the proportion of low-frequency vocabulary is typically less when 200,000 words are measured than with selections of 2,000 or 200 words, and it is in the short texts of beginning reading instruction that Lexile measures are more likely to be variable and inflated (p. 9). A further criticism of the Lexile Framework is that it fails to take into account the frequency of the morphological unit, resulting in the inflation of the Lexile measure of informational texts that have a high incidence of morphologically derived words (p. 9). The Lexile Framework, however, has been developed for L1 learners and educational materials rather than for English as a Foreign Language (EFL) learners and EFL materials. For reasons outlined above, it is not a suitable tool to measure the readability of graded reader texts.

The most recent research into factors that can lead to increased text comprehension by EFL learners (Crossley, Yang, & McNamara, 2014) took into account not only linguistic differences in text conditions, but also the participants' language proficiency, reading proficiency and background knowledge. In this research, the readability of authentic texts was compared with texts simplified by the researchers to a beginner and intermediate level. The researchers measured both text comprehension and reading speed as indicators of processing ease. Their results demonstrated that three main factors can increase text comprehension: i) cohesion (i.e., texts with more noun overlap, more causality and more semantic similarity) ii) lower lexical diversity (i.e., texts with more familiar words, and more meaningful words) iii) syntactic simplicity (i.e., texts that are easier to parse). They also concluded that simplification benefits beginner level readers more than intermediate to advanced level readers. Although they acknowledged that "a more fine-grained analysis" (p. 109) is still needed, these findings pro-

vide further evidence for the importance of close attention to these factors and the need for consistency when writing materials for low-level L2 readers.

With regard to measuring the readability of EFL materials written for ER programs, Claridge (2012) drew attention to the lack of consistency between publishers. She highlighted the wide variation in categorizing the level of reading material among the graded reader publishers who appear to be following a structural approach based on word and structure lists, and demonstrated how there are surprisingly large differences in the numbers of headwords deemed appropriate at each level, and how each publisher has a slightly different way of dividing the headwords. Although the Common European Framework of Reference (CEFR) should have simplified life by providing a benchmark, books that have 400 headwords are classified as A1/A2 by one publisher, while another has put books with 1100 headwords in its A2 list. Wan-a-rom (2008) also found that although publishers base their word lists on Michael West's (1953) general service list, they do not always conform to it, nor do their own lists match up.

Claridge (2012) also showed that there are different perceptions among publishers of how a text for non-native speakers must either be adapted or conceived in order for it to provide comprehensible input for the learner. She made the point that whether learners are able to read the texts from the stance of a pleasure reader depends partly on the quality of the in-built scaffolding provided by the lexical, syntactical, and information controls. The level of language is critical in this case as it must be easy if it is to be enjoyed, rather than worked at (p. 117).

In contrast to the structural approaches to simplifying graded reading material, an intuitive approach is subjective, depend-

ing on the author's experience and sense of the comprehensibility of a text, and the discourse processing ability of the target readers (Lotherington-Wolosyzn, 1993). Although both structural and intuitive approaches are common in the development of reading materials for L2 learners (Bamford, 1984), it seems that an intuitive approach is the most common strategy in L2 text simplification. Crossley, Allen and McNamara (2011, 2012) draw this conclusion based on research by Young (1999), Simensen (1987), Blau (1982), and Carrell (1987), who found that even when authors follow a structural approach and refer to word and structure lists, most nevertheless rely mainly on their intuition. This raises questions about the reliability of an author's intuition and the degree of consistency that can be expected between authors and publishers. Furthermore, no mention is made of how extensively the materials are piloted with their targeted EFL readers before publication.

In response to these issues, Simensen (1987) investigated the guidelines for adapted readers (texts that are adapted from authentic texts for EFL/ESL learners) from six publishers, which published a total of nine adapted reader series. She interpreted the principles of adaptation as principles of control and there are three categories: control of information, control of language, and control of discourse and text structure. She pointed out that these areas of control often overlap; they are not clear-cut. She also explained that "control" does not mean rules to be followed as much as suggestions to the adaptor. For example, on the matter of control of information, subplots, characters and descriptive passages seen as unnecessary to the main plot are often deleted. Publishers are also concerned that the adaptor keep in mind the background culture of the text and the readers, and so background notes are included. Alternatively, a setting or character might be substituted for some-

thing more familiar to the readers.

Of particular relevance to this current research are Simensen's findings that adaptors are warned against making the information of the adapted text too compressed or condensed and they are advised to avoid a mere synopsis of action. Information should be evenly distributed. A second kind of "load" in a text is where inferences need to be made. It is considered preferable to spell out the presuppositions even if it makes the adapted text longer.

Other advice concerns control of language. Simensen listed 11 aspects of language to be avoided (p. 50-51). Five of these are relevant to this research. They are:

- Idiomatic expressions
- Ambiguity
- Ambiguity in pronominal references
- Difficult stylistic sequences
- Unexplained allusions and figurative uses

Of an additional 14 aspects of language to take care with (p. 51), the six listed below are relevant to this research:

- Colloquialisms
- Unusual expressions
- Allusions
- Pronominal references
- Distance between pronoun and antecedent
- "the order of events" in a sequence of sentences/arrangement of events in chronological order

Simensen reported that: "In general, publishers stress the importance of verbally explicit language, i.e. language with ample surface clues. The aim seems to be to make the interpretation as easy and unambiguous as possible for the foreign language reader" (p. 52). She went on to say that two publishers were concerned with the organization of information in a text.

One publisher suggested that when a text is organized in accordance with “the order of events”, it is less likely to be a problem for the reader. For the intermediate level, another publisher suggested that events should be arranged “in chronological order” and that “too much time switching” should be avoided (p. 52). Although 22 of the 24 texts used in this study were original stories written for graded reader series and not adapted from authentic texts, the recommendations that Simensen outlines are equally appropriate to both original and adapted graded reader material. Nevertheless, they seem to have been overlooked sometimes by the publishers whose texts were used in this study.

Background to the current study

Extensive reading was first introduced in the English Department at Kyoto Sangyo University in 1987. In 2009, the ER program was expanded to include an additional 2,500 first-year students not majoring in English. ER became a required component of their mandatory English program and comprised 20% of their final grade. Books were borrowed mainly from the library and students took quizzes using an online software program called MoodleReader (Robb, 2010) to accumulate words read. The targets for the number of words read were set according to the English level of the students, which was determined by a placement test at the start of year.

My colleagues and I have been collecting data about the ER program since 2000. Results of a previous research project (Gillis-Furutaka, 2012) showed that many students recognized the positive effects of ER, but others were struggling to read the required number of words. These students reported that they were translating into their L1 extensively as they read, or stopping and thinking about what they had read in their L1. This finding has led to a 3-year study (funded by the Japanese

Ministry of Science and Education) to investigate more closely the ER experiences of Japanese EFL students from beginner to advanced level. The present study is based on the findings of the first two years, which focused on Japanese university and secondary school students. The research question that this paper explores is:

What aspects of graded reader texts, other than lexis and syntax, impede comprehension by Japanese EFL readers?

Method

This research has used a qualitative approach based on interviews and retrospective think-aloud protocols using graded reader texts.

Participants

In the first year, interviews and an adapted form of think-aloud protocols were carried out with university student volunteers (N = 30) from Kyoto Sangyo University. The same interview and think-aloud protocol procedure was carried out the following year with junior high (N = 26) and senior high school students (N = 27). The students came from two private and one public junior high school, and two private and one public senior high school. Eight of the junior high school students were participating in an ER program and eighteen of the senior high school students had participated in an ER program.

The university students were almost all unknown to me, and I was not teaching any of them at that time. There were 17 males and 13 females. They belonged to the following departments: Science (1), Life Sciences (1), International Culture (1), Business Management (2), Economics (2), Law (5), Foreign Languages (18). Only three of the Foreign Language Faculty students were English majors. The junior and senior high school students also volun-

teered, and they were all unknown to me. I aimed to interview three students from each of the three years, but one junior high school student missed her appointment and third-year students were too busy with entrance exams at the third school, so I interviewed two extra second-year students. There were 9 junior high males and 17 junior high females, but the ratio was more balanced among the senior high students with 14 males and 13 females. The first-year junior high students were in their second semester of required formal English study. All of them had experienced informal and irregular English lessons in the final year or two of elementary school. Some of them had attended English conversation schools when younger and/or were taking supplementary English classes outside school. Overall, they were at beginner level. The senior high school students were in their third to sixth year of formal English study and ranged from beginner or low intermediate to advanced level.

Interview method

The participants were interviewed individually in the privacy of my office or a designated room in the school, and with the approval of the university's ethics committee and the principal of each school. Each interview lasted roughly one hour and the students could choose which language(s) to use. The participants were compensated for their time and assured that their identity would be kept private at all times. The interviews were recorded and transcribed later by me.

They were asked about their reading habits in the L1 and L2, their experience of and attitude to ER if relevant, their use of the L1 when reading, and about their goals for their English studies. Next, they were asked to read aloud from the opening pages of different levels of graded readers that were not in the ER program

at their school or the university. This was a rather rough measure, but enabled me to quickly determine their general reading fluency level. There was a good spread of reading levels represented in the university sample, with the majority of students falling into the middle range of low intermediate to intermediate (CEFR A2–B1). There was also a range of levels among the junior high school students, depending on the age at which they had started learning English before starting mandatory English education on entering junior high. Many were still at the beginner level, but some were low intermediate (CEFR A1, A2). The range of levels among the high school students varied the most because several were academically gifted and likely to enter university courses at a higher level than the students at my own university. They ranged in level from CEFR A1–C1.

Next, a reading comprehension exercise was administered using, for the university students, the opening pages (about 700 words) of one of the four graded readers (CEFR A1, A2, B1, B2) deemed to be an appropriate level. For the junior high students, a wider range of reading texts was used, starting with 75 headwords through to 1300 headwords (CEFR A2/B1.) The senior high texts ranged from 250 headwords to 2,500 headwords (CEFR C1). These are the publishers' stated levels and headword counts. The very low level books only have the headword count because they are below the lowest CEFR level (A1). Some of the publishers use both number of headwords and CEFR level, and where they do, I include both, as can be seen in the tables. The students took between 8 and 20 minutes to read the texts, depending on their reading speed. They were also asked to write the answers to five comprehension questions in either English or Japanese.

The next stage of the interview included close scrutiny and discussion of the text

they had read. First, they were asked to summarize orally what they had read in Japanese or English. Next, they were asked to mark with a highlighter pen the places which were difficult to understand and where they had thought or translated into Japanese. They were invited to explain the reasons why. To my knowledge, this form of retrospective think-aloud protocol has not been used to research extensive reading. Think-aloud protocols have been used by many researchers of L2/FL reading skills in the past (Davis & Bistodeau, 1993; Champot, Kupper & Impink-Hernandez, 1988; Hosenfeld, 1984; Kern, 1994; Vann and Abraham, 1990), but think-aloud protocols in which the readers verbalize their thoughts in real time as they read are more suitable for investigating strategies used

and difficulties encountered when reading a difficult text for detailed understanding.

They do not seem suitable for researching the ER experience when students are reading easy texts fast and for general comprehension.

This more relaxed retrospective approach seemed more suitable. It also did not require any prior training of the participants. Below are three tables summarizing the reading comprehension materials used. Although graded reading material from only eight different publishers was included, it was taken from some of the major graded reading publishers and is representative of many graded reader library collections.

Table 1.

Materials used for the reading comprehension for the university students

CEFR Level	Publisher	No. of students
A1	A	4
A2	B	10
B1	B	10
B2	B	6

(None of the texts were in this University ER program at that time.)

Table 2.

Materials used for the reading comprehension for the junior high students

Headwords	CEFR Level	Publisher	No. of students
5		C	4
100		C	1
100		D	2
150		C	1
150		D	1
220		E	7
400	A1	E	5
700	A1/A2	C	1
	A2	E	3
	A2/B1	F	1

(Alternative 100 and 150 headwords texts by publisher D were used for students in an ER program that uses texts by publisher C)

Table 3.

Materials used for the reading comprehension for the senior high students

Headwords	CEFR Level	Publisher	No. of students
250		D	1
300		F	0
400		E	1
550	A1	C	3
900	A2	C	5
1,200	A2	H	3
	A2	B	3
1,600	B1	C	5
1,900	B2	C	2
	C1	B	4

(Alternative A2 text by publisher B used for students in an ER program that uses texts by publisher H)

The aim of this paper is not to single out individual publishers for criticism, but to use these examples to address issues related to readability that affect all publishers. For this reason, the publishers are not identified by name, but by a single letter of the alphabet.

Results and Discussion

As explained above (Simensen, 1987; Claridge, 2012), graded readers are written according to strict guidelines, especially with regard to semantic (measured by the number of headwords) and syntactic complexity. Although some unknown vocabulary items, difficult verb tenses, or embedded clauses impeded comprehension, additional factors that caused confusion were:

- Illustrations
- Cultural differences
- Pronouns and their referents
- Idiomatic and figurative expressions
- Onomatopoeia
- Inferences and other literary devices
- Unexpected changes in the flow of the narrative

Illustrations

Although not discussed in the literature related to L2/FL text readability, illustrations are extremely important as an aid to comprehension, especially when the number of headwords that authors can use is as low as 75. Illustrations allow the authors to use words such as “snowboards” (75 headword text by publisher C) or “elves” (100 headword text by publisher D), and “janitor” (220 headword text by publisher E) that fall outside their list, and give them greater scope to create more interesting stories. The lowest level students demonstrated how they were relying on the illustrations to work out the meanings of unknown vocabulary items and to follow the story line. Here is one example taken from an interview with a first-year junior high student (R). I am A and the Japanese in italic font has been transcribed into romaji. The English translation follows.

A: ... *Hai* ... “Elves” *wakatta?* (... Yes ... Did you understand elves?)

R: *Mm.* (Yes)

A: *Do iu fu ni wakatta?* (How did you understand what they were?)

R: *Do iu fu ni? Do darou? Nan ka atarashii tango ga dete kite koko ni "little" ga kaite kara kono picture wo mitara chichai koto wo mitsukete wakarimashita.* (How did I know? How was it? There was a new word and written here is "little" and looking at the picture I found something little and I understood it was the elves.)

In fact, it seems that publishers may wish to think again about the levels at which they stop including illustrations. Publisher B does not include illustrations at CEFR level B1, but it seems advisable to do so if they are using narrative devices that do not state explicitly the point of view from which a scene is being described. An example that came up frequently in the university student interviews was the opening sentence of the CEFR B1 text. The narrative is told in the third person, but the scene is described from the point of view of the main character without this being stated directly. The story begins: "The train was really flying along now. The buildings, fields and trees all seemed to race by." The readers have to infer that the scene is described as moving because this is what can be seen from inside the train carriage. Rather than creating the intended dramatic effect, it created confusion. The inclusion of an illustration of the young man looking out of the train window would have helped the students to understand the situation immediately. Perhaps the publisher considered that readers at this level no longer need illustrations. It would, nevertheless, be wise to use illustrations when the narrative style may cause confusion for the readers.

Although illustrations can aid comprehension and allow readers to 'see' the action of the story in their mind's eye, some illustrations in the texts used in this research created confusion. For example, in the 75-headword text by publisher C, there are two sisters. The reader is not told that

they are sisters until the very end of the story, and even then, this is not explained explicitly. The junior high school students who read this story understood from the illustration of the older sister that she was the little girl's mother. It was rather puzzling for them that the young man in the story was clearly attracted to her and very surprising that she was his new classmate at the end of the story.

Another example is the CEFR level A1 text from publisher A used with the university students. The protagonists in the story talk about going to stay in "a cottage in the middle of a field". There are two illustrations, but they are very unhelpful. One illustration is of a sizable, two-story house, and a second illustration is of the corner of a field lined by trees. For students from an Asian country like Japan, where fields are very small, crowded together, and often terraced, these illustrations did not help at all because they did not show the isolation of the small country house that is implied by the phrase "a cottage in the middle of a field". Moreover, the students were unable to infer the frustration of the child in the story at the prospect of spending his family holiday in such a lonely place. The isolation of the cottage could have been illustrated effectively in one picture.

Examples of cultural differences

Graded readers are written for students from all over the world and the stories are usually set in one country. It is therefore inevitable that cultural differences will cause confusion from time to time and that writers, illustrators, and editors may be unaware of all possible cultural misunderstandings. Several of the graded readers used in this research project, nevertheless, included examples of cultural confusions that probably could have been avoided.

In the CEFR level A2 text from publisher B used with the university students, two

boys speculate why a mobile phone they have stolen will not work. One says, "Maybe it hasn't got any credit left." In countries where pre-paid mobile phones are common, this would probably not cause confusion. In Japan, however, these days most people pay for their mobile phone service on a monthly basis. If a simpler, more universal explanation had been given for the phone not working, the students would probably not have stopped to try to puzzle out what this expression meant.

Another example is also related to mobile phones, devices that most young people are very familiar with. This is from a higher-level book from publisher B (CEFR level B1). The character is using his phone to look at a photo of his former girlfriend. Readers are told, "Then he chose 'My photos' from the menu on the side of the screen." In spite of their familiarity with mobile phone interfaces, the students interviewed were puzzled that the menu was at the side of the screen. The inclusion of this unnecessary detail caused them to interrupt their reading and doubt their understanding.

Closely related to the problem of catering to readers from a wide range of geographical backgrounds, is the difficulty of matching the level of the language and content to the world knowledge and life experience of the readers. The lowest level book used with a senior high student by publisher D (250 headwords) is written in easy language, but the plot is complicated, requiring a lot of inference by the reader, and the topic of smuggling drugs under a truck in Europe is far removed from the life experience of most Japanese senior high school students. The male senior high student who read this story understood almost all the vocabulary, but could not understand the situation or what the people in the story were doing at all, even with the help of plentiful illustrations. This illustrates

the necessity for writers of graded readers at this low level to make explicit the reasons for the actions of the characters and for their consequences. If they do not, the readers' cognitive resources become overloaded and they are not able to carry out the higher level processing required for comprehension.

A similar problem arose when using books by publisher C that deal with adult topics and viewpoints, such as marriage, house hunting, and pregnancy. Much of the irony, and the need to infer went over the heads of senior high students because these matters were not part of their life experience. Four of the five senior high students who read extracts from the CEFR level A2 book, and three of the five students who read extracts from the CEFR level B1 book experienced this sort of difficulty. It seems that the higher the linguistic level, the more adult the content of many series of graded readers becomes. There is a need for a wider range of content, especially for younger readers at higher levels, and also for clearer indications from the publisher about the targeted age group of their graded readers.

Pronouns and their referents

Learners have to learn to deal with pronouns in the English language. Nevertheless, there is a case to be made for avoidance of pronouns that cause unnecessary confusion in lower level graded readers. There were numerous examples in the texts used in this study. I will discuss just two that caused a lot of confusion. In the CEFR level A2 text used with university students by publisher B, two boys had stolen a woman's mobile phone and run away. While they are hiding, one of them regrets this action and the reader is shown his private thoughts: "Even school would be better than this, though I'd never tell Alex that." The students who read this text could not understand what "that" referred

to and so stopped to analyze the sentence in Japanese. It seems that this sentence is more complex for low-intermediate level readers than it may appear to more fluent English readers because there are two pronouns to deal with ("this" and "that"). The students interviewed could see that "this" referred to the current situation of the two boys, but they were unable to find the referent for the second pronoun "that". If the sentence had used the complete phrase "what I was thinking" instead of simply "that", the meaning would have been far less opaque to these learners and they would have been able to read and understand immediately. Using the pronoun "that" is certainly more natural from the point of view of a native speaker of English, but clarity should surely take priority in a graded reader of this level.

Another example of ambiguity in pronominal reference comes from the CEFR A2 short story by publisher B used with high-school students. The plural pronoun "they" is used to introduce a new, unidentified subject. A wife is referring to the owner(s) of a pub, but this is not made explicit in the text. She says, "Hey, they're a business, aren't they? They understand money all right, you just wait and see." The business owner, identified later in the following paragraph, is only one man, which added to the students' confusion. By replacing the initial "they" with the noun phrase "the owner of the pub" and using the singular form, such unnecessary confusion could have been avoided.

Figurative expressions

Ensuring that the lexical items used fall within the designated headword list for each level of graded reader does not always guarantee that learners will be familiar with their use in idiomatic expressions. As explained above, publishers often warn graded reader writers to avoid phrasal verbs for this reason. There is, however, a

need for vigilance over other kinds of idiomatic and figurative language. Below are some examples that students stopped to puzzle over.

At the lowest level, an apparently simple expression in the 75-headword story by publisher C was opaque to all the junior high students who read it. One of the main characters is described as, "the girl in red." This is an example of metonymy (a word that is used to denote something else related to it.) In this case, the students had to infer that this person was wearing red clothes. Their confusion was compounded by the illustration, in which the only red item of clothing she is wearing is a jacket.

Also incomprehensible to junior high students in the 220-headword story by publisher E is the warning by the school janitor, who had just washed the floor of the gym, "Don't fall over and break something." They all puzzled over what could get broken. They did not understand that "something" meant a bone.

In the slightly higher level 400-headword story by publisher E, third-year junior high school students found the following sentence difficult to understand: "He still couldn't see how to finish." Although using the verb "see" is natural idiomatic English, low-level students found it confusing in this context. Using the more familiar expression, "he still didn't know", would have helped them to read and comprehend more smoothly.

In an example from a story used with university students (CEFR level A2 from publisher B), one of the boys who stole a mobile phone feels guilty about his actions and tells the reader, "I couldn't get the woman's face out of my head." Later, the other boy is excited by the memory of the robbery and the reader is told, "His face came alive as he remembered." Most students puzzled over both of these figu-

rative expressions.

The slightly higher-level text used with university students (CEFR level B1 from publisher B) had many examples. Readers are told, "And then it hit him all over again." Students wondered what had struck the character in the story. They did not understand that "it hit him" refers to a painful memory that he suddenly recalls. If this had been expressed more explicitly, an unnecessary obstacle to comprehension could have been avoided. Moreover, the explanation of what happened is given in the following paragraph. Students did not realize they had to read on to find out what "it" was. Instead, they went back and translated what they had already read into Japanese to try to find what they thought they had missed.

Soon after in the same story, readers are told that when he took out his mobile phone, "... he opened it and watched the screen come to life." Such figurative language is easily understood by native speakers who are familiar with the concept of metaphor. However, it does not seem appropriate for this level of graded reader because this metaphorical meaning was not immediately obvious to the students who read this text.

One more example of an apparently simple lexical item used in an unfamiliar and metaphorical way, from the same text, is when the narrator describes the scene from the train window and tells us, "a bridge shot past the window...." Students associated "shoot" with guns, not bridges, and were confused by this expression.

Onomatopoeia

Although onomatopoeia is a common and effective literary device, its use in a low-level graded reader seems unnecessary, especially as it confused all the university students interviewed when they read,

"The stick made a THWACK sound when it hit the walls" (CEFR level A2 text from publisher B).

Inference

Although L2 learners may be fluent readers in their L1, conversant with narrative perspective and skilled in making inferences, this research showed that the requirement to do so while reading in the L2/FL is a barrier to comprehension.

In the CEFR level B1 story from publisher B referred to above, there are many examples of the need for readers to make inferences in the opening pages, which caused confusion. For example, the main character is thinking about his former girlfriend and the reader is told his thoughts, "So then to get jealous about a dog ...". In this case, the subject of the verb has to be inferred. It is, in fact his former girlfriend, but the students did not understand this. Greater clarity could have been provided by a simple rephrasing of the incomplete sentence.

The sentences that caused the most difficulty in this story were those in which a character or incident is referred to before it has been explained. One example, "And then it hit him all over again," has already been mentioned above because the figurative use of "hit" was problematic. In addition to this, "it" refers to the young man's sadness when remembering that he has recently broken up with his girlfriend. The reader has not been told about the girlfriend at this point in the story, so students wondered what they had missed and some re-read and translated into Japanese to try to find out. This did not help because only later is the reader told, "... the argument had come out of nowhere." Some publishers and instructors may argue that graded readers should prepare L2/FL readers for native speaker literature and that training in inferring meaning is

an important part of such preparation. The question is: At what level should this kind of literary device be introduced? If the main purpose of ER is to increase reading fluency by allowing students to read directly and smoothly in the L2, these interview findings suggest that expecting CEFR B1 level students to be inferring to such an extent is not helpful.

This problem is compounded when writers of very low-level graded readers expect their readers to infer as they read because readers with lower level L2/FL ability need to concentrate to an even greater extent simply to understand the basic, explicit meaning of the text. This was clearly the case for the university students who read the lowest level text (CEFR level A1 from publisher A). The reader is told, "Mr. Carter wants a problem to solve so he is not very happy." Readers need to infer that life is dull for Mr. Carter, but this is not stated explicitly. Readers are only told that he "has an angry face" and "an angry voice". The text is too condensed. More explanation is needed for Mr. Carter's character to be understood.

Similarly, even the lowest level text (publisher C, 75 headword story) expects the readers to infer that "the girl in red" (already mentioned above) is the older sister of the "little girl". If the relationship between the two girls had been explained explicitly at the start of the story, the students would have understood the situation easily.

Unexpected changes in the flow of the narrative

Another aspect related to the style of graded reader narratives that caused confusion was an unexpected change in the course of events. This is an essential narrative device, and one that is central to almost all plots. However, the language learners interviewed tended to doubt their language

skills when something unexpected happened in the flow of the narrative and felt the need to stop and think in their L1 in order to check their understanding. Similarly, misunderstandings arose when events were not related in chronological order and flashbacks were very hard for many students to identify and understand. The interaction of text structure with L2/FL comprehension was investigated by Riley (1993) and reported by Durgunoğlu (1997). In this study of 120 English-speaking college students taking various levels of French courses, all subjects read French texts in structured (i.e., following story grammar) or less structured (episodes with flashbacks interspersed, or episodes completely mixed, as in a modern short story) formats. The results showed that third-year students had equally high levels of recall for all three text types. However, first-year students recalled the most information from the structured text and the least from the completely mixed-episodes text. That is, text structure interacted with the reader's proficiency level.

In my own study, readers of CEFR level B2 and C1 texts were able to understand flashbacks and changes in the flow of the narrative without problems. The examples below are all from the lower level texts and seem to suggest that these features of narrative style are better avoided in graded readers up to and including the CEFR B1 level.

First are some examples from the CEFR levels A2 and B1 texts from publisher B that were read by the university students.

"He threw the stick into the air as hard as he could" (CEFR level A2). Almost all students were puzzled by this sudden, unexpected action that appears unconnected to the preceding scene. The author's intention is no doubt to illustrate the character's impulsive personality. Some form of lead-in expression or explanation about his

motive could have reduced the students' confusion.

"I could think of lots of places I would rather be, like at home waiting for Mum to get back from work" (CEFR level A2). The characters have just run away from a woman whose mobile phone they have stolen and are hiding by a public toilet. The unannounced switch from the third person narrator's description of the action to the character's internal dialogue confused many students.

The three senior high school students who read CEFR level A2 text, a short story by publisher B, all found this story very difficult to understand. This story starts by introducing two of the main characters, an American couple on holiday in the UK. However, the subsequent references to past events and to several characters who had lived in the past made it very difficult for these students to follow the narrative thread. Moreover, they were unable to make the inferences required to understand the personalities of all the main characters and their motivations. Consequently, the twist at the end of the story, despite being illustrated, was lost on all of them. The level of sophistication and complexity of the narrative style of this short story was far higher than the level of the English.

The use of flashbacks is another very common narrative device, but one that presented problems for the younger, less experienced junior and senior high school readers as well as the university students. Only one of the third-year junior high school students had a reading level high enough to be able to deal with the CEFR level A2/B1 text by publisher F. She was exceptional in that she had been in an English immersion program during her elementary school years and had passed the Japanese English proficiency test commonly known as Eiken at level 2, which

is the equivalent to CEFR B1. She read the story quickly and there were only three infrequent vocabulary items that she did not know. She could understand the main story line very well, and could infer information about the characters readily, but was confused by the insertion of a flashback. The story begins at dinnertime, but then the scene changes to earlier that day. The scene then returns to the events that took place after dinner. Although the flashback is signaled in the text ("It all started four hours earlier"), she had assumed that this scene followed the dinnertime events chronologically (i.e., that it took place the next day) and that the return to the same dinner scene was actually the following evening.

It appears that failure to recognize a flashback is not simply related to the age and amount of L1 and L2/FL reading experience of language learners. University students experienced the same difficulty. In the CEFR B1 text from publisher B read by the university students, where the young man is travelling on a train, the sentence already discussed above, "And then it hit him all over again," introduces a flashback. The flashback ends unannounced and the narrative returns to the present situation: "Hiro looked sadly out of the window again. Where were they?" Many students found this abrupt time shift confusing. If flashbacks are used in graded readers, it will help students if they are very clearly signaled.

One last observation in regard to readers having to deal with breaks in the flow of the narrative is when there are references made to previous stories in a series (i.e., to characters and events). This was unnecessarily confusing to the junior high school students who had not read the other stories in the CEFR Level A2 series by publisher E. Although series editors would ideally prefer students to read all the books in their series, they should not

assume that students will read all the other stories, or that they will read them in a certain order if the books are part of an ER program in which student self-selection of reading materials is a criterion.

All but the last of the problematic aspects explained above were shown by Simensen (1987) to be matters that publishers wish their adapted reader writers to pay attention to. It would appear that the same guidelines should be adhered to by writers of graded readers based on original stories.

Limitations

This initial research project needs to be replicated with a wider range of Japanese learners of English as well as with English learners with different linguistic and cultural backgrounds. Also, the number of graded reader texts used was rather small. A wider range of graded reader texts produced by more publishers needs to be used in order to see whether these findings can be generalized. In addition, this paper has focused on findings related to only a few aspects of the linguistic and narrative style of the graded readers used.

Conclusion

This study used a qualitative approach to explore reasons why Japanese secondary school and university students have difficulty in comprehending graded reading material at their linguistic level. The interview and retrospective think-aloud protocol findings suggest that, in many cases, the graded reading material was not easily readable. The reading experiences reported by these students reveal that measuring the readability of a graded reader text is far more complicated than limiting the number of headwords and grammatical structures. Publishers need to pay closer attention to other less quantifiable aspects of readability. These include cultural as-

pects, the age of the targeted readers, literary devices such as onomatopoeia, metaphor, idiomatic language, twists and turns in plot structure and the use of flashbacks. Although some such features can enhance the pleasure of reading in the L1, they impede the processing of a L2/FL text by lower-level learners and so, arguably, they should be avoided, or at least used with greater care, in graded readers below the CEFR B2 level. When students are reading at a suitable lexical level, but are still unable to read and comprehend directly in the L2, they are not getting the practice they need to build up automaticity and improve their overall linguistic skills, and the purpose of an ER program is undermined.

It is very difficult to find the right balance between providing a story that is both exciting to read and easy for readers to follow. Publishers can help maintain this balance by providing plenty of helpful illustrations and by avoiding potential cultural confusion. It should also be borne in mind that L2/FL learners find it harder to make inferences and recognize metaphorical expressions when reading in the L2/FL. Moreover, pronouns can often become barriers to comprehension and so are better avoided when potentially ambiguous.

It is not at all easy for authors, publishers, and even teachers to be able to predict what difficulties learners will encounter when reading a graded reader. Graded reader editors Bowler and Parminter (2011) said that: "The test of a well-written graded reader, we believe, is that a native speaker reading it would not necessarily be aware that the vocabulary and grammar in the story are restricted. They would simply enjoy the story" (p. 36). This study shows that it is unwise to rely only upon the evaluation of native speaker readers. Piloting new graded reading material with readers of the age group and from the cultural backgrounds at which it

is aimed, using a method similar to that used in this research project, would provide a more accurate readability measure and enable publishers to indicate to teachers more precisely the suitability of the graded reading material for their learners. The benefits that could be derived from a more accurate evaluation of graded reading materials could arguably offset the additional costs that would be incurred.

References

- Akamatsu, N. (2008). The effects of training on automatization of word recognition in English as a foreign language. *Applied Psycholinguistics*, 29, 175–193.
- Al-Homoud, F., & Schmitt, N. (2009). Extensive reading and a challenging environment: A comparison of extensive and intensive reading approaches in Saudi Arabia. *Language Teaching Research*, 13, 383–401.
- Allen, D. (2009). A study of the role of relative clauses in the simplification of news texts for learners of English. *System*, 37(4), 585–599.
- Amer, A. A., & Khouzam, N. (1993). The Effect of EFL Students' Reading Styles on their Reading Comprehension Performance. *Reading in a Foreign Language*, 10(1), 967–978.
- Bamford, J. (1984). Extensive Reading by Means of Graded Readers. *Reading in a Foreign Language*, 2(2), 218–260.
- Beglar, D., Hunt, A., & Kite, Y. (2012). The effect of pleasure reading on Japanese university EFL learners' reading rates. *Language Learning*, 62(3), 665–703.
- Bell, T. (2001). Extensive reading: Speed and comprehension. *The Reading Matrix* 1(1). <http://www.readingmatrix.com/articles/bell/article.pdf>
- Bernhardt, E. B., & Kamil, M. L. (1995). Interpreting relationships between L1 and L2 reading: Consolidating the linguistic threshold and the linguistic interdependence hypotheses. *Applied Linguistics*, 16, 15–34.
- Blau, E. K. (1982). The effect of syntax on readability for ESL students in Puerto Rico. *TESOL Quarterly*, 16, 517–528.
- Bowler, B. & Parminter, S. (2011). What is a graded readers series? Series editors' view (2). In R. Day et al. (Contributors), *Bringing extensive reading into the classroom* (pp. 34–43). Oxford: Oxford University Press.
- Carrell, P. L. (1987). Readability in ESL. *Reading in a Foreign Language*, 4(1), 21–40.
- Chamot, A. U., Kupper, L., & Impink-Hernandez, M. V. (1988). *A study of learning strategies in foreign language instruction: Findings of the longitudinal study*. McLean, VA: Interstate Research Associates.
- Claridge, G. (2012). Graded readers: How the publishers make the grade. *Reading in a Foreign Language*, 24(1), 106–119.
- Coady, J., & Huckin, T. (1997). *Second Language Vocabulary Acquisition: A Rationale for Pedagogy*. Cambridge: CUP.
- Crossley, S. A., Allen, D. B., & McNamara, D. S. (2011). Text readability and intuitive simplification: A comparison of readability formulas. *Reading in a Foreign Language*, 23(1), 84–101.
- Crossley, S. A., Allen, D., & McNamara, D. S. (2012). Text simplification and comprehensible input: A case

- for an intuitive approach. *Language Teaching Research*, 16, 89–108. doi: 10.1177/1362168811423456.
- Crossley, S. A., Yang, H. S., & McNamara, D. S. (2014). What's so simple about simplified texts? A computational and psycholinguistic investigation of text comprehension and text processing. *Reading in a Foreign Language*, 26(1), 92–113.
- Davis, J. N., & Bistodeau, L. (1993). How do L1 and L2 reading differ? Evidence from think aloud protocols. *The Modern Language Journal*, 77(4), 459–472.
- Day, R. (2011) Extensive Reading: the background. In R. Day et al. (Contributors), *Bringing extensive reading into the classroom* (pp. 10–21). Oxford: Oxford University Press.
- Durgunoğlu, A. Y. (1997). Bilingual Reading: Its Components, Development, and Other Issues. In A. M. B. de Groot & J. F. Kroll (Eds.), *Tutorials in Bilingualism* (pp. 255–276). Mahwah, NJ: Lawrence Erlbaum Associates.
- Erten, I. H., & Razi, S. (2009). The effects of cultural familiarity on reading comprehension. *Reading in a Foreign Language*, 21(1), 60–77.
- Farrell, T. S. C. (2009). *Teaching Reading to English Language Learners*. Thousand Oaks, CA: Corwin Press.
- Flesch, R. (1948). A new readability yardstick. *Journal of Applied Psychology*, 32, 221–233.
- Gillis-Furutaka, A. (2012). MoodleReader at Kyoto Sangyo University: Why is the program working for some, but not for others? *Extensive Reading World Congress Proceedings*, 1, 40–43.
- Grabe, W. (2009). *Reading in a second language: Moving from theory to practice*. New York: Cambridge University Press.
- Grabe, W., & Stoller, F. L. (1997). Reading and vocabulary development in a second language: A case study. In J. Coady, & T. Huckin (Eds.), *Second language vocabulary acquisition* (pp. 98–122). Cambridge: Cambridge University Press.
- Grabe, W., & Stoller, F. L. (2011). *Teaching and Researching Reading (Second Edition)*. Harlow: Pearson Education Ltd.
- Hiebert, E. H. (2009). *Interpreting Lexiles in Online Contexts and with Informational Texts*. Apex Learning. [http://www.apexlearning.com/documents/Research_InterpretingLexiles_2009-02\(1\).pdf](http://www.apexlearning.com/documents/Research_InterpretingLexiles_2009-02(1).pdf)
- Hosenfeld, C. (1984). Case studies of 9th grade readers. In Alderson, J. and Urquhart (Eds.) *Reading in a Foreign Language* (pp. 231–249). New York: Longman.
- Kern, R. G. (1994). The role of mental translation in second language reading. *Studies in Second Language Acquisition*, 16, 441–461.
- Kincaid, J. P., Fishburne, R. P., Rogers, R. L., & Chissom, B. S. (1975). *Derivation of new readability formulas (Automated Readability Index, Fog Count and Flesch Reading Ease Formula) for Navy enlisted personnel*, Research Branch Report 8–75, Millington, TN: Naval Technical Training, U. S. Naval Air Station, Memphis, TN.
- Koda, K. (2007). Reading and Language Learning: Crosslinguistic Constraints on Second Language Reading Development. *Language Learning*, 57, 1–44.

- Krashen, S. (2001). The Lexile Framework: Unnecessary and potentially harmful. *CSLA (California School Library Association) Journal* 24(2), 25-26. http://sdkrashen.com/content/articles/lexile_framework.pdf
- Lotherington-Woloszyn, H. (1993). Do Simplified Texts Simplify Language Comprehension for ESL Learners? In M. L. Tickoo (Ed.), *Simplification: Theory and application* (pp.140–154). Singapore: SEAMEO Regional Language Center.
- McLeod, B., & McLaughlin, B. (1986). Restructuring or automaticity? Reading in a second language. *Language Learning*, 36(2), 109–123.
- Robb, T. (2010). A digital solution for extensive reading. In *Bringing Extensive Reading into the Classroom*, (pp. 105–110). Oxford: Oxford University Press.
- Robb, T. N., & Kano, M. (2013). Effective extensive reading outside the classroom: A large-scale experiment. *Reading in a Foreign Language*, 25(2), 234–247.
- Robb, T. N., & Susser, B. (1989). Extensive reading vs skills building in an EFL context. *Reading in a Foreign Language*, 5(2), 239–251.
- Simensen, A. M. (1987). Adapted readers: How are they adapted? *Reading in a Foreign Language*, 4(1), 41–57.
- Stenner, A. J. (1996). *Measuring reading comprehension with the Lexile Framework*. https://lexile-website-media-2011091601.s3.amazonaws.com/m/resources/materials/Stenner_Measuring_Reading_Comprehension.pdf
- Tweissi, A. I. (1998). The effects of the amount and the type of simplification on foreign language reading comprehension. *Reading in a Foreign Language*, 11(2), 191–206.
- Upton, T. A., & Lee-Thompson, L. (2001). The role of the first language in second language reading. *Studies in Second Language Acquisition*, 23, 469–495.
- Vann, R. J. and Abraham, R. R. (1990). Strategies of Unsuccessful Language Learners. *TESOL Quarterly*, 24(2), 177–198.
- Wan-a-rom, U. (2008). Comparing the vocabulary of different graded-reading schemes. *Reading in a Foreign Language*, 20(1). <http://nflrc.hawaii.edu/rfl/April2008/wanarom/wanarom.html>
- West, M. (1953). *A General Service List of English Words*. London: Longman, Green and Co.
- Young, D. J. (1999). Linguistic Simplification of Second Language Reading Material: Effective Instructional Practice? *The Modern Language Journal*, 83(3), 350–366.