Scientific Structural Changes within Texts of Adapted Reading Materials

Anealka Aziz (PhD)
Academy of Language Studies, UiTM
Universiti Teknologi MARA, 40150 Shah Alam, Selangor, Malaysia
E-mail: anealka@salam.uitm.edu.my

Chan Yuen Fook (PhD) & Zubaidah Alsree (PhD)
Faculty of Education, Universiti Teknologi MARA, Malaysia

Abstract
ESL instructors generally use ready-made reading materials in commercially published coursebooks. However, it would be more effective for learners if ESL instructors are involved in developing reading materials for them. This is because they are closer, more sensitive and responsive to the needs of their learners and they are also aware of the reading ability of the learners. ESL instructors can adapt reading materials taken from non-instructional sources like magazines and newspapers and adapt them to suit the needs and reading ability of their learners. Adaptation enables ESL instructors to adjust the difficulty level of the materials. Adjustment can be made by making changes at sentence and word levels and this study intends to demonstrate the structural changes that take place when the materials are being adapted. This study uses the a tool named IDL Suite to measure the difficulty level of the materials before and after adaptation.

Keywords: Reading materials, Difficulty level, Adaptation, Sentence difficulty, Word difficulty

1. Introduction
Generally, ESL instructors use reading materials from coursebooks specifically written for instructional purposes when teaching reading skills to their ESL learners. They usually use coursebooks prescribed by their teaching institution. For those who want to add varieties in the selection of topics, they would use supplementary materials from other coursebooks as suggested by either their colleagues, publishers or based on their own preference. They usually use these materials and the accompanying tasks as they are, without making any changes to the reading materials nor to the tasks assigned.

Besides that, there are also ESL instructors who use reading materials taken from non-instructional materials such as newspaper articles, journal articles, brochures, etc. to teach the language to ESL learners. Materials taken from these non-instructional and/or authentic sources such as newspaper articles, journal articles, brochures, etc. are useful in developing general education values in learners. These materials provide information on what is happening around the world (Sanderson, 1999), keep the learners up to date with the language change that takes place (Sanderson, 1999; Martinez, 2002) and enable them to see a wide variety of vocabulary, sentence structures, text types and language styles, which are not found in conventional teaching materials (Martinez, 2002). However, these materials may not be suitable for lower level learners because the language could be too difficult for them to cope with (Sanderson, 1999; Martinez, 2002; Tan, 2006). Nation (1990) also claims that unsimplified materials usually contain vocabulary, which is hard for the learners to understand. Therefore, as suggested by Crandall (1995), these materials need to be adapted if they are to be used with for learners at lower level. In fact, studies have also shown that learners are able to comprehend simplified materials better than unmodified materials (Yano, Long & Ross, 1994; Oh, 2001). For that reason, ESL instructors need to adapt materials from these sources if they want to use these materials with lower level learners. ESL instructors can revert to the original versions of the materials once the learners reach the expected level of competency.

The most preferred adaptation strategy is simplifying strategy and it can be done in various ways such as replacing low frequency / specialized / idiomatic vocabulary with high frequency vocabulary, reducing the length and complexity of sentences, and expanding the materials by replacing pronoun references with noun phrase referents (Thornbury, 2005). ESL instructors usually rely on their judgment to determine the amount of simplification that they need to do so that the difficulty level of the materials matches the language ability of their learners.

The judgment method is the most commonly used method to estimate the difficulty level of reading materials adopted by ESL instructors. However, while ESL instructors are consistent in ranking the difficulty of reading materials, in corresponding to objective measures of difficulty, their ranking could be inaccurate and inadequate (Chall, 1958; Gunning, 2003). Besides that, this method is unreliable if it is done by untrained instructors (Klare, 1984). Klare (1984) also claimed that judgment of materials difficulty by untrained instructors could not be relied upon as their ratings often varied by several grade levels.

There is also evidence showing the inconsistency of rating materials difficulties that result from using personal judgment. In a survey conducted by Aziz (2010a), the participants were asked to rate the difficulty level of five reading passages. Results of the survey showed that there was no agreement in the rating of the difficulty level of these five reading passages. This inconsistency could be due to the length of experience the instructors had in
teaching the target ESL learner. As reported, 63.9% of the respondents had more than 15 years of experience teaching ESL learners, 22.2% of the respondents had between 10 to 15 years of experience and the remaining 13.9% had less than 10 years of experience.

The inconsistency that results from using personal judgment in determining the difficulty level of reading materials can also be seen in the work of Hancioglu and Eldridge (2007). In the study, the participants were asked to rank five reading materials from the easiest to the most difficult. The results showed that the ranking was inconsistent except for one reading material that was ranked as the easiest almost unanimously.

**IDL Suite**

ESL instructors should consider other options that can complement the subjective nature of instructors’ judgment. ESL instructors can consider using a tool named **IDL Suite** to estimate the difficulty level of reading materials for ESL learners at the intermediate difficulty level scale (IDL scale henceforth). The IDL scale starts at IDL (the least difficult level in the range) and ends at IDL (the most difficult level in the range). Table 1 shows the general descriptors of the IDL scale.

**IDL Suite** is constructed based on the intermediate difficulty level formula (IDL Formula henceforth) which is a readability formula developed to measure ESL reading materials at the IDL scale (Aziz, 2010b). The IDL Formula uses the Flesch Reading Ease (FRE) score and the coverage of the first 2000 high-frequency words (HFW) in the reading materials as predictors of materials difficulty. (Refer to Figures 1 & 2 for the calculation of FRE and HFW respectively)

**IDL Suite** is a set of tools that can be used to estimate the difficulty level of reading materials at the IDL scale. IDL Suite comprises of three support tools namely **IDL Guide**, **IDL Generator**, and **IDL Word Calculator** (Aziz, 2010b).

a. **IDL Guide** is a non-automated tool to estimate the difficulty level of reading materials at the IDL scale. It is a table that uses FRE and HFW scores as the predictors to estimate the difficulty level of the materials. ESL instructors need to match the FRE score (column) with the HFW score (row) to get the estimated difficulty level of the reading materials. IDL Guide is displayed in Figure 3.

b. **IDL Generator** is an automated tool to estimate the difficulty level of reading materials at the IDL scale. It is a computer application that runs using Microsoft Excel programme. ESL instructors are required to fill in the necessary information about the reading materials namely the FRE and HFW scores in the spaces provided. The computer automatically calculates the scores and displays the estimated difficulty level of the materials almost instantaneously. A pop-up message will appear if the scores entered are outside the specified range. The interface of IDL Generator is displayed in Figure 4.

c. **IDL Word Calculator** is an automated tool to calculate the coverage of HFW in reading materials. It also runs using Microsoft Excel programme and it is used to determine the HFW score of the material, which is one of the two scores required to estimate the difficulty level of reading materials using **IDL Generator** and **IDL Guide**. The interface of IDL Word Calculator is displayed in Figure 5.

**2. Objectives and Research Questions of the Study**

As mentioned earlier, the study intends to demonstrate the structural changes that take place in reading materials when the materials are being adapted. To do that, the study needs to perform a correlation analysis between the ranking of materials difficulty level (DL) and the text characteristics of the materials. In this study, text characteristics of the materials refer to the FRE scores, average sentence length (ASL), the use of simple/compound sentences (S/Cd), the use of complex/compound-complex sentences (Cx/CdCx), average word length (AWL) and the HFW of the materials.

The following research questions are constructed:

RQ1. What are the relationships between DL and the text characteristics of reading materials?

RQ2. What are the changes that take place in the text characteristics of reading materials at lower and higher IDL Scale?

**3. Methodology**

Reading materials taken from several reading coursebooks for intermediate and high intermediate ESL learners are selected and scanned to convert them into soft copies. Readability Statistics function in Microsoft Word is used to generate the FRE scores of the materials. The study also uses Writer’s Workbench 8.18 to extract information on the ASL, S/Cd and Cx/CdCx from the materials and WordSmith Tools 4.0 is used to extract information on the AWL and HFW of the materials.

The study also ensures that all the text characteristics meet the assumptions required prior to conducting inferential analysis as suggested by (Tabachnick & Fidell, 2001; Coakes, Steed & Dzidic, 2006). Firstly, the minimum requirement of the ratio of cases to independent variables should be met. Secondly, there should not be any outliers. Thirdly, there should not be any multicollinearity and singularity of variables. Finally, there should be normality, linearity, homoscedasticity and independence of residuals.
After all the relevant information is acquired and all the assumptions required prior to conducting inferential analysis are fulfilled, a correlation analysis of the DL and the text characteristics of materials is performed.

4. Findings

4.1 Relationships between the DL and text characteristics of reading materials

Pearson Product Moment Correlation Coefficient was used to determine the relationships between the DL of reading materials and the text characteristics of the materials, while Guildford’s interpretation of coefficient values (Guildford, 1973) was used to interpret the results of the analysis. The results display in Table 2 show that the DL of the materials has a significant correlation with FRE, ASL, AWL and HFW but no significant correlation with S/Cd or Cx/CdCx. The DL of the materials has a high negative correlation with FRE \((r = -.816, p<.05)\), a low positive correlation with ASL \((r = .362, p<.05)\), a high positive correlation with AWL \((r = .778, p<.05)\) and a high negative correlation with HFW \((r = .774, p<.05)\). An increase in the DL of the materials lowers the FRE scores and the percentage of HFW but increases the ASL and AWL scores. In other words, an increase in the DL of the materials increases materials difficulty at text, sentence and word levels. Values of FRE and HFW decrease while values of ASL and AWL increase. As a result, the materials become more difficult as the difficulty level of the materials increases.

On the other hand, a decrease in the DL of reading materials increases the score of FRE and the percentage of HFW and lowers the scores of ASL and AWL. In other words, a decrease in the DL of the materials lowers materials difficulty at text, sentence and word levels. Values of FRE and HFW increase while values of ASL and AWL decrease. As a result, the materials become easier as the difficulty level of the materials decreases.

In terms of the relationships among the text characteristics, FRE shows a significant correlation with ASL, AWL and HFW. FRE has a low negative correlation with ASL \((r = -.375, p<.01)\), a high negative correlation with AWL \((r = -.895, p<.01)\) and a moderate positive correlation with HFW \((r = .423, p<.01)\). ASL is the only variable that is significantly correlated with all text characteristics of the materials. Besides having a low negative correlation with FRE \((r = -.375, p<.01)\), ASL has a moderate negative correlation with S/Cd \((r = -.582, p<.01)\) and HFW \((r = -.435, p<.01)\), a moderate positive correlation with Cx/CdCx \((r = .582, p<.01)\) and a low positive correlation with AWL \((r = .286, p<.05)\). ASL of reading materials gets longer, the materials become harder to read. The materials also contain fewer number of simple and compound sentences and uses more of complex and compound-complex sentences. Words used in the materials get longer and less familiar too.

The types of sentences used in the materials are significantly correlated with one another. S/Cd has a perfect negative correlation with Cx/CdCx \((r = 1.00, p<.01)\). S/Cd also has a moderate negative correlation with ASL \((r = -.582, p<.01)\), while Cx/CdCx has a moderate positive correlation with ASL \((r = .582, p<.01)\). S/Cd and Cx/CdCx, however, do not have any significant correlation with FRE, AWL and HFW. As the number of simple sentence structures increases, the number of complex sentence structures decreases, resulting in a change in ASL, which later alters the difficulty level of the materials. The type of sentences used in the materials, however, does not directly affect the choice of words to be used.

AWL has a significant correlation with FRE, ASL and HFW. AWL has a high significant negative correlation with FRE \((r = -.895, p<.01)\), a moderate negative correlation with HFW \((r = -.475, p<.01)\) and a low positive correlation with ASL \((r = .286, p<.05)\). However, there is no significant correlation between AWL and S/Cd or Cx/CdCx. An increase in AWL reduces FRE, making the materials more difficult to read and understand. It also reduces the coverage of HFW in the materials and increases ASL of the materials.

HFW is significantly correlated with FRE, ASL and AWL. HFW has a moderate positive correlation with FRE \((r = .423, p<.01)\) and a moderate negative correlation with ASL \((r = -.435, p<.01)\) and AWL \((r = -.475, p<.01)\). A higher coverage of HFW increases the FRE score, making the materials easier to read and understand. It also shortens the length of sentences and length of words used in the materials.

Figure 6 summarizes the relationships between the DL and the text characteristics of reading materials. These relationships allow ESL instructors to adjust the difficulty level of reading materials by making changes at text, sentence and word levels.

4.2 Changes that take place in the text characteristics of reading materials at lower and higher IDL Scale

To illustrate the changes that take place in the text characteristics of reading materials at lower and higher intermediate difficulty scale, an extract of a reading material taken from one of the reading coursebooks is selected randomly. This reading extract is rated at IDL\(^2\). The reading extract is revised twice to illustrate the changes that take place when there is a change in any of the text characteristics. The revised versions have been doctored by the researcher to illustrate the changes that take place when the difficulty level of the extract changes. Figure 7 shows the reading extract and its text characteristics.

The first revision of the reading extract is to illustrate the changes that take place when the extract is made less difficult (IDL\(^1\)). In the first revision, there is an increase in the total number of sentences (from 5 to 7 sentences) but
a decrease in the total number of words used (from 135 to 119 words). The changes could suggest that long sentences have been broken down into shorter sentences and there could be fewer words to process as the sentence structures have been made simpler. The revised version of this extract is more readable than the original version. It can be seen that the revised version has a higher FRE score of (67.3) than the original version, which has an FRE score of 59.2.

Besides that, a lower ASL (17 w.p.s.) shows that generally the new version uses shorter sentences than the original version (22.5 w.p.s.). It can be seen that the use of S/Cd sentences is higher (28%) and the use of Cx/CdCx sentences is lower (72%) in the revised version, as compared to the original version (17% and 83% respectively). These results show that longer sentences have been broken into shorter sentences to reduce the burden of processing that readers have to do to understand the materials.

Apart from that, the AWL of the revised version also shows a reduction of the number of characters per word. It shows that the words used in this revised version are shorter (4.47 c.p.w.) than the words used in the original version (4.56 c.p.w.). In addition, the HFW of the revised version also supports that words in this version are shorter as higher frequency words tend to be shorter than those of lower frequency. It can be seen from a higher percentage of HFW of the revised version (74.0%) as compared to the original version (72.3%). Figure 8 shows the first revised version of the reading extract and its text characteristics.

The second revision of the reading extract is to illustrate the changes that take place when the extract is made more difficult (IDL³). In the second revision, there is a decrease in the total number of sentences (from 5 to 3 sentences) as well as a decrease in the total number of words used (from 135 to 98 words). The changes could suggest that short sentences have been combined to form longer sentences and because of that, some words have to be removed. Although there are lesser words to be processed, the difficulty level of this version increases as the same amount of information is delivered via lesser number of sentences and words. The revised version is less readable than the original version. It can be seen that the FRE score of the revised version is lower (39.8) than the FRE score of the original version (59.2).

Besides that, a higher ASL shows that generally the second revised version uses longer sentences (32.6 w.p.s.) than the original version (22.5 w.p.s.). It can be seen that this revised version is more difficult as it uses 100% of Cx/CdCx as compared to the original version (83% of Cx/CdCx and 17% of S/Cd). It shows that shorter sentences have been combined into longer sentences and it requires more effort from the readers to process these sentences in order to understand the extract.

Apart from that, AWL of the second revised version also shows an increase in the number of characters per word. It shows that the words used in this version is longer (4.8 c.p.w.) than the words used in the original version (4.56 c.p.w.). In addition, HFW of this revised version also supports that words in this version are longer as lower frequency words tend to be longer than those of higher frequency words. It can be seen from a lower percentage of HFW of the revised version (70.3%) as compared to the original version (72.3%). Figure 9 shows the second revised version of the reading extract and its text characteristics.

These examples illustrate the changes that may take place at text, sentence and word levels when the difficulty level of reading materials is altered. In order to increase the difficulty level of reading materials, ESL instructors may increase the ASL, Cx/CdCx, and AWL and lower the S/Cd and HFW of the materials and vice versa. The FRE scores change according to the changes that take place at sentence and word level. Table 3 summarizes the text statistics of the three versions as measured at the IDL scale.

**Conclusion**

Knowing how the text characteristics of reading materials interact to the changes in the DL of the materials and also among each other help ESL instructors make necessary changes at sentence and word level in order to make the materials less or more difficult. With the assistance of the *IDL Suite*, the process of adapting reading materials for learners can now be performed in a better way. ESL instructors can now combine their expert judgment together with an objective method in adapting reading materials for the learners.

**References**

Aziz, A. (2010a). Discrepancy among language instructors concerning the difficulty level of reading passages for learners at a diploma level. In S. Menon, & J. Lourdunathan, *Readings on ELT Materials IV*. Malaysia: Pearson Malaysia Sdn Bhd.

Aziz, A. (2010b). *Computational text analysis of intermediate and high-intermediate reading passages for ESL learners*. Unpublished Doctoral Thesis. Universiti Teknologi MARA, Malaysia.

Chall, J. S. (1958). *Readability: An appraisal of research and application*. Columbus, Ohio: Ohio State University.

Crandall, J. (1995). *The why, what, and how of ESL reading instruction: Some guidelines for writers of ESL textbooks*. In P. Byrd, *Material writer's guide*. US: Heinle and Heinle Publishers.

Coakes, S. J., Steed, L. & Dzidic, P. (2006). *SPSS 13.0 for Windows. Analysis without anguish*. China: Wiley.
Table 1. General Descriptors of the Intermediate Difficulty Level Scale

| Levels | General Descriptors |
|--------|---------------------|
| IDL¹   | A reading material at this level is less difficult than IDL². It has an overall difficulty level within the range of 59.0 to 65.0. It has 74% to 80% of familiar words that are frequently encountered. It uses different types of sentences and the sentence length of the material varies from 13 words per sentence to 25 words per sentence. |
| IDL²   | A reading material at this level is more difficult than IDL¹ but less difficult than IDL³. It has an overall difficulty level within the range of 44.0 to 65.0. It has 59% to 80% of familiar words that are frequently encountered. It uses different types of sentences and the sentence length of the material varies from 13 words per sentence to 25 words per sentence. |
| IDL³   | A reading material at this level is more difficult than IDL² but less difficult than IDL⁴. It has an overall difficulty level within the range of 30.0 to 65.0. It has 50% to 80% of familiar words that are frequently encountered. It uses different types of sentences and the sentence length of the material varies from 13 words per sentence to 25 words per sentence. |
| IDL⁴   | A reading material at this level is more difficult than IDL³ but less difficult than IDL⁵. It has an overall difficulty level within the range of 30.0 to 58.0. It has 50% to 78% of familiar words that are frequently encountered. It uses different types of sentences and the sentence length of the material varies from 13 words per sentence to 25 words per sentence. |
| IDL⁵   | A reading material at this level is more difficult than IDL⁴. It has an overall difficulty level within the range of 30.0 to 44.0. It has 50% to 64% of familiar words that are frequently encountered. It uses different types of sentences and the sentence length of the material varies from 13 words per sentence to 25 words per sentence. |

Table 2. Relationships between the DL and the Text Characteristics of Reading Materials

|       | DL     | FRE    | ASL    | S/Cd   | Cx/CdCx | AWL    | HFW    |
|-------|--------|--------|--------|--------|---------|--------|--------|
| DL    | 1.000  | -.816* | .362*  | -.083  | .083    | .778*  | -.774* |
| FRE   | -.816* | 1.000  | -.375**| .173   | -.173   | -.895**| .423** |
| ASL   | .362*  | -.375**| 1.000  | -.582**| .582**  | .286*  | -.435**|
| S/Cd  | -.083  | .173   | -.582**| 1.000  | -1.000**| -1.09  | .056   |
| Cx/CdCx| .083   | .173   | -.582**| 1.000  | -1.000**| -1.09  | .056   |
| AWL   | .778*  | -.895**| .286*  | -.109  | .109    | 1.000  | -1.075**|
| HFW   | -.774* | .423** | -.435**| .056   | -.056   | 1.000  |        |

** Correlation is significant at the .01 level; * Correlation is significant at the .05 level.

Note: FRE= Flesch Reading Ease; ASL=Average Sentence Length; S/Cd=Simple/Compound; Cx/CdCx = Complex/Compound-Complex; AWL=Average Word Length; HFW= High Frequency Words
Table 3. Text Statistics of the Original and Revised Versions of a Reading Extract

|                   | First Revision | Original | Second Revision |
|-------------------|----------------|----------|-----------------|
| **Tokens**        | 119 words      | 135 words| 98 words        |
| **Sentences**     | 7 sentences    | 5 sentences| 3 sentences   |
| **FRE**           | 67.3           | 59.2     | 39.8            |
| **ASL**           | 17.0 w.p.s.    | 22.5 w.p.s. | 32.6 w.p.s. |
| **S/Cd**          | 28%            | 17%      | 0%              |
| **Cx/CdCx**       | 72%            | 83%      | 100%            |
| **AWL**           | 4.47 c.p.w     | 4.56 c.p.w. | 4.8            |
| **HFW**           | 75.3%          | 72.3%    | 70.3%           |

**Figure 1. The Flesch Reading Ease Formula (Flesch, 1948)**

\[
Y = 206.835 - 1.015(X_1) - 84.6(X_2)
\]

- \( Y \) = Reading Ease Score
- \( X_1 \) = Average Sentence Length
- \( X_2 \) = Average number of Syllables per Word

**Figure 2. Formula to Calculate HFW of Reading Materials**

\[
HFW = \frac{\text{Total number of words that is within the high-frequency list}}{\text{Total number of different words in the reading material}} \times 100
\]

**Figure 3. IDL Guide**
Figure 4. IDL Generator

Figure 5. IDL Word Calculator

Figure 6. Relationships between DL and Text Characteristics of Reading Materials
Animals such as the narwhal may provide evidence, or give us a clue, that unicorns once existed on earth. The male narwhal is a huge sea creature that can be found in the Arctic seas. This animal has a single horn, or tusk, that projects from its forehead - similar to that of the unicorn. In fact, in the seventeenth century, people often confused the narwhal's tusk with the horn of the unicorn. The similarity of the male narwhal to the unicorn has led some people to believe that perhaps the unicorn is the narwhal's ancestor, or vice versa. Although narwhals have also been hunted for centuries, they may have escaped extinction for different reasons: they are only found in certain parts of the world, and are hunted by only a relatively small number of people.

Source: Para5, Animal in Mythology, Active: Skills for Reading, Thompson Learning

Figure 7. Reading Extract and its Text Characteristics

Animals such as the narwhal may suggest that unicorns once lived on earth. The male narwhal is a huge sea creature and it can be found in the Arctic seas. It has a single horn on its forehead, which makes it looks like the unicorn. In fact, in the seventeenth century, people often confused the narwhal's horn with the horn of the unicorn. This similarity has made some people think that the unicorn could be the narwhal's ancestor, or vice versa. Although, narwhals have also been hunted for centuries, they can still be found nowadays. This is because they are only found in certain parts of the world, and are hunted by only a relatively small number of people.

Figure 8. The First Revision of the Reading Extract and its Text Characteristics

Narwhals, huge sea creatures that can be located in the Arctic seas, may offer confirmation on the existence of unicorns on the surface of the earth. A sole protruding horn in the middle forehead of the male narwhals, similar to the horn of the unicorns, have led people in the seventeenth century to believe that Narwhals could be ancestors of unicorns, or descendants of unicorns. Narwhals, which can be found only in limited areas of the world and are preyed upon by only a relatively small number of people, have shrunk from extinction despite being sought after for centuries.

Figure 9. The Second Revision of the Reading Extract and its Text Characteristics