A case for linguistic simplification

Jim Smiley†

Abstract
This study investigated the effect on comprehension of a complex reading text after a limited and timed exposure to a simpler version. This design was an attempt to mirror teacher simplification actions during class contact time in order to study the change in medium-term language development. A baseline text was simplified and made more complex using readability scales. Subjects (N=81) were divided into three groups. Group S received the simple version of the text, group D received the difficult version and group B received both versions. A posttest comprehension test was administered. Statistically significant differences were observed between the S group and the D group, providing evidence that readability scales distinguish between levels in English language education, and significant differences occurred between the B and the D group which supported the study hypothesis that access to a simplified version would aid comprehension of a complex text. Group B’s exposure to the simple version was limited, and before they answered the comprehension questions, sufficient time was built in to the research design to minimise the effect of working memory. This study tentatively forwards the argument that linguistic simplification offers a direct route to language development beyond its currently accepted role of only providing comprehensible input over an extended study time.

†東北文化学園大学専任講師 Lecturer, Tohoku Bunka Gakuen University
email: jimsmiley@pm.tbgu.ac.jp
この研究では制限時間内に平易な文体で書かれたテキストを読んだ後
に複雑な文体のテキストを読む際、どのような影響があるかについて調
査した。基準となるテキストは平易な文体で書かれ、読み易さの目安に
基づいてさらに複雑な文体にされた。被験者（81名）は3つのグルー
プに分けられた。グループ S には平易な文体のテキストを、グループ D
には複雑な文体のテキストを、グループ B には両方のテキストが与え
られた。この手法は授業中に教員が説明を簡略化する様を正確に模し、
中期的な言語発達にどのような変化をもたらすかを調査した。調査後
には読解テストも合わせて実施した。グループ S とグループ B の間に
は統計的に有意な違いが見受けられ、今回用いた難易度基準は ELT に
おけるレベル差を明らかにするうえで有効であった。また、グループ B
とグループ D の間にも同様に有意な違いが見られ、平易な文体のテキ
スト読解後には複雑な文体のテキスト読解が容易になるという本研究
の仮説を裏付ける。グループ B は平易な文体のテキストに接する機会
を制限され、実際に読解問題に取り組む前に、十分に時間をおいて作業
に関わる記憶の影響が出ないように出来る限りの配慮をした。現在広く
受け入れられている考えによれば、言語の発達は長期間理解可能なイン
プットを与えるだけでよいとされている。しかし、本研究はその議論を
一歩進めて、言語の簡素化はむしろ言語発達に直結していると指摘する
ものである。

Concepts and arguments
When students fail to comprehend texts, a very natural response by teachers is to simplify the
material. Simplification strategies include: code-switching, contextualisation, grammar structure
explanation, linguistic modification and paraphrasing (Goodman & Freeman 1993; Sibayan et al.
1993). A less intuitive teacher reaction would be to instruct students to ignore the difficult text
entirely. Asking students to do this seems unethical; it goes against the grain of wishing to inspire:
a love of learning the unknown; a desire to search for meaning; a persevering attitude in times of
difficulty. Yet according to current beliefs about the role simplification plays in language
development, asking students to avoid complex texts well be the most beneficial thing to suggest
to the student. Although teachers may believe that by simplifying, learners develop their language
abilities, how those abilities are developed through simplification is an open question. Indeed, not
only might the act of linguistic simplification be only useful in “coping with the conceptual and
experiential chaos that surrounds us [in the classroom]”, it might be “a potentially dangerous
process, too” (Brumfit 1993, p.3). The motives underlying the act of simplification include fulfilling immediate procedural requirements, such as helping students get through the task in hand and avoiding demotivation when students find material too difficult, and realising pedagogical principles. Accordingly, the idea that a simplified text can act as a vehicle of the increased comprehension of complex material has received little support in theory. Ellis sums up this view, “simplified input may help comprehension, but is a waste of time where acquisition is concerned” (1993, p.57). Simplification serves the here and now as a teacher action and provides comprehensible input over the long term, and there is sufficient agreement that simpler passages are understood better (Long & Ross, 1993 p.10; Tewissi 1998).

Indeed, the strong version of the argument against simplification sees writers on simplification noting the reverse effect so-called 'simplification' has on difficult texts. Susser and Robb conclude that “the process of simplification . . . leaves writing that is more difficult to understand than the original” (1990, cited in Mehrpour & Riazi 2004, p.4), a view supported by Bernhardt who comments that “syntactic simplicity may decrease text cohesion and thereby hinder comprehension” (1984, cited in Mehrpour & Riazi 2004, p.4). Cervantes and Gainer (1992) claim that “input premodified by decreased complexity and increased redundancy did not assist comprehension” (1992, p.767). Honeyfield claims that “Simplification of syntax may reduce cohesion and readability” (1977, p.435), a view echoed by Strother and Ulijn who take up Honeyfield’s position by commenting that “simplifying syntax does not necessarily lead to more readable texts” (1987 cited in Mehrpour and Riazi 2004, p.5). Weaker versions argue for no definite position. Even though Rivera and Stansfield found an original test text that was answered better than a corresponding simplified version, they caution that their results “lacked statistical power due to the small sample size” (2001, p.3). Young concurred and commented that “recall scores for the simplified texts were not superior to the authentic ones” (1999 p.350). Yano et al. believe that simplified input may not be beneficial for language learning because even though “removal of possibly unknown linguistic items from a text may facilitate comprehension”, it is “worse . . . [because it] will simultaneously deny learners access to items they need to learn” (1994, p.191).

A challenge to the limited view role of simplification comes from studies in glossing. Glosses that explain the in-context value (‘value glosses’) or the definition (‘definition glosses’ Widdowson 1978, p.86, elaborated by Roby 1999, p.94) help subjects outperform unglossed text on various posttests (Ko 2005; Yoshii, 2006; Lomika, 1998). Furthermore, Yoshii found that posttests two weeks after glossing input, subjects had better recall of the L1 glosses, suggesting that simplification in L1 may result in better retention (2006, p.93) and that simplification (through glosses in Yoshii’s case) may lead to more immediate language development. This is a
highly contentious issue, and other studies point to better memory retention of L2-contextualised instruction. Ellis 1993 reports on a study by Brauer in which retention of L2-contextualised lexis was higher than L1 translation (1993, p.59). Leow summarises the situation: “The issue of the role and effects of simplification on learners’ comprehension and intake remains contentious” (1997, p.291).

*Linguistic simplification*

Two kinds of linguistic simplification may be observed: additive and subtractive. Subtractive simplification involves (taking a baseline text) shortening the sentences in an effort to reduce clause numbers, which is supposed to render texts easier; identifying and rewriting idiomatic expressions into more standard ones; replacing low frequency words with high frequency ones; shortening entire texts to reduce the overall complexity; identifying and replacing oblique contextual support with more concrete language. Apart from this last technique, which may be defined as pragmatically oriented, the others are linguistically based. An example from the present study is:

Higher level: She was the daughter of an aristocratic family.
Simplified: Her family was upper class.

Here, we can see that the simplified version is shorter, the term ‘aristocratic’ has been replaced with ‘upper class’ and the meaning in the clause ‘she was the daughter’ is subsumed in the single word ‘her’. To take Leech’s categorisation of seven types of meaning as a useful base, we can see that there is no loss of conceptual meaning, although connotative, stylistic and collocative meanings have all disappeared (1974, p.11. Writers have criticised subtractive simplification intensely. For example, Honeyfield claimed that reading simplified “material may lead students to develop reading strategies that are inappropriate for unsimplified English.” (1977, p.431). Yano makes a similar claim: “elimination [or subtraction] prevents exposure to items that learners eventually should know” (1994, p.71). In other words, not only does simplification deprive students of valuable and complex input, it may actually lead them to conceptualise the L2 ineffectively.

Additive techniques do not take away from the base text: they add to it. When we elaborate during natural speech, we rephrase misunderstandings, we explain lexical items that are not understood, we paraphrase meanings and we supply opaque contextual information. An example from Long and Ross (1993, p.31) goes like this:
Native Speaker Baseline Version: Because he had to work at night to support his family, Paco often fell asleep in class.

Elaborated Version: Paco had to work at night to earn money to support his family, so he often fell asleep in class next day during his teacher’s lesson.

They also give a simplified version by way of comparison.

Simplified Version: Paco had to make money for his family. Paco worked at night. He often went to sleep in class.

They claim that their elaborated version “sounds more like a natural sample of spoken or written English” (p.32). I will leave it up to you to decide if their example chosen was the best one to demonstrate their case. I cannot decide if a text that, at the same time, sounds like both spoken and written language is actually a good thing or not. Personally, I think that there are very few crucial differences between elaboration and good simplification. Other writers have said the same thing. Claridge challenged Honeyfield’s claim that “Simplification of syntax may reduce cohesion and readability” (1977, p.435) in relation to graded readers and found that well-written simplifications preserve “the essential features of a 'normal' English text” and are no less meaningful than their originals” (2005, p.157). The question is still very open. In this present study, no distinction is drawn between subtractive and additive simplification. This decision follows Young’s precedent in her 1999 Modern Language Journal article on the same topic where she places both kinds under the same banner of text simplification (1999, p.350).

Readability
Readability refers to how easy a text is to read. The method of getting readability data comes from mathematics. Readability is, basically, a mathematical way of analysing language. It’s based on the some of the same concepts we looked at in subtractive linguistic simplification: shorter words, shorter sentence lengths and fewer clauses. Notice that these three terms were all placed around the cognitive space of ‘short’. This can be described easily in terms of mathematics. A readability scale will take a passage and count it. It will count the number of words overall, the number of paragraphs, the number of sentences, the number of words in a sentence, the number of syllables in a word. It will return a ‘simple’ verdict to a sentence that is made up of short elements. Here is an example:
The cat sat on the mat.

This sentence returns a value of Flesch Reading Ease of 100.0 and a Flesch-Kincaid Grade Level of 0. The first figure is out of 100 and 100 means easy. Difficult texts would score lowly. The Flesch-Kincaid Grade Level places the text into the appropriate US grade school level. In other words, a text with the score of, say 6 would mean that a US 6th grader should be able to read it.

The conventional wisdom about these scales is that they are a useful tool, but that they shouldn’t be taken too seriously. Numerous studies, both experimental research based and theoretical works, have been written on the appropriateness of using scales in L1 education. A body of similar research has been done in L2. This topic is huge. Here, I won’t begin to survey the various arguments. I’ll give just one example to show how the scales may be misinterpreted.

The oaf sat on a lea.

This sentence also returns roughly the same values as the cat on the mat example did, but is most likely incomprehensible to many students without a dictionary. It must be said that it took me quite a while to find ways of devising such a sentence. More spontaneous, or natural, thinking strategies would more likely produce more ‘normal’ scores. In the present study, readability scales were used to force a difference between the difficult and simple versions.

A case for linguistic simplification
Previous studies have compared comprehension and recall scores on simple versions with baseline versions, with different types of simplification and glossing, and at different levels. But none has compared the comprehension scores of a student group that has received help and is answering questions on the difficult text only, with the control groups of those who were exposed to only the simple or the difficult texts. This means, effectively, that teachers are still unaware of the effects of their help during lessons. To find that out, we need to give that help to some and not to others. Then we need to give some time for the help to vanish from the working memory. Then we need to test the comprehension of the groups. There was a need, therefore, to investigate the effect of simplification (linguistic, pedagogic, or elaboration) on comprehension after a short delay. The length of this delay was crucial. If it were too short, pure working memory of the simplified text may have become a too large factor in the answering of the comprehension questions. If it were too long, any effect the exposure to the simplified text may have had might
be lost.

Participants
The participants in this study were 81 non-English language majors aged between 18 and 19 (male 38, female 43) taken from three separate class groups majoring in Occupational and Physical Therapy at Tohoku Bunka Gakuen University in Sendai, Japan. Subjects were randomly assigned to three study groups based on their student identification number, which contained no sex or level data. Group S received the simple version of the text, group D received the difficult version and group B received both versions. A mock TOEIC Bridge Test was administered to all three classes. Students who missed this test were eliminated from the study data. As the study group allocation was random, no attempt was made to balance the study groups’ sexes, or levels based on Bridge Test data in advance, or to even the groups’ numbers post hoc. As a result, the groups’ numbers were slightly uneven.

Instruments
A passage from the Longman Beginner Text, What a Life, which described the life of Diana, the Princess of Wales, was selected on account of its assumed close psychological distance to the subjects and because the base level of the text allowed it to be modified in simplification and in complexity. The Flesch-Kincaid Grade Level (FKGL) scale was used to grade the alteration process. The base text at FKGL 9.0 was adapted to create two texts: a simple version at FKGL 7.4, and a difficult version at FKGL 12.0 (see table 1).

<table>
<thead>
<tr>
<th></th>
<th>Flesch Reading Ease</th>
<th>Flesch-Kincaid Grade Level</th>
</tr>
</thead>
<tbody>
<tr>
<td>Diana – original</td>
<td>45.4</td>
<td>9.0</td>
</tr>
<tr>
<td>Diana – difficult</td>
<td>26.3</td>
<td>12.0</td>
</tr>
<tr>
<td>Diana – simple</td>
<td>54.3</td>
<td>7.4</td>
</tr>
<tr>
<td>Diana – questions</td>
<td>71.1</td>
<td>5.3</td>
</tr>
</tbody>
</table>

Table 1. Text Complexity Figures

Twenty posttest comprehension questions were written that could be answered without recourse to the syntax or lexis of either passage. These questions were designed to be at a lower FKGL than the reading texts to avoid problems of question incomprehensibility. The concurrently-generated Flesch Reading Ease figures reinforced the perceived simplicity/complexity of the resultant test texts and questions. The original text may be ignored from this
study as its purpose was to provide a useful starting point for adaptation.

The issue of memory was controlled for by having a 2-minute interval between the final exposure to the simplified text and answering the comprehension questions. A key point in the methodology of the study, therefore, was the issue of timing. Test timing may be seen in Table 2.

<table>
<thead>
<tr>
<th>Group</th>
<th>Minute 0</th>
<th>6</th>
<th>12</th>
<th>15</th>
<th>17</th>
<th>42</th>
</tr>
</thead>
<tbody>
<tr>
<td>S</td>
<td>Read simple version</td>
<td></td>
<td></td>
<td></td>
<td>Comprehension</td>
<td>End</td>
</tr>
<tr>
<td></td>
<td>Questions Start</td>
<td></td>
<td></td>
<td></td>
<td>Test</td>
<td></td>
</tr>
<tr>
<td>D</td>
<td>Read difficult version</td>
<td></td>
<td></td>
<td></td>
<td>Comprehension</td>
<td>End</td>
</tr>
<tr>
<td></td>
<td>Questions Start</td>
<td></td>
<td></td>
<td></td>
<td>Test</td>
<td></td>
</tr>
<tr>
<td>B</td>
<td>Read difficult version</td>
<td>Read simple version</td>
<td>Compare both versions</td>
<td>Read difficult version</td>
<td>Simple version removed and Comprehension Questions Start</td>
<td>End</td>
</tr>
<tr>
<td></td>
<td>Questions Start</td>
<td></td>
<td></td>
<td></td>
<td>Questions Start</td>
<td>Test</td>
</tr>
</tbody>
</table>

Table 2: Timing Schedule

Maximum reading times were controlled strictly, and timing was an important issue in the set up of the test. To control for the effect of working memory of the simplified version on the answering of the comprehension questions, the Group B had the simplified version taken away from them, and they were asked to re-read the difficult version for 2 minutes before attempting the questions. Note that the Groups S and D had 17 minutes to read their texts. The Both group had only 6 minutes to read the difficult text, then 6 minutes for the simplified version, and then two minutes in which they were instructed to compare the versions. The simple text was then removed and questioning began for all groups. Students were allowed to use their dictionaries throughout the study (except for the TOEIC Bridge Test). This was to keep the reading styles as close to the normal ones students usually adopted. Also, this decision was taken to counter the contention that the Difficult version was probably too hard to read without any support.

The research questions are:
1. Will the group who received the extra instruction (Group B) on the Difficult text outperform those who did not, Group D? (In other words, does help help?)
2. Will the group who read the Simple version, Group S, outperform those who read the Difficult text? (In other words, do readability scales distinguish between higher and lower levels in ELT?)
3. Will those who received the extra instruction (Group B) score lower than those who read the Simple version, Group S? (I.e. what is the effect of simplification?)
A case for linguistic simplification

The hypotheses are:
1. The Group B will outperform the Group D but not the Group S.
2. There will be a significant difference between Groups S and D.
3. There will not be any significant differences between Groups B and S.

Results
Before analysing scores, each test was checked for internal validity. An item analysis on both sets of data from the Bridge and the Diana tests was conducted. This comprised looking for instances where higher level students scored lower than the lower level students, and vice versa. This would indicate that the item needs to be questioned. Often in problematic cases, we find that the question is too subtle for lower proficiency test takers to realise the problems, problems which subsequently floor the higher proficiency takers. The method used for this was to take the top 25% of ranked-by-score scores and then take the bottom 25%. Then, for each question, the total correct responses were tallied for each proficiency group. The reverse was also done: the number of incorrectly answered questions should be higher for the lower proficiency group. There were no instances of irregularity in the Diana study. There were 6 items that were problematic in the Bridge Test. These items were removed from the study.

The statistical software package Minitab version 14 was used for all statistical operations. The next stage was to see if the general proficiency of each student, as measured in the Bridge Test, was the same in the Diana test. If so, it could be said that the Diana test, although it tested a different skill base from the Bridge Test, was still able to separate general proficiency levels. In other words, did the two tests correlate? A Pearson correlation was measured at 0.46 and the p-value was p<0.00. This indicates that there was a weak correlation between the two sets of data. The significance value showed that this result was significant. The two tests are comparable in that they distinguish the same groups of student by proficiency level.

A one-way ANOVA was run on the different groups in both data sets.

<table>
<thead>
<tr>
<th>Source</th>
<th>DF</th>
<th>SS</th>
<th>MS</th>
<th>F</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>Factor</td>
<td>2</td>
<td>367</td>
<td>184</td>
<td>1.49</td>
<td>0.232</td>
</tr>
<tr>
<td>Error</td>
<td>78</td>
<td>9617</td>
<td>123</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>80</td>
<td>9985</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 3: Diana Groups by Bridge Scores

The significance value of the Bridge Test was p = 0.23 (at an alpha level set at 0.05). This
indicated that there was no significant difference between the three groups. In other words, the students in each of the study groups can be considered similar in terms of general language proficiency.

<table>
<thead>
<tr>
<th>Source</th>
<th>DF</th>
<th>SS</th>
<th>MS</th>
<th>F</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>Factor</td>
<td>2</td>
<td>3127</td>
<td>1564</td>
<td>3.66</td>
<td>0.030</td>
</tr>
<tr>
<td>Error</td>
<td>78</td>
<td>33331</td>
<td>427</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>80</td>
<td>36458</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 4: Diana Groups

On the other hand, the p-value returned from the Diana data indicated a significant difference at p=0.03. This may have been expected as the following summary data shows that the means for the Difficult group and the Simple group are very far apart.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Count</th>
<th>Mean</th>
<th>StDev</th>
<th>Min.</th>
<th>Q1</th>
<th>Median</th>
<th>Q3</th>
<th>Max.</th>
<th>Range</th>
</tr>
</thead>
<tbody>
<tr>
<td>Both</td>
<td>28</td>
<td>67.00</td>
<td>19.24</td>
<td>35.00</td>
<td>50.00</td>
<td>70.00</td>
<td>84.50</td>
<td>97.00</td>
<td>62.00</td>
</tr>
<tr>
<td>Difficult</td>
<td>28</td>
<td>56.82</td>
<td>21.96</td>
<td>5.00</td>
<td>45.00</td>
<td>60.00</td>
<td>75.25</td>
<td>93.00</td>
<td>88.00</td>
</tr>
<tr>
<td>Simple</td>
<td>25</td>
<td>71.76</td>
<td>20.73</td>
<td>17.00</td>
<td>65.00</td>
<td>80.00</td>
<td>85.00</td>
<td>95.00</td>
<td>78.00</td>
</tr>
</tbody>
</table>

Table 5: Descriptive Statistics

A few more points may be drawn from this data. The ranges for all three groups are very wide. This suggests the possibility that students were not entirely comfortable with the test format. Those for whom the test was a new procedure may have been put off by the question forms. The Bridge Test simply requires takers to choose from multiple choices, but the Diana test asked them to write in the answer on a blank line. This format would favour those more comfortable with selecting and manipulating syntax from a text. However, the overall test did distinguish between proficiency levels, so this aspect was not considered to be a serious problem. A more reliable indicator is the range between the 1st and 3rd Quartiles. These figures eliminate the very low and the very high scorers. At a glance, there is roughly the same kind of distribution as in the overall figures, that is, the Simple highest, Both in the middle and Difficult lowest. For statistical analyses, however, the overall figures were used.

A second point is that the standard deviations for each group are similar (which indicates
A case for linguistic simplification

a good test validity) but are very wide. The equivalent data from the Bridge Test shows much
tighter figures. The means are similar, the standard deviances are almost half of those in the Diana,
the ranges are closer and the restricted 1st quartile to 3rd quartile range is very narrow.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Count</th>
<th>Mean</th>
<th>StDev</th>
<th>Min.</th>
<th>Q1</th>
<th>Median</th>
<th>Q3</th>
<th>Max.</th>
<th>Range</th>
</tr>
</thead>
<tbody>
<tr>
<td>Both</td>
<td>28</td>
<td>57.57</td>
<td>9.67</td>
<td>40.00</td>
<td>52.00</td>
<td>55.00</td>
<td>62.50</td>
<td>77.00</td>
<td>37.00</td>
</tr>
<tr>
<td>Difficult</td>
<td>28</td>
<td>53.04</td>
<td>12.43</td>
<td>23.00</td>
<td>44.50</td>
<td>53.00</td>
<td>65.25</td>
<td>74.00</td>
<td>51.00</td>
</tr>
<tr>
<td>Simple</td>
<td>25</td>
<td>53.16</td>
<td>11.02</td>
<td>26.00</td>
<td>44.50</td>
<td>52.00</td>
<td>63.00</td>
<td>71.00</td>
<td>45.00</td>
</tr>
</tbody>
</table>

Table 6: Descriptive Statistics by Diana Group

There were statistically significant differences between some of the groups in the Diana study. To
isolate where these differences lay, 2-sample T-tests were carried out. The research hypothesis
was built into the t-test, that is, the t-test of difference was set up to answer the ‘is greater than’
equation.

<table>
<thead>
<tr>
<th>Group</th>
<th>N</th>
<th>Mean</th>
<th>StDev</th>
<th>SE Mean</th>
</tr>
</thead>
<tbody>
<tr>
<td>DianaB</td>
<td>28</td>
<td>67.0</td>
<td>19.2</td>
<td>3.6</td>
</tr>
<tr>
<td>DianaD</td>
<td>28</td>
<td>56.8</td>
<td>22.0</td>
<td>4.1</td>
</tr>
</tbody>
</table>

Difference = mu (DianaB) - mu (DianaD)

Estimate for difference: 10.1786

95% lower bound for difference: 0.9418

T-Test of difference ≠ 0 (vs >): T-Value = 1.84 P-Value = 0.035 DF = 53

Table 7: Two-Sample T-Test Both and Difficult Groups

<table>
<thead>
<tr>
<th>Group</th>
<th>N</th>
<th>Mean</th>
<th>StDev</th>
<th>SE Mean</th>
</tr>
</thead>
<tbody>
<tr>
<td>DianaS</td>
<td>25</td>
<td>71.8</td>
<td>20.7</td>
<td>4.1</td>
</tr>
<tr>
<td>DianaD</td>
<td>28</td>
<td>56.8</td>
<td>22.0</td>
<td>4.1</td>
</tr>
</tbody>
</table>

Difference = mu (DianaS) - mu (DianaD)

Estimate for difference: 14.9386

95% lower bound for difference: 5.1075

T-Test of difference ≠ 0 (vs >): T-Value = 2.55 P-Value = 0.007 DF = 50

Table 8: Two-Sample T-Test Simple and Difficult Groups
The final T-test between the Both and Simple groups was carried out again at the ‘is not equal to’ alternative.

The results fully supported all three research hypotheses. The Both group outperformed the Difficult group significantly at p=0.035. The Simple group outperformed the Difficult group significantly at p=0.007. And no significant differences could be found between the Simple and the Both groups at p=0.196 for the equation ‘S is greater than B’ and at p=0.392 for the alternative ‘S is not equal to B’.

**Discussion**

What does it mean by the claim that the Both group significantly outperformed the Difficult group? Exposure to the simple may change the student in a number of ways. The first, and the point of this study, is that the simple text offered a direct passage not only to the temporary, and transient, understanding of the complex text, it actually helped them develop their understanding
beyond the immediate working memory. The old argument runs like this: when a student hears a simplification, it is the simplification that they understand, not the complex text (Ellis, 1993). Now, we may be able to say that: yes, they do understand the complex text on its own merits. There is the possibility that in the complex vocabulary will be understood in terms of their simpler equivalents. This one-to-one link possibility must take into account the third element: the equivalent meaning in the question lexis. For the one-to-one link to have been correct, the question lexis must have additionally triggered that connection through lexical set association. In the 28 answer sheets, not one subject used the lexis from the simple text in answering the questions on the difficult text. All answers were taken from the difficult text (i.e. paraphrased or quoted directly). This suggests that the simple version was not being remembered, and that the lexis of the difficult text was understood on some level.

One might wonder if the two minutes was enough to purge the working memory. This may be countered in a number of ways beyond the traditional measurement of working memory in typical terms of seconds, rather than minutes or hours. A critical factor in the research design was that of timing. The space between contact with the simplified text was one, but more important was the time-on-task the Both group had. The time-on-task the Difficult and Simple groups had was much more than that of the Both group. Those other groups had 17 minutes to read and re-read their texts. According to other studies rereading is one of the best methods of improving comprehension and recall (see Stephen Brown, 2002, “in L2 (English) research, rereading has also led to improved reading fluency and recall”). In other words, we might expect the Difficult and Simple groups to outperform the Both group on this time-on-task aspect alone. In the end, the Simple group did do the best, but not significantly better than the Both group. A further level was the factor of reading speed. The Simple text had 1315 characters (and 289 words), and the difficult one 1403 (249 words, but remember that they were longer). At a reading rate of 50 words (or 250 characters) per minute, the initial 6 minutes given to the Both group would have stretched their cognitive processing abilities. Then, the Both groups was asked to read another similar length text in the same time. If 50 words per minute was a fast rate, this would have been another reason to expect that the Both group should have scored lowly. They would have been pressed for time. Certainly, they would not have had time to memorise the simple text.

Allowing participants to use a dictionary complicates things a bit. A few problems spring immediately to mind: it is possible that skilled dictionary users get more information than unskilled ones; faster readers might have had more time to access dictionaries than slower ones, and faster does not necessarily mean ‘better’. There was no controlling factor to account for various strategy uses of the dictionary. A further and separate test of dictionary skills may have helped to inform the results of this study. Besides this, the more common sense conclusion of the
dictionary issue is that the dictionary effect should be equal for all groups and thus not really an issue in the study.

The second hypothesis was supported that the readability scales would produce texts that were significantly different in terms of difficulty level. The third hypothesis was also returned. The effect of the simplification on the Both group was enough to render the result significantly close to that of the Simple group. A potential threat to internal test validity is the claim that the simplified reading was so close temporally to the question period that students remembered the simplified version when answering the questions. This claim may be countered in part by the controlling of time-on-task, which was designed to prevent students from memorising either text. Even if the Both group got close to the Simple group’s comprehension score, memory alone does not account for the failure of the Difficult group to do the same, again at least in terms of lexis. Dictionary access was allowed giving all groups access to conceptual meanings. There was no reason any text would be inaccessible to any group significantly, at least in terms of lexis. As for syntax, now maybe that plays a far greater role in the comprehension of texts than was tested in this study. This obviously points to the need for further research.

Conclusion
This study examined the effects of teacher involvement in helping students get through a difficult passage. The types of involvement are numerous. Sibayan, Bautista and Gonzalez, among others, analysed the kinds of teacher interaction and came up with a sizable list of teaching techniques. To turn Sibayan’s list into research terminology, we could say that it would be impossible to construct a research design that simultaneously controls for and measures even more than a few variables observed after receiving a number of treatments. Choosing what this researcher believed to be the single most important technique of supplying teacher aid, glossing, and to eliminate teacher variables, a ‘help’ text was prepared. Basically this text acted as a text-wide gloss in L2 of the difficult version. Having a ‘help’ text allowed students to answer questions on the higher-level text significantly better than those who didn’t. The idea that teacher help during class contact time actually supports more direct language acquisition was upheld significantly.

Acknowledgments
I would like to thank Professor Ian. Gleadall for his helpful comments, support and effort in the preparation of this manuscript and in the study, and Michiyo Masui and Hitoshi Eguchi for allowing their classes to participate in the study.
References


**Appendix**

*The Simplified Diana Text*

Diana Spencer was born in 1961. She was born in Norfolk, in England. Her family was upper class. Diana’s parents divorced when she was very young. The children lived with their father. Diana’s school was in Switzerland. It was for girls only. She worked in England as a nursery school teacher.

Prince Charles was a friend of the Spencer’s. Diana and Charles started to date when Diana was a teacher. They got engaged. They got married in 1981. Many millions of people around the world watched their wedding on TV. Princess Diana was the most liked person in the royal family. She was tall, beautiful and stylish. Women wanted to look like her. The media took her photo all the time. She was the most photographed woman in the world.

Diana and Charles were not happy together. They began to live apart in 1992. Diana gave a famous interview on TV in 1995. She talked about her private life. She told people why she was not happy. Royal people never talk about private things. The TV interview was not usual. People liked watching Diana’s honesty. Diana and Charles divorced in 1996.
Diana still worked to help people after the divorce. She worked with poor people, people who had AIDS and with people who had drug problems.

Diana started to date the Egyptian Dodi al-Fayed in 1997. He was a very rich man. They were in a car in Paris one night. The media chased their car. The car was going very fast. It crashed, and Diana and Dodi died. It was August 31, 1997. She was only 36 years old. People all over the world were very sad. People will always remember her. They called her the “People’s Princess”.

*The Difficult Diana Text*

Born Diana Spencer in 1961 in Norfolk, England, the future princess was raised by her father following her parents’ divorce in her early childhood. After attending a Swiss private girls’ school, she returned and was employed as a nursery school teacher.

Through her family’s acquaintance with Prince Charles, they began dating, became engaged, and in 1981, their marriage ceremony attracted millions of television viewers around the globe. Her position as the most popular royal was reflected by women worldwide clamouring to look like her and by her becoming the world’s most photographed person. Being tall, beautiful and stylish virtually guaranteed her media coverage wherever she went.

Following conjugal difficulties, the royal couple announced their official separation in 1992. Breaking the royal protocol never to publicly reveal private affairs, Diana appeared in an unendorsed television interview in which she discussed personal matters and the source of her unhappiness. Her honesty appealed to viewers, but undoubtedly accelerated her divorce, which was finalised in 1996.

Diana’s efforts against poverty, AIDS, land mines and drug abuse continued after the divorce, reinforcing popular affection in her.

By 1997, Diana had begun a relationship with an Egyptian millionaire, Dodi al-Fayed. After an evening out, photographers trailed their limousine resulting in a high-speed dash through Paris, ending dramatically and drastically with the crash that killed both Diana and al-Fayed. The death of the 36-year-old princess caused world to grieve. The date of August 31, 1997 and the “Peoples’ Princess” will forever be in their memories.