広島文教女子大学高等教育研究　1，2015

【原著】

Increasing Comprehensibility of Language Classroom Learning Material Instructions

Arthur Rutson-Griffiths and Judith Runnels

外国語教室の教材指示についての理解度の向上

Arthur Rutson-Griffiths and Judith Runnels

Abstract

The comprehensibility of written task instructions from language classroom learning materials is an important area to examine since language learners need to understand what is expected of them in order to perform classroom tasks successfully. This study describes how teacher-written instructions on classroom materials developed for a General English curriculum were analyzed and modified to improve comprehensibility. Written instructions were modified to simplify vocabulary, reduce sentence length, eliminate extraneous information, and reflect the sequencing of task performance. An analysis using three readability formulas showed that readability increased for all indices. Both the pre and post-change task instructions were then rated by language learners for comprehensibility; Rasch analysis and ANOVAs comparing the comprehensibility ratings on the pre- and post-change instructions revealed that 78% of the instructions were rated as easier to understand than their pre-change equivalent, illustrating that the modifications were effective at increasing both readability and comprehensibility.

Introduction

There are numerous factors that have been shown to increase the difficulty of texts for second language (L2) learners. The role of vocabulary in comprehending texts is one such area which has been extensively researched. Studies have suggested that there is a strong correlation between vocabulary knowledge and reading comprehension, with an understanding of 95−98% of the vocabulary in a text being necessary for unassisted comprehension (Hirsh & Nation, 1992; Hsueh-Chao & Nation, 2000; Laufer & Ravenhorst-Kalovski, 2010). Similarly, grammatical structures that are beyond readers' comprehension impair their understanding of a text (Shiotsu & Weir, 2007). The length of sentences also affects comprehension, with longer sentences being more difficult to understand (see Koda, 2005). Factors that increase the difficulty of a text go beyond the level of vocabulary and grammar however, including also the structure of the text and characteristics unique to the learners themselves such as confidence, motivation and prior learning experiences (Brindley, 1987 as cited in Nunan & Keobke, 1995). For instance, texts that are written in paragraphs in continuous form (rather than using numbers or bullet points) affect

learners' perceived difficulty of even easy texts (Nunan & Keobke, 1995).

While these studies focused on the text of the task itself, the findings likely also apply to the instructions given to learners to introduce the learning task. In fact, the relationship between successful completion of a task and the quality of given instructions on how to complete the task has been well-researched within the field of healthcare (see DeWalt, Berkman, Sheridan, Lohr & Pignone, 2004; DuBay, 2004). For the field of education however, and particularly for language learning, comparatively less research has been performed. Indeed, further exploration of the comprehensibility of task instructions on language classroom learning materials is required since "when texts exceed the reading ability of readers, they usually stop reading" (DuBay, 2004, p. 1). This finding particularly concerns language teachers who employ a task-based teaching and learning (TBL) approach in their classroom; given that in TBL tasks are used as "the core unit of planning and instruction" (Richards & Rodgers, 2001, p. 223), if learners fail to comprehend how they are supposed to complete a task, they are also failing to engage with the core instruction in a lesson.

Measuring readability

Readability refers to the ease with which a text can be read. It is typically gauged using readability formulas, which give an estimate of the reading level required to comprehend a text. Many studies of comprehensibility of written information have used readability formulas as a research tool, in fields such as medicine (Wegner & Girasek, 2003) and education (Gallagher, Fazio & Gunning, 2012). However, very few of these studies focus on the readability of a foreign language, and are often reflective of the abilities of the general population reading in their native language. Indeed, it has been suggested that some of the most common readability formulas that are used to analyze texts in a learner's native language (L1) are not appropriate for analyzing texts in the learner's language of study (L2); they do not account for textual factors such as syntactic complexity and organization of arguments and do not closely align with the cognitive processes involved in second language reading (see Crossley, Allen & McNamara, 2011). Studies have found mixed results, both negative (Brown, 1998) and positive (Greenfield, 1999, cited in Crossley et al., 2011), for how well readability formulas predict L2 reading comprehension. Additionally, many readability formulas are trialed and tested on long passages of text (McLaughlin, 1969), and the accuracy of such formulas is thus unclear when used on written instructions of only a few sentences. In other words, even if readability is improved according to any index, this may not necessarily be associated with increased comprehensibility by actual users of the text. Using formulas to analyze readability may not be sufficient for determining whether instructions are indeed easier to understand by the target population of users, and comprehensibility judgments from intended users of the texts should also be included in readability or difficulty analyses.

The current study was thus designed to examine how making changes in vocabulary, syntax, sequencing and formatting affected general readability as well as student comprehensibility on task instructions from classroom learning materials from an English as a Foreign Language class.

All task instructions were developed by teachers who had created the learning materials for classroom usage. It should be noted that the pre-change instructions were not understandable or unusable; in fact, in their pre-change form, the instructions and their associated tasks had been previously employed for at least two academic years prior to the current study. It was only when a teacher who had not participated in the development process remarked that the instructions appeared to be more difficult to understand than the task that the need for this study became apparent. Essentially, following modifications of the existing task instructions according to vocabulary, syntax, sequencing and formatting, readability analyses were performed on both versions of the instructions to test for any improvements according to the readability formulas. Both pre- and post-change instructions were then administered to students for the purposes of obtaining difficulty or comprehensibility ratings for comparison. It was expected that the post-change instructions receive both higher readability formula scores and higher comprehensibility ratings from learners.

Methods

Participants

89 first year students from a small, private university in Japan participated voluntarily in this study. Participants were in one of five disciplines of study: Early Childhood Education, Welfare, Nutrition, Psychology and Global Communication. The survey was administered during the second semester of the academic year, meaning that students of the former four majors had completed at least one semester of twice-weekly 90-minute English classes at the tertiary level. The Global Communication majors (a total of 27% of participants) had completed one semester of full-time English tuition. All participants were familiar with all of the instructions on the handouts as they had encountered and interacted with them throughout the duration of their study.

Instrument

Table 1 illustrates examples of how the task instructions were modified according to the findings of previous research. The main changes were to simplify vocabulary, reduce sentence length, sequence the instructions according to the order of performance, and format the instructions into numbered points rather than in a paragraph.

Table 1. Examples of task instruction modification

Example	Type of change	Pre-change instruction	Post-change instruction
1	Vocabulary	Choose the appropriate word.	Choose the best word.
2	Brevity	Look at the words below. Try and match them with their Japanese translation. The first one has been done for you.	Look at the words below. Match them with the Japanese translation.
3	Sequencing and formatting	Match the vocabulary cards that your teacher gives you.	1. Your teacher will give you vocabulary words. 2. Match them.

Vocabulary was simplified initially according to teachers' intuition and the English Vocabulary Profile (2012) was also used to assist decision-making. Vocabulary was also repeated wherever possible. For example, instructions asking students to work with a partner were always written as "Find a partner" rather than "Work with a partner" or "Find someone to work with" in the post-change versions of the instructions. To simplify and shorten the instructions any extraneous information was removed, even if the content was intended to help or guide the students (Table 1, Example 2). Sentences were also shortened as much as possible by breaking sentences with more than one clause into separate sentences. For sequencing and formatting, instructions were re-written in the same order that students were expected to perform them, and numbering was used to break up blocks of instruction text with the intention of making them easier to follow (Table 1, Example 3).

A total of nine instruction pairs were selected as representative of all classroom material instructions for inclusion on the student survey. Each pair consisted of the pre-change and post-change versions of the same instruction (see Appendix 1 for all included instruction pairs).

Procedure

The pre-change and post-change versions of the instructions were administered to participants online. Participants were asked to imagine they were reading instructions in a lesson handout and to indicate the extent of their understanding of how to perform the task on a four-point Likert-esque scale. A total of 18 instructions were presented in random order, where nine of the items consisted of pre-change instructions and nine of its post-change equivalent. A four-point scale was employed due to Japanese tendencies to select a neutral response if one is included in the survey response scale (Dornyei & Taguchi, 2010).

Analysis

Two readability formulas were applied to the pre- and post-change instructions. The formulas employed were the McAlpine EFLAW Readability Score and the Flesch Reading Ease. The McAlpine EFLAW Readability Score was developed to measure texts intended for L2 readers (McAlpine, 2007). The calculation is based on the average number of words and "miniwords" (of three characters or less) per sentence. McAlpine argues that miniwords are difficult for L2 learners because their many meanings are often difficult and time-consuming to decipher from a dictionary, and often cluster in areas of difficulty such as wordy clichés, colloquial expressions and phrasal verbs. The lowest possible score is 1, with anything over 30 considered confusing. The Flesch Reading Ease formula uses the average number of words per sentence and syllables per word to calculate a reading score (Flesch, 1948), and is one of the most reliable and well-used readability formulas (DuBay, 2004). It is measured on a scale from 1-100, where ease of readability increases with score. Although some researchers have raised doubts over the validity of using readability formulas for L2 learners (Nation, 2001), in this study they provided a useful initial measurement of pre- and post-change instructions to indicate whether improvement had been achieved.

Descriptive statistics were also calculated for all pre- and post-change instructions, including the

total number of words and the numbers of hard words, different words and miniwords. The number of sentences and average number of words per sentence were also calculated. Winsteps® Rasch software Version 3.72.4 (Linacre, 2010) and SPSS, Release Version 18.0.0 were used to analyze the results. To determine the difficulty of each instruction, the Rasch measure was calculated for all instructions. The Rasch measure is the probability of a person correctly responding to a given item and is related to their ability and the difficulty of the item (for a detailed overview of Rasch analysis, see Bond & Fox, 2007). A Rasch measure is calculated for each instruction in logits (the units of the Rasch measure) and is calculated with the following formula:

$$\log[pni/(1-pni)]=Bn-Di(1)$$

where Bn is the ability of a person n and Di is the difficulty of item i (from Beglar, 2010).
A 0.3 logit difference between each instruction is required for a significant difference in difficulty (Miller, Rotou & Twing, 2004; Lange, Greyson, Houran, 2004). ANOVAs were also performed to measure significant differences between pre- and post-change instructions.

Results

Descriptive statistics for the pre-change and post-change instructions are shown in Table 2. All post-change instructions underwent a decrease in difficulty of at least 7.2% and up to 65.4% according to the listed criteria (Table 2). Following the changes, the overall number of words and the frequency of hard words were reduced by over two-thirds. The number of different words and miniwords also fell significantly. The overall number of sentences was not greatly reduced, but the average sentence length of the post-change instructions was less than half the length of the pre-change instructions. The readability formulas employed also show improvement; the McAlpine EFLAW Readability score was more than halved, and the Flesch Reading Ease score increased to from 81.4 to 87.3.

Of greater importance, however, is whether participants, the actual users of the instructions, gave the post-change instructions lower difficulty ratings. An ANOVA confirmed that the post-change instructions (M=3.38, SD=1.3) were rated as easier than the pre-change instructions (M=3.13, SD=1.2; F=4.06, $p<0.05$). This was also found for the Rasch analysis, whereby the mean difference in difficulty between pre-change and post-change instructions was 0.80 logits, thus demonstrating a significant difference in difficulty between instruction types.

Table 2. Descriptive Statistics for Pre- and Post-Change Task Instructions.

	Pre-change	Post-change	Difference	Change (%)
Total words	2434	867	1567	−64.4%
Hard words (3 or more syllables)	159	55	104	−65.4%
Different words	453	265	188	−41.5%
Miniwords (3 or fewer characters)	923	402	521	−56.4%
Sentences	277	253	24	−8.7%
Average words per sentence	8.9	4.0	4.9	−54.6%
McAlpine EFLAW Readability Score	12	5.6	6.4	−53.3%
Flesch Reading Ease	81.4	87.3	5.9	+7.2%

Figure 1. The difference in difficulty between pre-change and post-change instructions for each instruction pair (in logits). A positive measurement indicates that post-change instructions were rated as less difficult than pre-change. The horizontal black line represents the 0.3 logit difference required for significance.

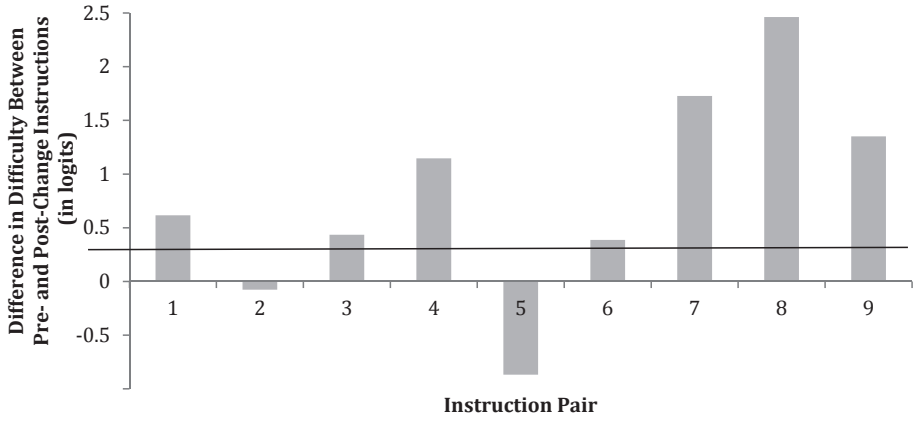

Figure 1 shows the difference in Rasch measures between the pre-change and post-change instructions where a difference of 0.3 logits represents a significant difference for difficulty (and is indicated as a horizontal solid line within the figure). Seven of the nine post-change instructions were rated as significantly easier to comprehend than their pre-change equivalent. For instruction pair number 2, no significant difference in pre- and post-change instructions was found. In Figure 1, it is also evident that one of the instruction pairs (5) exhibited a negative difference: this means that the pre-change instructions were rated as less difficult than the post-change instruction. Further exploration of these instruction pairs is perhaps required.

Discussion

In the current study, participants found that 78% of written task instructions which had been modified to include simpler vocabulary, simplicity and brevity in sentence structure and greater accuracy in sequencing less difficult to understand than task instructions that had not been modified. However, there were two instruction pairs that did not receive higher comprehensibility ratings from participants. In the case of instruction pair 2 (Table 3), there was no significant

Table 3. Instruction pairs with negative changes

Pair	Pre-change Instruction	Post-change Instruction
2	Work with a partner and role-play the conversation below.	1. Find a partner. 2. Role-play the conversation below.
5	Now, make a group of three. Where is the best place for a honeymoon? Rank (from 1−10) the destinations in Japan below.	1. Make a group of three. 2. Rank (from 1−10) the honeymoon destinations. 3. Write a reason for each destination.

difference between the pre- and post-change instruction, likely because the difference in the instructions is seemingly very small: the post-change instruction eliminated a total of two words but the number of characters or overall length remained essentially the same. Furthermore, participants were likely already familiar with the expression “Work with a partner”, having come across it numerous times throughout the semester of English classes. It was therefore rated as easy to understand. This finding raises the question of familiarity with instructions, and that in addition to vocabulary, brevity and accurate sequencing, familiarity with the instruction may also impact comprehensibility and subsequent task performance.

This is unlikely the case however, for instruction pair 5 (Table 3) whereby the post-change instruction was rated as more difficult than its pre-change equivalent. In this instruction pair, the pre-change instruction provides more contextual information. In the participant survey, all instructions were presented in isolation, and in this instance the task to be completed is fairly evident in the pre-change instruction. Furthermore, the frequency of hard words remains the same across both instructions (honeymoon, destination) as does potentially challenging vocabulary (rank). Lastly, the post-change instruction even includes an additional task (to write a reason for each destination). This highlights the importance of testing for difficulty, because teacher perceptions or estimations of difficulty may not be in agreement with those of the language learners, as was shown by the difficulty ratings on pair 5 (Table 3). Despite that the instructions had previously been used in the classroom and that the majority of the changes successfully increased the comprehensibility of the instruction, an oversight meant that the post-change instruction was not modified appropriately. This result highlights that teachers should constantly be mindful of keeping instructions as simple and succinct as possible and that even after modifications, subsequent checks or tests should be performed to ensure greater comprehensibility of the instructions.

Conclusions

Although it appears that the recommended guidelines (derived from the work of Hirsh & Nation, 1992, Koda, 2005 and Nunan & Keobke, 1995) were successful in increasing the comprehensibility of the task instructions, differing results may have been found if the instructions had been presented in the context for which they were designed. Furthermore, whether easier to understand task instructions indeed led to higher performance on the tasks or which of the four main changes (vocabulary, brevity, sequencing and formatting) had the greatest impact on comprehensibility were not examined. Familiarity was also suggested to impact comprehensibility. Further research should certainly aim to explore the readability, learner comprehensibility or familiarity with task instructions from L2 classroom learning materials and task performance since considering either of these in isolation reflects a limited perspective. Ultimately, despite the fact that task instructions created by teachers had been previously and extensively used by teachers in the classroom, it was shown that modifications improved both readability and comprehensibility for language learners. Teachers who create and utilize their own written task instructions should be mindful of keeping them simple, short, and presented in point-format

which reflects the temporal sequence of the actions required by the task.

References

Beglar, D. (2010). A Rasch-based validation of the Vocabulary Size Test. *Language Testing*, *27*, 101–118.

Bond, T. G., & Fox, C. M. (2007). *Applying the Rasch model: Fundamental measurement in the human sciences* (2nd ed.). Hillsdale, NJ: Erlbaum.

Brown, J. D. (1998). An EFL readability index. *JALT Journal*, *20*(2), 7–36.

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), 84–102.

DeWalt, D. A., Berkman, N. D., Sheridan, S., Lohr, K. N., & Pignone, M. P. (2004). Literacy and health outcomes. *Journal of General Internal Medicine*, *19*(12), 1228–1239.

Dornyei, Z., & Taguchi, T. (2010). *Questionnaires in second language research: Construction, administration, and processing*. New York: Routledge.

DuBay, W. H. (2004). *The principles of readability*. Costa Mesa, CA: Impact Information.

English Vocabulary Profile (2012). *About the English vocabulary profile.* Retrieved from http://vocabulary.englishprofile.org/staticfiles/about.html

Flesch, R. (1948). A new readability yardstick. *Journal of Applied Psychology, 32*(3), 221–233.

Gallagher, T. L., Fazio, X., & Gunning, T. G. (2012). Varying readability of science-based text in elementary readers: Challenges for teachers. *Reading Improvement*, *49*(3), 93–112.

Hirsh, D., & Nation, I. S. P. (1992). What vocabulary size is needed to read unsimplified texts for pleasure?. *Reading in a foreign language*, *8*, 689–689.

Hsueh-Chao, M. H., & Nation, I. S. P. (2000). Unknown vocabulary density and reading comprehension. *Reading in a Foreign Language, 13*(1), 403–430.

Koda, K. (2005). *Insights into second language reading: A cross-linguistic approach*. Cambridge: Cambridge University Press.

Lange, R., Greyson, B., & Houran, J. (2004). A Rasch scaling validation of a 'core' near-death experience. *British Journal of Psychology, 95,* 161–177.

Laufer, B., & Ravenhorst-Kalovski, G. C. (2010). Lexical threshold revisited: Lexical text coverage, learners' vocabulary size and reading comprehension. *Reading in a Foreign Language, 22*(1), 15–30.

Linacre, J. M. (2010). Winsteps® (Version 3.70.0) [Computer Software]. Beaverton, Oregon.

McAlpine, R. (2006). *From plain English to global English.* Retrieved from http://www.webpagecontent.com/arc_archive/139/5/

McLaughlin, G. H. (1969). SMOG grading: A new readability formula. *Journal of Reading, 12*(8), 639–646.

Miller, G. E., Rotou, O., & Twing, J. S. (2004). Evaluation of the .3 logits screening criterion in common item equating. *Journal of Applied Measurement, 5*(2), 172–177.

Nation, I. S. P. (2001). *Learning vocabulary in another language*. Cambridge: Cambridge University Press.

Nunan, D., & Keobke, K. (1995). Task difficulty from the learner's perspective: Perceptions and reality. *Hong Kong Papers in Linguistics and Language Teaching*, *18*(1), 1–12.

Richards, J. C., & Rodgers, T. S. (2001). *Approaches and methods in language teaching*. Cambridge: Cambridge University Press.

Shehadeh, A (2005) Task-based language learning and teaching: Theories and applications.

Shiotsu, T., & Weir, C. J. (2007). The relative significance of syntactic knowledge and vocabulary breadth in the prediction of reading comprehension test performance. *Language Testing, 24*(1), 99–128.

Skehan, P. (1996). A framework for the implementation of task-based instruction. Applied Linguistics 17, 38–62.

Wegner, M. V., & Girasek, D. C. (2003). How readable are child safety seat installation instructions? *Pediatrics*, *111*(3), 588–591.

—平成27年１月31日　受理—

Appendix 1

All instruction pairs included on the student survey

Pair	Pre-change Instruction	Post-change Instruction
1	Make a new pair. Together, unmix the questions to ask your classmates. Use the example conversation. Write the correct questions in the boxes below.	1. Find a new partner. 2. Unmix the questions. 3. Write the correct questions in the box. Use the conversation below.
2	Work with a partner and role-play the conversation below.	1. Find a partner. 2. Role-play the conversation below.
3	You will use these words for the next two tasks. Work with a partner. Match each vocabulary word with its meaning.	1. Find a partner. 2. Match the words on the left with the Japanese meanings on the right.
4	Now that you have completed the postcard, check your postcard with another group.	1. Work with another group. 2. Check your answers.
5	Now, make a group of three. Where is the best place for a honeymoon? Rank (from 1−10) the destinations in Japan below.	1. Make a group of three. 2. Rank (from 1−10) the honeymoon destinations. 3. Write a reason for each destination.
6	Listen to the song and if you hear the word or phrase, put a tick next to the word. Then answer the questions at the bottom.	1. Listen to the song. 2. If you hear the word or phrase, put a tick next to the word. 3. Answer the questions at the bottom.
7	Work with a partner, read the questions below. Then use the information in the article to answer the questions. The first one has been done for you. Write your answers in full sentences.	1. Find a partner. 2. Read the article. 3. Answer the questions.
8	Ask two classmates to answer the questions below. Write their answers in the spaces provided.	1. Find a partner. 2. Ask her the questions below. 3. Write her answers. 4. Ask another classmate!
9	Match the vocabulary on the left with the appropriate picture or meaning. The first one has been done for you.	Match the words on the left with the pictures or meanings on the right.