



TEXT DIFFICULTY AND ACCESSIBILITY: READING FORMULAE AND EXPERT JUDGEMENT

GLENN FULCHER

English Language Institute, University of Surrey, Guildford, UK, GU2 5XH

Text difficulty, or text “accessibility” is an important but much neglected topic in Applied Linguistics. Establishing text difficulty is relevant to the teacher and syllabus designer who wish to select appropriate materials for learners at a variety of ability levels. It is also critical to test developers in selecting reading texts at appropriate levels for inclusion into the reading sub-tests of examinations. Writers of texts for various audiences also need guidance related to the range of factors which make texts more or less accessible. In all these cases, however, decisions are still made very much on intuitive grounds. This research specifically addressed the concerns of text writers, but the findings are still relevant to the first two concerns. The research involved the analysis of a corpus of texts, and shows that factors which make the texts difficult, or less accessible, include poor linguistic structure, contextual structure, conceptual structure, and unclear operationalisation of the reader–writer relationship. It is argued that factors which are not considered in traditional readability formulae are more important determinants of text accessibility. © 1997 Elsevier Science Ltd

1. INTRODUCTION

The general concern to be addressed in this research was the accessibility of texts produced by the Overseas Development Administration (ODA)¹ for a large and varied audience. The ODA is sensitive to the need to produce texts which are accessible by many readers, and the research described in this paper represents a first but vital step in identifying factors which make texts more or less accessible. The practical application of such an investigation needed to be addressed in terms of writers of texts, and guidelines which could be formulated to help them in their task.

The focus of the research does, however, have much more general applications. This is especially the case in the selection of reading material for pedagogic purposes, or the selection of texts for inclusion in reading sub-tests of examinations. Research in this area is particularly pertinent at the present time, as text difficulty is re-emerging as an area of great concern in language teaching, materials writing, and establishing equivalent difficulty levels of texts in reading tests.

Readability, or “text difficulty”, has been, and remains, an area of concern for all those who need to establish the appropriacy of a given text for a pedagogic purpose. And these concerns have historically been more practically than theoretically oriented. At the most basic level, teachers recognise that giving students reading material that is “too difficult” is damaging to the learning process, and demotivating to the student. It has been recognised that there are a range of reader factors which effect the reading process, including motivation, background knowledge and previous reading experience; but as these are essentially variables beyond the control of the teacher, it is facets of the text which have received the most attention. Among those most often used to assess text difficulty are sentence length and vocabulary, and it is precisely those facets of the text which have been used to attempt to *predict* the difficulty of texts.

The notion of *prediction* has been central to virtually all text difficulty studies. If learner variables are responsible for much of the variance in reading ability, or the ability to comprehend text, then the best we are able to do is estimate the difficulty level of a given text for an idealised, “average” reader of a given age. The result of this quest has been a set of formulae which help teachers or testers to estimate the difficulty level of texts, and select the appropriate ones for their intended audiences.

Studies which use readability formulae for this purpose often claim increased accuracy of match between text and reader (Jones, 1995), and the readability formulae themselves are being widely used in the development of editing programs for desk-top computer packages (Giles, 1990). The use of readability formulae in this field may encourage the naive application of simplistic calculations to the automatic advice given to novice writers—a practice that should be discouraged. Readability formulae have also been used in the field of testing, most notably as a “control” on the difficulty of text levels in reading tests (Davies and Irvine, 1996). Although Davies and Irvine (1996: p. 170) acknowledge that:

Such a measure [Gunning–Fog index] is inevitably crude... Nevertheless, unsophisticated measures of this kind are useful approximations of the ordering of texts in terms of difficulty and complexity.

Despite the variety of uses, it has always been known that the prediction of difficulty was only partial at best, as many of the reader variables had been left out of the equation. It is now possible to go further than this, and claim that the readability formulae developed from the late 1940s through to the 1970s do not take into account any more than the grossest aspects of what make a text difficult, and other more important aspects will be dealt with in this paper, using texts from ODA publications as examples. Studies such as those of McAdams (1993) which indicate that topic/reader interest are more important than sentence length, and Agnihotri and Khanna (1992) which show that conceptual difficulty and textual organisation are crucial factors on estimating difficulty, point the way forward to more accurate assessments of text accessibility.

1.1. Research questions

The research which is described in this paper was designed to answer three key questions, which are:

- What are the difficulty levels of a range of texts published by the ODA?
- What makes texts difficult or less accessible?

What advice/guidelines can be given to writers of texts to make texts more accessible?

1.2. Methodology

The methodology used in this research was designed to answer the explicit questions above, and to develop an insight into the differences between estimating text difficulty between quantitative measures and the judgement of experts in the field of reading.

1.2.1. Stage 1. A range of texts (the corpus) were collected from the ODA and converted into computer readable format. The Flesch reading index was calculated for each of the texts. This particular readability measure was used, because it is the only reading difficulty measure which was written for the adult rather than the child reader (Harrison, 1980: p. 77). As our primary concern is with texts aimed at adult readers, it was clear that no other alternatives were equally appropriate. The Flesch rating was then used to (a) select a smaller number of texts for analysis in stage 2 of the project, and (b) report predicted range of text difficulty.

1.2.2. Stage 2. In stage 2, a group of five experts in the field of reading/writing were presented with extracts from the texts selected in stage 1.

The experts were asked to:

- sequence the texts in order of perceived difficulty individually (step 1),
- agree on a sequence of text difficulty (step 2),
- agree on the reasons for their decisions in step 2 (step 3).

Step 1 was carried out by the experts individually, at their own pace and in their own time. The experts were then requested to attend an all-day meeting to agree on a sequence of text difficulty, and agree on reasons for the final sequence. On this day, detailed notes of the discussions were kept, and used when writing up a report for the ODA, and this paper.

The experts were also explicitly informed on the research day that the focus of the discussion should be related to providing possible guidance for writers of texts.

1.2.3. Statistical analysis. The following statistical analysis was carried out:

- Spearman rank order correlation between individual experts which can be interpreted as a measure of inter-expert agreement prior to conducting step 2,
- Spearman rank order correlation between agreed difficulty rate and predicted difficulty rate, which is a measure of agreement between predicted and expert difficulty assessments.

2. TEXT DIFFICULTY ESTIMATES (PREDICTED)

2.1. List of selected texts

Texts from the following publications (provided by the ODA) were put into computer readable format:

- Action on Health and Population (foreword and body) (1992)
- Britain and Bangladesh (1993)
- British Overseas Development Newspaper (1996*b*)
- Britain and Ghana (1996*a*)
- Guide to Working Overseas (1995*c*)
- Annual Review, 1994 (introduction and body) (1994)
- A selection of Press-releases
- Making aid work for women (1995*a*)

2.2. The Flesch formula

2.2.1. *Calculation of the Flesch formula.* Difficulty estimates were predicted using the Flesch formula (Flesch, 1949). This index was chosen because it is easily the best known readability formula in current use, and it has the advantage of producing a numerical result which, unlike alternatives, is not linked to a United States grade level (year at school) estimate. Flesch was one of the few readability researchers who was interested in texts written for adults rather than children, and so he designed an abstract scale which runs from 0 to 100, with more difficult texts having a lower score, the average being 50. Alternatives are linked to United States Grade Levels, the equivalent of attempting to link the difficulty level of a given text to a school reading age. As the texts used in this study were written for adults, the Flesch is the most appropriate measure for our purposes. Flesch formula is:

$$\begin{aligned} \text{Reading ease score} &= 206.835 \\ &- (0.846 \times \text{number of syllables per 100 words}) \\ &- (1.015 \times \text{average number of words per sentence}) \end{aligned}$$

2.2.2. *Estimation of text difficulty.* It is now possible to present the texts used in this study (see 2.1. above) in Table 1, with the Flesch reading ease score, and an estimate of the reading age required for the text to be comprehended. Table 1 demonstrates that all the texts in this sample are predicted to be fairly difficult to difficult, with no text falling within the range which an "average 16 year old" would be expected to comprehend.

The figures presented in Table 1 were used in the selection of a smaller number of extracts from texts for the analysis in stage 2 of the research. Extracts were used from those

Table 1. Difficulty estimates for texts used in this study

Text	Flesch reading ease score	Estimated reading age
Action on Health and Population * (AHP) (body) [Text 2] (1992)	36.4	21
Action on Health and Population * (AHPF) (foreword) [Text 4] (1992)	53.9	17
Britain and Bangladesh* (BB) [Text 7] (1993)	44.8	19
British Overseas Development* (BOD) [Text 3] (1996 <i>b</i>)	43.1	20
Britain and Ghana (BG) (1996 <i>a</i>)	46.4	19
Guide to Working Overseas* (GVO) [Text 6] (1995 <i>c</i>)	47.5	19
Annual Review 1994 * (AV) (body) [Text 5] (1994)	40.1	20
Annual Review 1994 (AVI) (introduction) (1994)	44.0	19
Press release* (PR) [Text 8] (1995 <i>b</i>)	39.5	20
Making aid work for women* (MAW) [Text 1] (1995 <i>a</i>)	49.3	19

documents marked with an asterisk (*) in Table 1. This selection gave extracts from a range of publications and difficulty levels.

It should be noted that in the case of Action on Health and Population, the foreword was included as an extract in its own right, and a separate extract from the body of the publication was also included in the sample for analysis. This was because the foreword was estimated to be much easier than the text in the body of the document. Indeed, these two texts were the easiest and the most difficult respectively. Finally, the text number after each selected extract refers to the random sequence in which the extracts were ordered for judges to read.

2.2.3. Limitations of Flesch and other predictors. It is clear that the Flesch readability formula is a function of the length of sentences and the size of words in a text, and nothing else. These factors are not, in themselves, what makes a text difficult or easy; but they are *predictors* of text difficulty. Indeed, as Harrison (1980: p. 45) indicates, studies have shown that predictors such as the Flesch formula tend to correlate at around 0.64–0.70 with other measures of text difficulty, such as cloze, or teacher (pooled judgement) estimates of text difficulty for particular groups of students. Predictors are short-cuts to making tentative decisions in the absence of more precise information. This is so, because there is no direct link between estimates (or predictions of) text difficulty, and reader comprehension. Readability formulae do not take into account any factors which relate to other aspects of text, nor do they take into account the *reader*, and so do not take into account a number of key variables in the reading equation. What do we ignore by using predictors like the Flesch formula?

Other aspects of text, including:

- font size and type
- illustration and colour
- specialist use of lexis
- conceptual and propositional complexity
- textual organisation
- syntax

Factors relating to the reader include:

- background knowledge
- interest in the subject
- level of general education
- reading speed and strategies

It should therefore be stressed that there is no simple linear relationship between estimated text difficulty and reader comprehension. In one piece of research, however, it is simply not possible to cope with all these variables. It is essential to concentrate on a small number of variables which affect text accessibility in any one study, and build upon findings as research progresses. In the present case, the emphasis was upon the text *as text*. A deliberate decision was taken to remove illustrations and graphics from all texts, so that

Table 2. Spearman rank order correlation matrix for inter-judge agreement, and agreement between final expert difficulty ratings and predicted difficulty (Flesch)

Judge	Judge 1	Judge 2	Judge 3	Judge 4	Judge 5	Agreed ranks
Judge 2	.41					
Judge 3	.21	.26				
Judge 4	.57	.60	.50			
Judge 5	-.29	.43	.00	-.21		
Agreed ranks	.43	.86	.64	.67	.38	
Predicted ranks	.81	.43	-.02	.41	.05	.31

the focus was on the writing. It was envisaged that illustrations would be investigated in a separate study at a later date, which could take into account the ways in which graphics set up schematic expectations which resulted in better understanding of a text, irrespective of the difficulty of the text itself.

3. INTER-EXPERT AND EXPERT-PREDICTED DIFFICULTY AGREEMENT

An enduring problem of research which relies on expert judgement is that of agreement between experts (Alderson, 1993). In this research, Spearman rank order correlations were calculated for the rank orders of individual experts, which can be interpreted as a measure of inter-expert agreement prior to the research day on which they were asked to discuss the texts. It was anticipated, in line with previous research, that agreement would be minimal, and for this reason the research design included steps 2 and 3 in stage 2 (see 1.2.2. above). They were also calculated for agreed expert difficulty rating and predicted difficulty rating. The rank order correlation coefficients are presented in Table 2.

The average inter-judge agreement is only 0.25, indicating that these judges, working alone, were not able to agree substantially on the rank order of these texts. The correlation between predicted difficulty and agreed difficulty fared little better, at 0.31. However, this does not mean that there are no patterns in the data. In Fig. 1, we isolate those texts on which the judges agree, when ranking them independently.

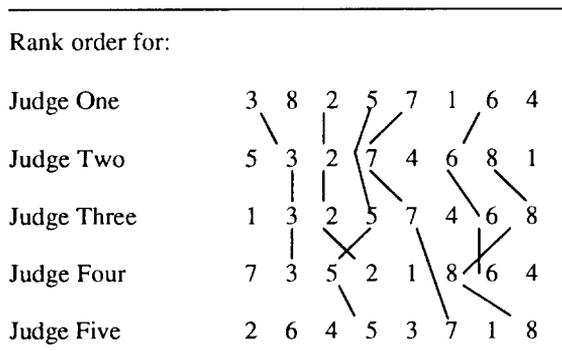


Fig. 1. Judge agreement.

We may make the following observations from this data:

1. The expert judges were not, on the whole, able to agree in rank ordering the difficulty of the texts when working independently. This was made explicit during the research day when the judges were asked to come to agreement on the rank order of difficulty, when each began to make their own criteria explicit. The low initial agreement confirms previous research which indicates that expert judgement can differ radically, and justifies research designs in which these experts are brought together and asked to come to an agreed position. Nevertheless, it should be noted that most of the disagreement is based around just two extracts: nos 1 and 4.
2. The correlation between agreed and predicted ranks was low, and this would have been worrying, had it not been for the fact that the experts explicitly said that one of the criteria they were *not* using in making difficulty judgements was sentence length: sentence length being one of the two criteria used in predicting difficulty in the Flesch formula. Indeed, in Section 4.1. below, it will be seen that very specific criteria were used by the expert judges working as a team, which concentrate on aspects of discourse and organisation. It is not therefore surprising that the rankings of predictors and expert judges did not coincide.

The judges did agree that the texts they were asked to rate were generally difficult, as predicted by the Flesch study, but this research has highlighted the fact that predictors of text difficulty do not correlate with expert judgements. The question which cannot be avoided is whether it is preferable to rely on the information provided by the predictor, or the expert judgements. This will be a question to which we return again, but here it should be stressed that the purpose of this research was not to ask judges to rank order texts according to a well defined construct of reading difficulty or text accessibility, but to establish the construct. It is therefore more appropriate to conduct a further study in which the difficulty ranks established here are confirmed or rejected on the basis of the triangulation of data from independent sources. This is an issue of profound theoretical and practical importance which should be treated independently of the inherently useful criteria and guidance which were developed during this research.

4. EXPERT JUDGEMENTS AND TEXT DIFFICULTY

Section 4 is an account of the discussion of the expert judges, which took place at a full-day meeting on 17 June 1996.

4.1. *Explicit criteria*

The expert judges stated explicitly that their criteria for deciding on text difficulty was not based upon consideration of factors such as sentence length. In the initial general discussion, four areas of concern were agreed upon:

- linguistic structure
- contextual structure (purpose and audience, context of use, information gaps, layout and visual support)

- conceptual structure (degree of familiar and unfamiliar text content, informational relationships, degree of abstractness, and temporal structure)
- reader–writer relationship (use of pronouns, tense and voice).

It was quickly agreed that there was very little to differentiate these texts in terms of purely linguistic structure. The texts were found to score differently on other criteria, and the extent to which each judge weighted the criteria was agreed to be the cause of discrepancies in individual ranking (and hence low correlation coefficients). Key factors were, however, agreed.

The most important factor was the notion of the concept of *Audience*, the extent to which the writer has a clear concept of who the text is meant for. Writers always have some concept of audience in mind, whether this is clear or vague. However, the clearer the notion of audience, the more coherent the text is likely to be, and hence easier to read. Some of the texts in the corpus analysed did not show a clear sense of audience, and some of them did not show a clear concept of text *Purpose*, the second most important criterion to emerge from the discussions. With an unclear concept of audience, it is frequently the case that the purpose of the text, the message which the writer wishes to communicate and the topical focus, is confused and unclear.

Directly related to these two concepts are the ability of the reader to predict the content of a text, especially information which is to come. *Prediction* is important in reading, and if predictions are not met, the reader may wonder what the text is about. Failure to meet reader predictions comes about as a result of lack of purpose and topical focus. One factor which can lessen the reader's ability to predict is unmarked *Topic Switching*. The *Information Load* which the texts carry is also an important variable. Some texts require the reader to retain large chunks of information in short term memory, but this can sometimes be counter-balanced by extremely clear textual organisation.

With reference to Section 3 above, it can be seen why the use of these criteria did not generate agreement between the Flesch difficulty estimates and the expert judges' agreed difficulty estimates. Indeed, it was even suggested that in using these criteria the word "difficulty" might be inappropriate, and text "accessibility" would be more appropriate. In what follows, the terminology of this report does switch between the notion of "difficulty" and "accessibility" in a fairly fluid way, depending upon whether a negative or positive implication is required.

Using the criteria outlined above, the team agreed upon a rank order, and organised the texts into three clusters:

- Cluster 1—Difficult texts 3, 2 and 5
- Cluster 2—Moderately difficult texts 7, 4 and 1
- Cluster 3—Easier texts 6 and 8.

It is worth comparing this sequence with the difficulty estimates in Table 2 again. Notice that text 8 was predicted to be the second most difficult text, and yet after discussion, it

was agreed by the experts that this was in fact the easiest to read. The one judge who originally decided that text 8 was in fact very difficult to the alteration of the sequence. This alone reduces significantly the correlation between the expert judges' sequence and the predicted difficulty levels. Texts 4 and 1 which the judges could not agree upon independently, were clustered into the moderately difficult category, perhaps as something of a compromise.

In what follows, we will treat each cluster in turn, selecting texts 3, 7 and 8 for discussion, as representative of each cluster, to highlight specific aspects of text accessibility. The following sections (Sections 4.2.–4.4.) reflect the discussions which took place between the assembled expert judges, and their agreed position.

4.2. Cluster 1 (Difficult): Text 3—Extract from British Overseas Development (1996b) (Appendix 1)

Text 3 (Reproduced as Appendix 1) appears to have two purposes which conflict with each other. One of its purposes is informational, to provide the reader with details of the *Reflect* programme, but the other purpose is clearly to improve public relations. Adding to its difficulty is the fact that topic development is somewhat unclear, and information load is exceptionally high.

At the beginning of the text there is a lexical problem with connecting "Reflect" in the second sentence to the programme introduced in the first sentence. It is necessary to make this connection before the text can be comprehended, and even for good readers, it may be necessary to re-read the first sentence to make sense of the second. Even if this connection is made, there is still the problem that the writer *assumes* the reader will know more about "Reflect" than he or she can reasonably take for granted. However, this is not as noticeable as the attribution of opinion in the text, and it is this which gives the text much of its sense of dual purpose. It is disruptive to read an opinion or view, and be told whose opinion it is at the end of the sentence, as in this example:

Education continues to be a major emphasis in the British aid programme, she said.

These problems are compounded by poor text organisation, and a lack of signalling devices.

Take these two sentences from the text:

The majority of children in developing countries only go to primary school, and the ODA will strongly support good quality and wider access to educational primary level. Literacy is not a magic wand, Lady Chalker continued.

It is clear that the second sentence requires a the marker "However" in order to make it interpretable in relation to the first sentence. This is true of most sentences. As there is no clear structure to the text, and no discourse markers to help interpretation, it is necessary for the reader to continually make inferences of this sort which may be potentially misleading. Relations between propositions need to be clearly signalled in written text.

In terms of text organisation, there is one extremely misleading signal. This occurs here:

The ODA has recently refined its activities into four policy aims. One of these....

Normally, this signals enumeration. The following text is to be organised: One of these.... The second is.... Thirdly.... Finally.... In this extract (which is in fact the entire text) we are only given one of the four policy aims.

It should be noted that the sentences in text 3 are not particularly long (one of the two criteria in calculating the Flesch index), but here this may be a disadvantage to aiding comprehension. They do not contain the information necessary to be able to process the text without a great deal of commitment from the reader.

The text reads as if it had been written up from the notes of someone listening to a speech, and the editorial process has left the weaknesses intact. The accuracy of this interpretation may be checked by a reader introspecting on what they are doing when reading in order to reach an understanding of what the text is about: in order to construct meaning it is necessary to read it as if reading aloud, for it is the imagined intonation of the voice superimposed on the text that provides the meaning, rather than the text itself.

In order to demonstrate this, we will rewrite the first sentences of the text in order to render it in acceptable written English which makes it more accessible to the reader:

A radical new adult literacy *programme called Reflect* will be launched at the end of the month by Action-Aid, and could dramatically improve world literacy rates. Pilot trials of *the programme* have shown that it could be two to three times more effective than conventional approaches to reading and writing.

This is *because* it abandons the traditional textbook *in favour of* groups of students creating their own learning materials related to local issues.

It was agreed that text 3 would be difficult to read not because of long sentences, difficult syntax or lexis, but because the writer had not taken the necessary care to ensure that the readers would be able to get a message from the text without taking a good deal of trouble.

4.3. Cluster 2 (Moderate): Text 7—*Britain and Bangladesh (1993) (Appendix 2)*

Text 7 has a clear introduction, and the organisation is chronological. There are, however, places where there are inconsistencies. Problems were noted with the relationship of general to specific material, in particular. Movements between information types should be clearly signalled.

The text was also found to have a relatively high information load, but this is compensated for by good organisation. An example of this is in the third paragraph:

A growing concern: health and population Bangladesh's population is growing at a rate of 2.3 per cent a year, and the mortality rate for mothers, newborn babies and infants is also very high. These statistics underline the close relationship between health and population. In an overcrowded city or village complex, where there is little or no clean water or sanitation, conditions such as diarrhoea can spread rapidly. And if family planning and other health services are inadequate, fewer men and women will have the choice to limit their families, and populations may continue to grow.

Here, the sentence beginning “In an overcrowded city...” avoids problems encountered in other texts with abstract information, by providing exemplification. This clearly signalled pattern of abstract–exemplification is of major benefit in structuring texts which are reader friendly.

Nevertheless, as the “most difficult” text in the moderately difficult cluster, some problems were isolated. The first of these was the complex syntax. In many of the sentences there is considerable distance between the subject and the verb, and there are long subordinate clauses separating or preceding them. Although the sentence beginning “In an overcrowded city...” is an excellent use of abstract–exemplification patterning, it is a difficult sentence to process:

In an overcrowded city or village complex [subordinate clause 1], where there is little or no clean water or sanitation [subordinate clause 2], conditions such as diarrhoea [subject] can spread rapidly. [verb group].

Here is another example:

Although its culture is rich and its people resilient, Bangladesh is one of the world’s poorest countries. And the challenges it faces—high population growth, low levels of literacy and a shortage of economically useful natural resources—are complicated by a volatile climate and difficult terrain.

In this case, the second sentence should really have been rendered as two sentences:

Bangladesh faces the challenges of high population growth, low levels of literacy, and a shortage of economically useful natural resources. The challenges are complicated by a volatile climate and difficult terrain.

Secondly, the text assumes that the reader can process the value of information given. This is not always the case, and when assumptions are not accurate, readers may fail to comprehend. Take this example, which provides particularly acute problems:

Bangladesh’s population is growing at a rate of 2.3 per cent a year, and the mortality rate for mothers, newborn babies and infants is also very high.

It is not clear to an uninitiated reader that 2.3% is particularly high. In connection with inflation, for example, this figure is very low. It is only at the end of the sentence when the reader encounters “also very high” that it is necessary to reprocess the preceding information, understanding that in this context 2.3% is actually very high.

4.3. Cluster 3 (Easy): Text 8—A Press Release (1995b) (Appendix 3)

Text 8 was predicted to be the second most difficult from the Flesch study, but expert judges agreed after discussion that it was actually the easiest text in the corpus. The Flesch formula takes into account number of syllables in words, and this text has many multisyllable words, including those for numbers. The only problematic part of this text appears to be the second paragraph, which is not connected to anything as it stands. This paragraph should have been part of the first paragraph, operating as an example of the general sentence which is currently the last of paragraph 1. It should perhaps also be noted that the writer assumes that the reader will understand the acronym NGO, which cannot be taken for granted. But as this occurs in examples, it is less important than it would be at the beginning of the text.

Although information content is high, it is quite accessible because of the way in which information is structured, and the figures, which constitute most of the information load, do not really matter; the reader only needs to comprehend the fact that they are large. The message can then easily be comprehended: The ODA gives a lot of money, and it is spent in x, y and z ways.

The text is well focused on Africa, and the first two paragraphs (which we have noted should have been one) provide the key general message that the ODA (Britain) is providing a large amount of aid to help Africa, which is suffering from drought. The rest of the text enumerates examples. It should also be noted that in each of the successive examples, a similar pattern is used in presenting information [In (country) Britain (is doing) X]. Repetition of structure means that once a reader has comprehended one of the examples, the others are much easier to read, and each successive example emphasises a single focused point. Text purpose is transparent, and everything in the text contributes to communicating a single message.

5. FEATURES OF CLUSTERS: ADVICE FOR WRITERS

The panel of experts agreed on the following defining features for each of the text clusters.

5.1. Cluster 1

The texts in cluster 1 have in common the absence of an awareness of the reader, and a lack of clear guidance regarding text goal. This results in a shift of focus as the text develops. The reader is not easily able to predict what is going to come next in these three texts, and this makes them much more difficult to comprehend.

It is also suspected that texts 3 and 5 began life as spoken discourse. For writers, it is important to be aware of the need to change this to acceptable, cohesive written form. (Some features of spoken discourse, such as the question–answer pattern, can be used to great effect in writing, but this is not the case here).

There is also a great deal of topic switching in these texts, as a result of no clear identification of the purpose of writing. Acronyms are widely used without explanation, and unjustified background knowledge assumptions are made. This is not necessarily shared background with the reader, and makes the texts inaccessible to the non-initiate.

These features increase an already heavy cognitive load, and there is little interactivity or personalisation which could help. Textual features such as titles and subtitles could have been used to provide the reader with help, but these are not present; indeed, this absence of such organisational devices is probably a reflection of the lack of clear focus. Writers should be asked to give each paragraph a title which is clearly related to the title which comes before and after. If they cannot do this, they should be encouraged to question what the text is about!

The passages in this cluster are difficult mainly because they are poorly written.

5.2. Cluster 2

Texts in cluster 2 were more accessible than those in cluster 1 because of the use of personalisation and exemplification. Information structure was, however, rudimentary.

5.3. Cluster 3

One of the main criteria for selecting these texts as being easier is the reiteration of structure and information. Reiteration of structure and the use of exemplification is a much sounder basis on which to build comprehension.

There are also linguistic marks of clearer focus, as a direct result of the writers clear conceptualisation of the text's audience, purpose, and topic.

5.4. Guidelines for writers

Once the individual texts had been assessed by the expert judges for their accessibility, the next stage was for them to transform their perceptions of the types of problems which had been identified in the texts into advice for text writers.

From the analysis of these texts the researchers found it possible to frame a number of recommendations for writers of this kind of text and, indeed, for writers in general. These are:

- Have a clear idea of your audience, including age, background knowledge, interest in your topic, reason for reading this text.
- Define the purpose of the text explicitly before beginning to write. Is the purpose to inform, for example, or to promote the role of an organisation?
- Carefully select which topics are going to be used to achieve the purpose of the text. This is a serious question whenever the purpose of the text is not explicitly to inform the reader of, for example, the current situation in country x as a result of disaster y.
- Decide what the relationship between you and the reader is going to be. Do you wish it to be somewhat distant, or do you wish to involve the reader through the use of appropriate pronouns (we, us, I)? What other methods of personalisation can be introduced to involve the reader in the subject.
- Decide on an overall structure for the text before beginning to write, in relation to the message to be conveyed, the audience, and the relationship to be established with the readers. Is it more appropriate to structure the text as a narrative with examples, an abstract-example pattern, according to geographical areas, chronologically, a question/answer pattern, or a problem/solution structure.
- Consider the level of abstraction required in the text. If the level of abstraction is high, information load will be increased. Consider carefully how and where clear examples which characterise the abstractions can be introduced.
- What typeface, layout, illustrations, and other physical characteristics of the text (such as indenting sections of the text, bullet points, bold face or italics) can be used to effectively enhance the accessibility of the text?

All of these considerations are important in the writing of effective texts, as reflected in current thinking on writing (see, for example, Fulcher, 1997, in which each contributor stresses the importance of audience and purpose specification).

6. PROBLEMS, OMISSIONS, AND FUTURE DIRECTIONS FOR RESEARCH

6.1. The absence of the reader as a variable

In this study, the reader has been absent as a variable. The follow-up study suggested in 6.1. would introduce the reader to the investigation, and help triangulate on data already obtained. However, as discussed in 2.2.3 above, there are other reader characteristics which should not be overlooked. Of particular importance to this research is level of interest and background knowledge. These factors would almost certainly influence the willingness to read texts and, when reading, the ability to comprehend (see Clapham, 1996).

This is a potentially complex problem, in which many variables are at work. In this study we have been able to investigate only one of these variables: the text. In any future study the number of variables can be expanded, based upon the work which has already been done.

6.2. Layout and illustrations

Layout, typeface and illustrations are critical to understanding the message of a text. A key illustration or photograph can activate background knowledge and ability to predict what comes next in a text, even if the text itself is badly written. In this research, however, a conscious decision was taken to exclude all of these variables from the study. The focus was the text itself: the words on the page. Any additional material was removed, to reduce the number of potentially confounding variables.

Nevertheless, it is clear that at some stage these variables must be tackled. Layout, typeface and illustrations interact with the text and the reader to help or hinder the reading process. Little research has been undertaken on this aspect of text production, but needs to be taken into account as one of the other variables in a complex process.

7. CONCLUSIONS

This research has successfully identified a number of areas in which advice and guidance could be given to writers of texts. These guidelines should be of considerable use in helping writers to produce texts which are more accessible to readers; it may also be possible to device "writing workshops" based on these guidelines, to improve writer efficiency.

However, the findings should be of general interest to all those who need to assess the "difficulty" or "accessibility" of a text for any of the purposes listed in the introduction to the paper. The common reliance on the results of predictor readability tests which are currently in use cannot be taken uncritically as providing a short cut route to solving selection problems for pedagogic or testing purposes.

This study began as a piece of applied research, but is significant in that it has generated a number of questions and issues which are of wider importance and significance to language teachers, testers and those interested in reading pedagogy. It is timely to ask whether tools which were developed 40–50 years ago can continue to serve the purpose of helping us to choose texts wisely, or whether we need to develop a new awareness of the complexity of text and the nature of reading.

Acknowledgements—The author would like to thank the Overseas Development Administration, in particular Mr Ron Fosker (Deputy Head of Information) and Amanda Hewett (Publications) for supporting this research, and giving permission for the publication of this paper. I would also like to acknowledge the crucial involvement of the following academics who gave up their time to take part in the research described in this paper. Without their co-operation and lively discussion, this research would not have been possible: Jonathan Charteris-Black (University of Surrey), Smiljka Gee (University of Surrey), Amos Paran (University of Reading), Linda Taylor (University of Cambridge), Cyril Weir (University of Reading).

NOTES

¹The Overseas Development Administration is a Department of the United Kingdom Government concerned with “promoting the development or maintaining the economy of a country or territory outside the United Kingdom, or the welfare of its people”. Its main areas of activity relate to supporting economic reform, increasing production, financing activities which relieve poverty, promoting human development through better education and health care, promoting the status of women, helping tackle environmental problems, and helping achieve good government.

REFERENCES

- Agnihotri, R. K. and Khanna, A. L. (1992) Evaluating the readability of school textbooks: an Indian study. *Journal of Reading* 35(4), 282–288.
- Alderson, J. C. (1993) Judgements in language testing. In *A New Decade of Language Testing Research: Selected papers from the 1990 Language Testing Research Colloquium*, ed. D. Douglas and C. Chapelle, pp. 46–50. Teachers of English to Speakers of Other Languages Inc., Washington, DC.
- Clapham, C. (1996) *The Development of IELTS: A Study of the Effect of Background Knowledge on Reading Comprehension*. Cambridge; Cambridge University Press.
- Davies, A. and Irvine, A. (1996) Comparing test difficulty and text readability in the evaluation of an extensive reading programme. In *Performance Testing, Cognition and Assessment: Selected papers from the 15th Language Testing Research Colloquium, Cambridge and Arnhem*, ed. M. Milanovic and N. Saville, pp. 165–183. Cambridge, Cambridge University Press.
- Fulcher, G. (1997) (ed.) *Writing in the English Language Classroom*. Prentice Hall International/Macmillan, Hemel Hempstead, in conjunction with the British Council. ELT Review Series.
- Flesch, R. F. (1949) A new readability yardstick. *Journal of Applied Psychology* 32, 221–233.
- Giles, T. D. (1990) The readability controversy: a technical writing review. *Journal of Technical Writing and Communication* 20(2), 131–138.
- Harrison, C. (1980) *Readability in the Classroom*. Cambridge, Cambridge University Press.
- Jones, K. H. (1995) Readability of textbooks for technology education. *Technology Teacher* 55, 28–32.
- McAdams, K. C. (1993) Readability reconsidered: a study of reader reactions to fog indexes. *Newspaper Research Journal* 14, 50–59.
- ODA (1992) *Action on Health and Population*. London: ODA.
- ODA (1993) *Britain and Bangladesh: Partners in Development*. London: ODA.
- ODA (1994) *British Overseas Aid, Annual Review*. London: ODA.
- ODA (1995a) *Making Aid Work for Women*. London: ODA.
- ODA (1995b) *South Africa Drought: Britain Sends More Aid*. Press Office News Release, 16 November. London: ODA.
- ODA (1995c) *The ODA Guide to Working Overseas for the Aid Programme*. London: ODA.
- ODA (1996a) *Britain and Ghana: Partners in Development*. London: ODA.
- ODA (1996b) *British Overseas Development*, Issue 46. London: ODA.

APPENDIX 1

TEXT 3

World literacy rates could be improved dramatically as a result of a radical new adult literacy programme launched by the charity ActionAid this month. Pilot trials of the programme in three continents have shown Reflect could be two to three times more effective than conventional approaches to reading and writing. Reflect abandons the traditional textbook and hands over the curriculum to learners. Working in groups, students create their own learning materials relating to local issues. Reflect teachers use these materials to introduce numbers and the written word. Minister for Overseas Development Baroness Chalker, who launched the programme on 11 March, said Reflect is important because it enables students to relate to what they are learning. Despite some signs of improvement, she said, world literacy rates remain low. Inadequate access to education and poor quality education in many countries means about 950 million adults are functionally illiterate, and over 100 million children are missing school. Teaching people to read and write provides them with the means to develop, Lady Chalker said. The ODA is therefore glad to play a part in the Reflect programme. Education continues to be a major emphasis in the British aid programme, she explained. The ODA has recently refined its activities into four policy aims. One of these is "to help people achieve better education and health, and to, widen opportunities—particularly for women". Women will continue to be a key target group for literacy programmes, as will children, who need to acquire the basic skills of literacy and numeracy on which all their further education and development depend. The majority of children in developing countries only go to primary school, and the ODA will strongly support good quality and wider access to education at primary level. Literacy is not a magic wand, Lady Chalker continued. Teaching people to read does not make all their problems disappear. Literacy skills do not improve people's lives automatically. In the past many literacy programmes failed because students and teachers discovered the magic wand was an illusion. New programmes such as Reflect are much more practical, she said. They take into account local needs, attitudes, conditions and thoughts. Systems such as Reflect emphasise dialogue and enable people to participate locally. They are about empowerment as well as passing on skills. This promises to be just the start of a new era in the teaching of literacy, she said.

APPENDIX 2

TEXT 7

Since the birth of Bangladesh as a nation in 1971, its government has worked in partnership with Britain towards a common goal: the development of Bangladesh and a lasting improvement in the living and working conditions of its 110 million people. Although its culture is rich and its people resilient, Bangladesh is one of the world's poorest countries. And the challenges it faces—high population growth, low levels of literacy and a shortage of economically useful natural resources—are complicated by a volatile climate and difficult terrain.

British aid to Bangladesh aims to help the country meet its own objectives for development, which include help for the poorest people, the balanced use of natural resources, and improved energy supplies and communications. The health and rapid growth of the population are areas of increasing concern, as are education and the involvement of women in development. Most women in Bangladesh, burdened by the multiple pressures of community and family duties and work outside the home, need education and opportunities that are tailored to their specific needs.

A growing concern: health and population Bangladesh's population is growing at a rate of 2.3 per cent a year, and the mortality rate for mothers, newborn babies and infants is also very high. These statistics underline the close relationship between health and population. In an overcrowded city or village complex, where there is little or no clean water or sanitation, conditions such as diarrhoea can spread rapidly. And if family planning and other health services are inadequate, fewer men and women will have the choice to limit their families, and populations may continue to grow.

The Government of Bangladesh is tackling these issues with several programmes. The ODA is providing over .20 million for the Fourth Health and Population Project, which concentrates on: improving family planning and mother and child health care by supporting the organisations involved in those areas—medical and nursing colleges, NGOs, and the Ministry of Health's Management Development Unit; backing the development of a Human Resource Development Master Plan for the health sector; improving efficiency by promoting better economic and financial analysis of the health budget. The ODA also funds primary health care through larger national and international NGOs, such as the Bangladesh Rural Advancement Committee (BRAC) and Care.

APPENDIX 3

TEXT 8

Britain is responding generously to calls for assistance to combat the severe drought in parts of southern Africa. Working with governments, UN agencies, non-governmental organisations (NGOs) and local partners, the Overseas Development Administration has undertaken a range of measures across the region to help vulnerable groups, some of whom may not yet have recovered fully from the effects of the 1992 drought.

Erratic rainfall and water shortages have caused failed harvests, chronic food shortages and the spread of disease across the southern African region. In some areas the rural credit system has collapsed.

In Malawi, the UK is providing £4.1 million to support a programme providing free seed and fertiliser to around 650 000 farming households in drought affected areas. Given reasonable rains it is hoped that an extra 160 000 tonnes of maize can be harvested as a result. In addition, Britain is supplying 10 000 tonnes of relief maize to Malawi through the World Food Programme and NGOs.

In Zambia, Britain is providing nearly £4 million of maize for immediate distribution through the World Food Programme. In addition, grants to NGOs of over £1.3 million will fund supplementary feeding, medical treatment, agriculture and water resource recovery, and logistical support for drought relief and recovery activities.

In Zimbabwe, where an estimated five million people are affected, Britain has contributed over £5.3 million in aid, mainly for feeding programmes for children and other vulnerable groups. The ODA is also supporting water supply projects, including the drilling and deepening of wells.

In Lesotho, Britain has donated cereals valued at over £1 million. In addition, grants to NGOs are funding food distribution, relief and rehabilitation programmes for vulnerable households and food-for-work projects.

Mozambique has received 7 500 tonnes of food aid to help vulnerable groups in areas affected by the drought, particularly in the central provinces. This assistance is being provided through NGOs.