

## Readability in ESL<sup>1</sup>

Patricia L. Carrell

Department of Linguistics, Southern Illinois University, Carbondale,  
USA

This article reviews the literature critical of readability formulas from the perspective of their use in second language reading contexts. Relevant empirical research (Davison & Kantor 1982; Johnson 1981; Blau 1982; Floyd and Carrell 1987) which casts doubt on the efficacy of syntactic simplification/adaptation is also reviewed. The paper argues against using readability formulas not only as guides to text production or adaptation/simplification, but also as measures of the difficulty of naturally occurring texts. The paper argues that valid measures of a text's comprehensibility require consideration of textual phenomena at the level of discourse, of syntactic and lexical choices other than those which affect length, of logical/rhetorical ordering of ideas and progression of topics and comments, as well as — most importantly — background knowledge presumed of the reader.

### INTRODUCTION

ESL reading teachers and materials developers need reliable ways of matching the difficulty of second language reading materials to second language readers. If materials are too easy, students are unchallenged and bored, and no learning occurs; if materials are too difficult, students are frustrated and withdrawn, and again no learning occurs. Optimal learning occurs when the difficulty levels of instructional reading materials are appropriately matched to the capabilities of the readers. The problem lies in how to achieve this ideal.

To those who have looked to, or are contemplating looking to, readability formulas and other text-based criteria (e.g. T-unit measures and other syntactic and/or lexical criteria) as a way of measuring the difficulty or comprehensibility of reading materials, the purpose of this article is to sound a cautionary note and to suggest other factors which can and should be considered. Readability in the narrow sense of readability formulas will be considered first, followed by consideration of readability in a broader sense. The paper will conclude with some recommendations of things that ESL reading teachers and materials developers can do to better match the reading difficulty of materials to the abilities of ESL readers.

### READABILITY FORMULAS IN ENGLISH AS A MOTHER TONGUE

Readability formulas, which have been used more widely in first language

---

Patricia L. CARRELL is Professor of Linguistics and EFL/ESL at Southern Illinois University at Carbondale. Her research interests include the role of background knowledge and the interaction of text and reader in text comprehension. She has presented numerous conference papers on this topic and her research in this area has appeared in the *TESOL Quarterly*, *Language Learning*, *Language Learning and Communication* and *On TESOL '82*. She may be contacted at the Department of Linguistics.

<sup>1</sup>Paper presented at the Twenty-first Annual TESOL Convention, Miami, Florida, April 22—25, 1987.

readability than second, and especially in the United States of America, are concerned with ways of measuring the ease of understanding a text, owing to the style of writing or the specific language in that text. That is, their focus is on the *text* as an artifact in itself. Early work on readability formulas began between 1915 and 1920 in the search for objective methods to supplement subjective judgments of individual reading abilities, specifically with the advent of standardised reading tests. Thorndike's *The Teacher's Word Book* (1921) was the first milestone; his tabulations of the frequency of 10,000 printed words in sample texts set the stage for readability formulas. In fact, Thorndike encouraged such use by providing guidelines for grade levels where particular words could be appropriately taught. This led Lively and Pressey (1923) to develop the first actual "readability formula," which provided several indices of the difficulty of a text, largely in terms of Thorndike's word frequency values. Lively and Pressey's criterion of validity—that is, the independent measure against which to judge the validity of the formula—consisted only of their judgments of the relative difficulty of sixteen different pieces of reading material. Subsequent formulas have used more objective criteria to validate themselves, for example, correlating readability with standardised reading test scores, with large groups of readers' and/or teachers' judgments.

Vogel and Washburne's 1928 formula took the form of a regression equation involving more than one language variable, and thus became the prototype for most formulas since. Gray and Leary (1935) tried to take a more comprehensive look at the many possible contributors to readability. A sample of librarians, publishers, and teachers and directors of adult education classes contributed 288 suggested "factors," classified under four headings: (1) content factors, (2) style factors, (3) format factors, and (4) general factors of organisation. Nonetheless, Gray and Leary considered only style factors for the readability formulas they subsequently developed. Formula developers since have almost universally followed the same pattern of omitting content, format, and organisation variables.

The increased need for readers to cope with scientific and technical materials during the second world war and afterward served to bolster the use of readability formulas in the USA not only in educational settings, but also in the military, in government, in business and industry, and in law (cf. Fry 1987). By 1973 various formula developers had tried well over 200 different language/style variables, and had in fact developed almost as many different formulas. Computers have helped to make possible more complex analyses, combinations, and comparisons of variables than in the past. However, due to considerations of ease of application, the two most common factors in the formulas continue to be word length/frequency/familiarity on the one hand, and sentence length on the other. According to Zipf's law (1935), word length is inversely related to word frequency, and, dating back to Thorndike's 1921 word frequency thesis, infrequent words are

deemed to be less familiar and thus more difficult than more frequent words. Sentence length is similarly correlated with more difficult texts: the longer the sentences, the more difficult the text. Four of the most common and widely used readability formulas in English as a first language follow; the reader will note that they all rely on measures of word and sentence length.

Fig. 1 Sample Readability Formulas (from Klare 1984)

*Flesch Reading Ease Formula*

$$RE = .39 (\text{words/sentence}) + 11.8 (\text{syllables/word}) - 15.59$$

(Simplified version)

$$RE = .4 (\text{words/sentence}) + 12 (\text{syllables/word}) - 16$$

*Gunning*

$$\text{Readability index} = .4 (\text{mean sentence length} + \% \text{ words over two syllables})$$

*Fog Count*

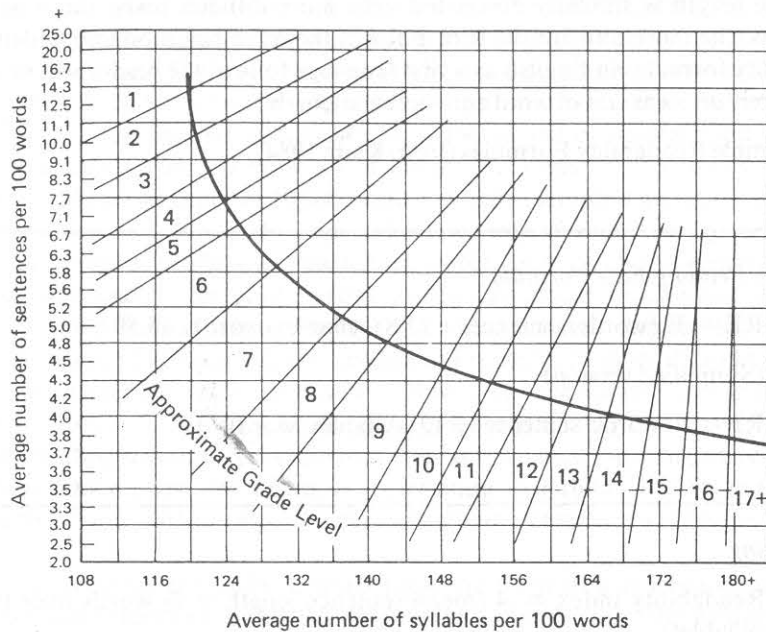
$$GL = \frac{\left[ \frac{\text{easy words} + 3(\text{hard words})}{\text{sentences}} - 3 \right]}{2}$$

GL = grade level; easy words = 1 or 2 syllables; hard words = more than 2 syllables

*Fry*

Grade level = intersection of value for sentence length and word length measured in syllables on the Fry reading graph (factors weighted differently for different grades).

Fig. 2 Fry's Extended Readability Graph



In English as a first or native language, the formulas do work to a certain extent. That is, they achieve high correlations (as high as the low .90s) with their criteria (that is, with other measures of readability, e.g. reading test scores and teacher's judgments) and with each other. However, they work because they are statistical formulas developed over and intended for use over large samples of texts. They do not work well with small samples. Consider the following sentences:

- 1) The uneven numbers are one, three, five, seven, nine, eleven and thirteen.
- 2) The squares of the absolute values of the transition amplitudes are summed.

Each of the two sentences is equally "readable" according to a readability formula which counts sentence length and word length in syllables; each sentence consists of 1 sentence, 12 words, and 19 syllables. That the first sentence contains a simple subject, a copula verb and list of eight familiar numbers, while the second consists of an agentless passive, a complex subject containing multiply-embedded prepositional phrases, and quasi-technical and technical vocabulary unfamiliar to most non-scientists is not detected by the formulas. While admitting that any readability analysis must be applied to larger chunks of connected discourse, not just to individual sentences as in these examples, nonetheless such larger chunks of

text could differ as substantially as these two sentences and still yield the same readability index.

Criticisms of readability formulas and their estimates and grade-level equivalents have abounded for about as long as there have been formulas. These have included:

a) **Statistical Consideration** A major reason for the high correlations of modern formulas lies in the nature of the criterion against which they are validated. If the criterion consists of a large number of test passages of varying content and with a wide range of difficulty and if a wide range of readers is added, then correlations are increased. However, the predictability of the formulas – that is, the high correlations – drops dramatically when more limited ranges of reader abilities, text subjects and numbers of test passages are considered. This is not always borne in mind by those who use readability formulas, often using them for limited ranges of reader ability, and limited ranges of text contents.

b) **Reliability of the Criterion of Validity** For many years, the McCall-Crabbs *Standard Test Lessons in Reading* were the common criterion used to validate readability formulas. This set of standardised test lessons vary widely in difficulty and content; comprehension is measured by objective questions. Recently, however, use of this criterion has been seriously questioned – and by McCall himself who is reported to have been surprised to learn that the lessons had been used in formula development. Stevens (1980) reports that McCall had considered them to be only practice exercises, with their grade level scores only rough equivalents for students to track their own progress.

The cloze procedure has recently become the main source of criterion passages. However, the cloze procedure can be, and often is, misused as a criterion. The most common abuse is to use only one of the  $n$  forms of a fixed-ratio, every  $n$ th deletion format, to collect criterion data. Studies have shown that all  $n$  forms of an every  $n$ th fixed-ratio deletion cloze are seldom equal in difficulty. Another criticism of cloze is that it may be more a measure of statistical redundancy rather than comprehensibility (Kintsch and Vipond 1979). Both correlations and cloze procedures have been characterised as "rubber yardsticks" in this regard (Carver 1977–1978: 31).

With the early criterion in question, and with the more recent cloze-based criterion also in question, validity of these two common criteria is a serious question. (Kintsch and Vipond 1979: 13, have suggested use of immediate free-recall and summarisation tasks, as well as reading time, as preferable criterion measures to multiple-choice or cloze measures.)

c) **Grade Levels** Providing readability estimates in terms of grade levels appeared early and has been embraced by most formula developers because these values

could be compared so conveniently to school grades and standardised reading test scores. But different readability formulas can yield different grade-level ratings, some by as much as three years difference. For this reason, the practice of relating readability estimates with grade levels has recently been condemned by the International Reading Association (*Reading Research Quarterly*, 1981). Manzo (1979) has also condemned the comparison of formula estimates with those derived from basal readers as a means of providing grade-equivalents as "incestuous" and has said that this made readability research "a construct without a point of reference."

d) **Sampling Variability** Another problem concerns the number and length of the text samples selected for analysis. The traditional practice of analysing selected 100-word samples from within a text rather than a complete text may obscure a great deal of variability within a text. How many such samples are necessary? The current practice (encouraged by Fry 1977) to use only three such 100-word samples makes sampling variability an especially important concern. Coke and Rothkopf (1970), for example, found a range of Flesch Reading Ease values from approximately 35 to 120 in a single book (roughly from below fourth-grade level to college level).

e) **Declining Reading Skills** The almost obsessive concern with readability in L1 classrooms in the USA during the past few years is sometimes seen as a cause of recent declines in inferential reading skills. Chall (1979), a co-developer with Dale of the highly regarded Dale-Chall readability formula, has found evidence for an association between the decline in textbook difficulty over a thirty-year period and the then-current decline in Scholastic Aptitude Test scores (Chall, Conrad and Harris 1977). She also described publishers as saying that elementary school teachers were requesting science and social studies textbooks two years below the grade placement of their children. This is seen as a vicious cycle: lower readability levels of classroom materials leads to lower reading abilities of students, which in turn leads to further lowering of readability levels of classroom materials, and so on.

f) **Factors Omitted from the Formulas** As previously mentioned, there are many text-based factors which are omitted from readability formulas – not to mention reader-based variables which are totally ignored by formulas. Most obvious of the text-based factors are those having to do with syntactic complexity, rhetorical organisation, and propositional density. Hunt's (1965) T-unit measures are attempts to reflect syntactic complexity of a very general sort (simple independent clauses, conjoined clauses, subordinate clauses). Several studies have been conducted by now to show that rhetorical organisation of a text affects comprehension (Mandler 1978; Meyer 1977a and 1977b; Meyer, Brandt & Bluth

1980; Meyer and Freedle 1984; Kintsch and Yarbrough 1982). Several other studies have been conducted to show that what Kintsch calls "propositional density" also affects comprehension (Kintsch and Keenan 1973).

g) **Comprehension** The most basic problem with readability formulas is that they assume we know what comprehension is, that it is a unified phenomenon susceptible to measurement via text/style variables. "Comprehension" is a complex term which not only means different things to different people, but it is a complex concept which covers multiple behavioral and cognitive factors.

## MOTHER TONGUES OTHER THAN ENGLISH

There have been attempts to extend readability formulas to languages other than English. Most of the languages have been European languages whose syntax and word structure are not very different from English, and most of the formulas have been based on those developed for English. For example, three recent formulas have been developed for Spanish, but they are all based upon the Fry (1977) graph. One of these formulas (Garcia 1977) adapted the Fry graph by changing both the vertical and the horizontal numbers to reflect language differences in sentence length and syllable length, respectively. Another one (Gilliam, Pena, and Mountain 1980) changed only syllable length, arguing that a change in sentence length was not necessary. Non-European languages such as Korean and Vietnamese have also recently been the object of readability formulas (Park 1974, Nguyen and Henkin 1982). In the case of Korean, the formula contains variables for "easy words," "different words," "different hard words," "simple sentences," and "pronouns." In the case of Vietnamese, the formula contains variables for word length in letters, sentence length in words, and percentage of "difficult words." As in English, these formulas were developed using already existing graded textbooks and experienced teachers' ratings as the criterion data.

Quite different problems face language communities which have no educational tradition for the languages in question, nor currently existing stocks of texts used for teaching reading. For example, the Yupik community of Alaska is trying to preserve their language (along with English) by teaching their children to learn to read with Yupik as the medium. Needless to say, the sentence and word structure of this language are very different from English. Without intervention, the language will soon be lost as children learn only English. In this situation, it would not be wise to teach reading by using text materials which are too hard for the children, or which are too simple and not appropriate for older children. However, instead of trying to adapt English-based readability formulas to Yupik, the members of the community have tried to draw on their own knowledge and experience as speakers of Yupik. They are analysing how people give information, tell stories, explain procedures to children, etc. and are then transferring these oral texts to the written

medium and trying them out on groups of children. In this way, texts in Yupik can be created for different age groups, though not necessarily graded into very fine grade-level distinctions. In this situation, rather than adapting or developing readability formulas, attention is being paid to actual readers, texts and a wide variety of features of language. (Information on the Yupik situation from Davison 1985.)

### READABILITY FORMULAS IN SECOND LANGUAGE/FOREIGN LANGUAGE SETTINGS

Readability in second or foreign language settings is very much an issue. Readability, which is ultimately a question of comprehension, is related to the broader question in our discipline of the role of comprehensible input in language development and what constitutes comprehensible input (Krashen 1980; Long 1981, 1985; Chaudron 1983; Pica 1986; Gass and Varonis 1985; Kelch 1985). This discussion will be limited to readability and written texts. As in the first or native language literature, the focus in second/foreign language readability has often been exclusively on features of the text. Standard foreign language reading materials, and EFL/ESL materials have often been "graded" for difficulty or readability; graded readers are a common phenomenon. (In German language teaching, for example, Van Nostrand publishes a series of Cultural Graded Readers; Gerard F. Schmidt has published *Spass und Spannung: A beginning graded German reader*, New York: Macmillan 1979.) The grading has often been done by intuitive means, i.e. the authors or editors judging reading difficulty based on their experience, or by more structured means. Although readability formulas *per se* have been used relatively infrequently, other text-based, formal criteria – specifically syntactic and lexical criteria – have frequently been applied. For example, the Regent's series of Graded Readers has been based on the Longman's syntactic structural criteria, and on the *Cambridge English Lexicon*, which is itself ultimately based on word frequency studies. (Nuttall 1982, Appendix C, contains a list of British publishers of graded readers for students of English as a second or foreign language.)

### MISAPPLICATIONS OF READABILITY FORMULAS OR OTHER FORMAL CRITERIA

**Text Simplification – Writing to Formula or Simplifying to Formula** One of the worst abuses of readability formulas is using them to produce a text at a pre-determined readability level, or to modify an existing text in an effort to simplify or lower the readability level of that text. As Klare (1984) points out, using readability formulas to answer the question "How can I tell how readable a piece of writing is likely to be for intended readers?" (the so-called prediction question) is quite different from using readability formulas to answer the question "How can I write readably (or more readably) for intended readers?" (the so-called production

question). Readability formulas were designed for prediction on already-existing texts, not production, and their use in production has been far less successful than their use in prediction. Klare (1985) has criticised using a readability formula in an attempt to produce readable writing – or as he calls it "writing to formula" – as being a bit like "expecting to raise the effective temperature in a room by holding a match under the thermometer" (1985: 244). The following discussion will focus specifically on the use of readability formulas or other structural criteria to modify existing texts in an effort to simplify or lower the readability of that text. (This seems to correspond to Widdowson's (1978) notion of "simplified versions" versus "simple accounts.")

Davison and Kantor (1982) compared two versions of four texts, the original versions intended mainly for adult-native English readers, and the adapted versions intended for less skilled readers. In discussing specific changes made in apparent attempts to make the texts conform to certain levels of readability as defined by readability formulas – i.e. to make the texts easier to read – they show that many of the changes actually function to make the adapted texts harder to understand.

A common change made in adaptation is to shorten overall sentence length (and thus lower the readability level) by splitting complex sentences into independent sentences. The following is an example from Davison and Kantor:

3) (Original) "If given a chance before another fire comes, the tree will heal its own wounds by growing new bark over the burned part."

(Adaptation) "If given a chance before another fire comes, the tree will heal its own wounds. It will grow new bark over the burned part." (1982: 192)

The subordinate gerund clause "by growing..." in the original text is turned into a separate sentence. But, as Davison and Kantor point out, removing the "by-ing" gerund marker of subordination results in a loss of relevant meaning. Whereas the original expresses explicitly that "growing new bark" is the means by which the tree heals its wounds, the adaptation does not make that meaning relationship clear. The reader must now infer this information. For a reader who has some background knowledge about trees such an inference may not be problematic; however, as Davison and Kantor point out, it would be easy for a reader without such knowledge to miss this connection and to infer only a sequential relation between the sentences. (Not incidentally, Pearson (1974-75) found that children were able to understand causal relationships better in sentences like "Because the chain broke, the machine stopped" than in two separate simple sentences like "The chain broke. The machine stopped." He concluded that simple sentences may actually extract some of the explicitness from intersentential propositional relationships, and consequently be more difficult to comprehend.) In other words,

syntactic simplicity may not be an aid, but a hindrance to comprehension since simplified syntax may decrease explicit textual cohesion.

The next example is also from Davison and Kantor:

4) (Original) "This small sand-dwelling animal emerges at night and secretes a luminous mucus as it moves about."

(Adaptation) "This small animal, which lives in the sand, comes out at night. As it moves about, it secretes a luminous substance." (1982: 198)

In this adaptation, breaking the original conjoined sentence into two separate sentences results in two focuses which receive equal emphasis: "comes out at night" and "secretes a luminous substance." The effect communicated is not the same, and not as coherent, as in the longer original, which subordinated a lot of background information about the luminous worms (e.g. their coming out at night) to the main point of the luminous trails. Thus, as shown by Davison and Kantor, one of the negative effects of splitting sentences in order to lower readability levels is to foreground information which should be backgrounded.

Davison and Kantor also show that many successful adaptations often violate readability formulas, and actually make sentences in the resulting texts longer, not shorter. (For example, an original text read as follows: "We had water to drink after that. We set out basins and caught the raindrops." The adapted version read: "We set out basins to catch the raindrops so that we would have water to drink.") They also point out that sensitive adaptations often result in a Domino Effect, that a change in one part of a text (e.g. shortening sentence length) requires other changes elsewhere in the text in order to compensate for the first change. They show that successful adaptations pay attention to factors such as text structure and content, not just sentence and word length.

## L2 STUDIES OF 'SIMPLIFIED' SYNTAX

At least five different studies have explored the effects on ESL reading comprehension of simplifying the difficulty of the language of the text. Johnson (1981) investigated the effects of both the complexity of the language and the cultural origin of prose on the reading comprehension of Iranian ESL students in the USA. Subjects read unadapted or adapted English texts of two cultural origins, as Iranian and an American folktale. Adaptation was based on a contrastive analysis of Farsi and English, and consisted of simplifying such English syntactic structures as relative clauses, sentential complements, compound and complex sentences, low-frequency vocabulary, and figurative language. Johnson's results revealed that the level of syntactic and semantic complexity of the text had a lesser effect on the reading comprehension of the ESL readers (as measured by recall

protocols, and multiple-choice questions on explicit and implicit information in the texts) than the cultural origin of the texts. Only the culturally unfamiliar text was easier for the ESL readers in adapted form.

Blau's 1982 study directly challenged the usual sentence length criterion of readability formulas which deem short sentences easy to read. Two groups of Puerto Rican subjects – college students and eighth graders – read three versions of a text. Vocabulary and content were held constant, while sentence structure varied across three versions: version 1 consisted of short, simple sentences; version 2 of complex sentences with clues to underlying relationships left intact; and version 3 of complex sentences without such clues. Contrary to what readability formulas would predict, version 2 rather than version 1 yielded the highest comprehension scores (on multiple-choice comprehension questions). Blau concluded that not only does lower readability level material, as measured by common readability formulas, not facilitate comprehension for these ESL students, but that it may actually impede comprehension.

Strother and Ulijn (1987) investigated the relationship between syntactic rewriting (simplification) and reading in English for Science and Technology (EST). They simplified syntactic structures commonly encountered in EST writing – namely passives, nominalisations, and participles. In a study of both expert and novice EST-ESL readers (Dutch and Chinese experts, Spanish and Arabic novices) on measures of reading time and correct answers to comprehension questions, they found no effect for simplified syntax – for either the novices or the experts. They conclude that what is needed is not *syntactic* rewriting of professional texts, but *lexical* rewriting to increase readability. They argue that concentrating on lexical skills and lexical rewriting may increase learning from texts and text readability, respectively.

In a pedagogical study addressing the question "Can we improve ESL students' reading by helping them build background knowledge on the topic prior to reading?", Floyd and Carrell (1987) used pre- and post-tests with a heterogeneous group of intermediate level ESL students enrolled in an intensive pre-university program in the USA. Half of the subjects were provided with the opportunity to build appropriate prior background knowledge on text content; the other half were not. Half of each group read a syntactically more complex version, half read a syntactically less complex version, of semantically comparable test passages. (The syntactically more complex passage had T-units of average length 11.34 words; the syntactically less complex passage had T-units of average length 8.67 words.) Results showed that by providing students with first-hand experiential knowledge, reading comprehension (as measured by objective tests as well as a free written recall test) was facilitated; differences in the syntactic complexity of the reading passages showed no significant effect on ESL reading comprehension.

Brown (1985) compared ESL readers on three versions of a text: (1) native, authentic, unmodified text, (2) the same text modified for syntax and vocabulary – which he referred to as the “modified input” text in the sense of Long (1981); and (3) the same text modified for structure, elaboration, clarification, and explanation – which he referred to as the “modified interaction” text (Long 1981). Although the modified interaction text was at a higher grade level of readability (9th), it was comprehended by 9th through 11th grade ESL students in Taiwan equally as well as the modified input text which was at a much lower readability level (5th grade). These results parallel Long’s (1981) finding for listening comprehension, namely that modified interaction may facilitate comprehension more so than modified input.

### READABILITY IN THE BROAD SENSE

As the foregoing studies demonstrate, the real problems with readability formulas lie not only in the textual features they ignore (syntactic complexity, textual cohesion, propositional density, rhetorical structure) but in the fact that they ignore the reader and reader variables such as background knowledge. They ignore the interactive nature of the reading process – the interaction of the reader with the text. They ignore the “fit” between the various features of a text, including its content domain, its rhetorical organisation, its perspective on the topic on the one hand, and on the other hand, the various schemata of the reader – the reader’s schemata for content and rhetorical organisation, the reader’s background knowledge, experiences, beliefs. This can be devastating in second or foreign language reading, particularly if the reader and text come from different cultural backgrounds. By now the second language reading literature contains many studies showing the crucial nature of the interaction between a text and a reader’s content schemata (Gatbonton and Tucker 1971; Steffensen, Joag-dev and Anderson 1979; Johnson 1981; Carrell 1981, 1987a; Alderson and Urquhart 1985; Floyd and Carrell 1987). Several other studies have shown the crucial nature of the interaction between a text’s rhetorical organisation and the reader’s formal schemata (Urquhart 1984; Carrell 1984a, 1984b, 1987a; Hinds 1983a, 1983b).

In addition, I have argued elsewhere (Carrell 1987b) that a reader’s background knowledge includes not only content, formal and linguistic schemata, but that several other, more specific factors may play a role in a reader’s background knowledge of, or schemata for, subject matter, genre, sociocultural and general world knowledge, in addition to linguistic knowledge of text code. These factors encompass de Beaugrande and Dressler’s (1981) so-called “standards of textuality,” – what makes a text a unified, meaningful whole, rather than just a string of unrelated sentences, what makes a text a communicative interaction. These include not only *cohesion* and *coherence*, but *intentionality*, *acceptability*, *informativity*,

*situationality*, and *intertextuality*. *Intentionality* is the text producer’s attitude that a cohesive, coherent text is created for some goal. *Acceptability* is the corresponding attitude from the text receiver’s perspective. *Informativity* is an index of the extent to which text occurrences are or are not probable or predictable in their context. *Situationality* is a measure of the text’s relevance to a current or recoverable situation, and whether the text merely monitors a situation or attempts to manage that situation. *Intertextuality* is the principle whereby the production or comprehension of a given text depends on knowledge of other specific texts or text-types.

These principles of textuality help characterize what it is that makes the following texts effective and interesting.

5) “Call us before you dig. You may not be able to afterwards.” (Telephone Company) (de Beaugrande and Dressler, 1981: 8).

6) “Twenty-year-old Willie B. is a diehard TV addict. He hates news and talk shows, but he loves football and gets so excited over food commercials that he sometimes charges at the set, waving a fist. Says a friend ‘He’s like a little child.’

Willie B. is a 450lb gorilla at the Atlanta Zoo. In December a Tennessee TV dealer heard about Willie B.’s lonely life as the zoo’s only gorilla and gave him a TV set.” (TIME Jan 22, 1979) (de Beaugrande and Dressler 1981: 151).

The text in (5) shows that texts which require the receiver to make important inferences (provided the reader/receiver is capable of those inferences) are often more effective than versions which make everything explicit. The text in (6) shows that texts which manipulate the level of informativity – first leading the reader to assume that the text is about a human being, and then leading the reader to have to change that assumption and to reprocess the text – markedly increase the impact of a text on a reader. (Note how dull the text would have been had the two paragraphs been reversed.)

De Beaugrande and Dressler’s (1981) standards of textuality not only remind us of the important role of the reader in determining what constitutes a text in the first place, a perspective obviously missing from readability formulas, but they also suggest that readability cannot be optimised simply by striving for the best possible match between text-presented knowledge and prior world knowledge of the reader. The resulting text would possess too low a degree of informativity and hence be devoid of interest. As de Beaugrande and Dressler themselves say: “Readability must *not* be defined as the expenditure of the least effort (despite Hirsch 1977), but rather as the appropriate proportion between required effort and resulting insights” (1981: 213).

Silberstein (1987), in commenting on Carrell (1987) vis-a-vis de Beaugrande and Dressler's (1981) standards of textuality, has called attention to an approach to literary criticism which is also exploring the mysteries of textual interaction – namely reader response theory. This approach to literary criticism developed as a challenge to the traditional notion of a literary text as an autonomous aesthetic object. Scholars affiliated with this school of criticism converge on the following assumptions: that meaning is not contained in the text, but is derived from an interaction between the content and structure of the author's message and the experience and prior knowledge of the reader; that readers comprehend differently because every reader is culturally and individually unique; and that examining readers' responses to text is more valid than establishing one "correct" interpretation of text meaning. First-language reading teachers are already at work exploring ways to translate reader response theory into sound instructional practice (Chase and Hynd 1987).

Yet, in addition to the individuality of interpretation emphasised by reader response theory, Fish (1980), one of its central proponents, points out that members of a particular "interpretive community" often show a great deal of consensus regarding text interpretation, just as the members of a residential community might all live in the same general type of house. Thus, in addition to individual interpretations, and personal factors, we also have what are called "discourse communities." The notion of discourse communities is very much a current concern in reading English for Scientific/Technical purposes. As Widdowson has observed (1979), different disciplines, such as physics, constitute sub-cultures of their own. Texts and modes of communicating via texts in each discipline-subculture vary. Much current research in EST is explicitly centered on differences, as well as transferability, among different discourse communities (Swales 1986; Johns 1987; Carrell 1986; Crandall 1986).

## RECOMMENDATIONS

What is needed is a clearer theoretical approach to readability, one which takes a broader range of reader as well as text variables into consideration. As Kintsch and Vipond have said, the problem of predicting reader comprehension "is not with the formula but with our theories" (1979: 335); they believe that readability must be defined for specific texts and specific readers. Valid measures of a text's comprehensibility, particularly to a second or foreign language reader, require consideration of a number of factors. Those with responsibility for selecting reading materials which will appropriately match ESL readers' abilities should consider the following:

1. Consider whether the reading is to be done extensively, outside the classroom and without the support of a teacher, or intensively, inside the classroom and with

the support of a teacher. Obviously, more difficult readings can be tackled inside the classroom where the teacher is there to guide and support the reading. As Klare has observed: "Training readers calls for more challenging material than merely informing or, especially, entertaining them" (1984: 730).

2. Take into account other recognised contributors to comprehension, especially higher levels of interest, motivation or prior knowledge. Particularly interesting texts and appropriate levels of prior knowledge can stimulate motivation and keep challenging material from being frustrating.

3. Bear in mind that readability formula scores, deriving from counts of style difficulty, become poorer predictors of difficulty for older, more mature readers, where content weighs more heavily.

4. Factor into readability judgments, consideration of textual phenomena at the level of discourse, for example cohesiveness, coherence, the flow of topics and comments, and propositional density – factors which have been shown to play a part in comprehensibility.

5. Factor into readability judgments, consideration of syntactic and lexical choices other than those which affect length – e.g. quasi-technical vocabulary, or non-technical vocabulary used in technical ways.

6. Consider the rhetorical structure of texts (Carrell 1985), and the transparency of the signalling of that rhetorical structure (Meyer, Brandt, and Bluth 1980), both of which have been shown to affect reading comprehension.

7. Pay special attention to any specialised background knowledge which the text may presume of the reader, and whether the reader has such background knowledge, or can and will be helped in some way (e.g. through pre-reading activities) to relate existing background knowledge to that presumed by or presented in the text.

8. Consider the trade-off between effort to be expended and value received by the reader in terms of information, entertainment, or whatever the purpose for reading.

## REFERENCES

- Alderson, J.C. and A.H. Urquhart (1985) This test is unfair: I'm not an economist. In P. Hauptman, R. LeBlanc, and M.B. Wesche (Eds.) *Second Language Performance Testing* Ottawa: University of Ottawa Press.
- Blau, E.K. (1982) The effect of syntax on readability for ESL students in Puerto Rico *TESOL Quarterly*, 16, 517–528.



- Brown, R.L. (1985) *A comparison of the comprehensibility of modified and unmodified ESL reading materials*. Unpublished Master's thesis, University of Hawaii at Manoa.
- Carrell, P.L. (1981) Culture-specific schemata in L2 comprehension. In R.A. Orem and J.F. Haskell (Eds.) *Selected Papers from the Ninth Illinois TESOL/BE Annual Convention and the First Midwest TESOL Conference*. Chicago: Illinois TESOL/BE.
- Carrell, P.L. (1984a) Evidence of a formal schema in second language comprehension. *Language Learning*, 34, 87—112.
- Carrell, P.L. (1984b) The effects of rhetorical organisation on ESL readers. *TESOL Quarterly*, 18, 441—469
- Carrell, P.L. (1985) Facilitating ESL reading comprehension by teaching text structure. *TESOL Quarterly*, 19, 727—752
- Carrell, P.L. (1986, December) *Implications and future directions of research on second language reading*. Paper presented at the Ninth Annual Meeting of the American Association of Applied Linguistics, New York, NY.
- Carrell, P.L. (1987a) The interaction of content and formal schemata in ESL reading. *TESOL Quarterly*, 21 (forthcoming)
- Carrell, P.L. (1987b) A view of written text as communicative interaction: Implications for reading in a second language. In J. Devine, P.L. Carrell, and D. Eskey (Eds.) *Research in Reading in English as a Second Language*. Washington, DC: TESOL
- Carver, R.P. (1977—1978) Toward a theory of reading comprehension and reading. *Reading Research Quarterly*, 13, 8—63
- Chall, J.S. (1979) Readability: in search of improvement. *Publishers Weekly*, 40—41.
- Chall, J.S., S. Conrad and S. Harris (1977) *An Analysis of Textbooks in Relation to Declining S.A.T. Scores*. Princeton, NJ: College Entrance Examination Board.
- Chase, N.D. and C.R. Hynd (1978) Reader response: An alternative way to teach students to think about text. *Journal of Reading*, 30, 530—540.
- Chaudron, C. (1983) Simplification of input: Topic reinstatements and their effects on L2 learners' recognition and recall. *TESOL Quarterly*, 17, 437—458.
- Coke, E.U. and E.L. Rothkopf (1970) Note on a simple algorithm for a computer-produced reading ease score. *Journal of Applied Psychology*, 54, 208—210.

- Crandall, J.A. (1986, December) *Linguistic problems in math: A further look at math register*. Paper presented at the Ninth Annual Meeting of the American Association for Applied Linguistics, New York, NY.
- Davison, A. (1985) *Readability - The Situation Today*. (Technical Report No. 359.) Champaign, IL: University of Illinois, Center for the Study of Reading.
- Davison, A. and R.N. Kantor (1982) On the failure of readability formulas to define readable texts: A case study from adaptations. *Reading Research Quarterly*, 17, 187—209.
- de Beaugrande, R. and W. Dressler (1981) *Introduction to Text Linguistics*. London: Longman.
- Fish, S. (1980). *Is there a Text in this Class?* Cambridge, MA: Harvard University Press.
- Floyd, P. and P.L. Carrell (1987) Effects on ESL reading of teaching cultural content schemata. *Language Learning*, 37 (forthcoming)
- Fry, E.B. (1968) A readability formula that saves time. *Journal of Reading*, 11, 513—16, 575—578.
- Fry, E.B. (1969) The readability graph validated at primary levels. *The Reading Teacher*, 22, 534—538.
- Fry, E.B. (1977) Fry's readability graph: Clarifications, validity, and extension to level 17. *Journal of Reading*, 21, 242—252.
- Fry, E.B. (1987) The varied uses of readability measurement today. *Journal of Reading*, 30, 338—343.
- Garcia, W.F. (1977) *Assessing Readability for Spanish as a Second Language: The Fry Graph and Cloze Procedure*. Unpublished doctoral dissertation, Columbia University Teachers College.
- Gass, S.M. and E.M. Varonis (1985) Task variation and non-native/non-native negotiation of meaning. In S.M. Gass and C.B. Madden (Eds.) *Input in Second Language Acquisition* Rowley, MA: Newbury House Publishers.
- Gatbonton, E. and G.R. Tucker (1971) Cultural orientation and the study of foreign literature. *TESOL Quarterly*, 5, 137—143.
- Gilliam, B., S.C. Pena and C. Mountain (1980) The Fry graph applied to Spanish readability. *Reading Teacher*, 33, 426—430.
- Gray, W.S. and B.E. Leary (1935) *What Makes a Book Readable*. Chicago: University of Chicago Press.

- Hinds, J.L. (1983a) Contrastive rhetoric: Japanese and English. *Text*, 3, 183—195.
- Hinds, J.L. (1983b, March) *Retention of information using a Japanese style of presentation*. Paper presented at the 17th Annual TESOL Convention, Toronto.
- Hunt, K.W. (1965) *Grammatical Structures Written at 3 Grade Levels*. Champaign, IL: National Council of Teachers of English.
- International Reading Association (1981) *Reading Research Quarterly*.
- Johns, A.M. (1987) *The discourse communities dilemma: Identifying transferable skills for the academic milieu*. Paper presented at the Ninth Annual Meeting of the American Association of Applied Linguistics, New York, NY.
- Johnson, P. (1981) Effects on reading comprehension of language complexity and cultural background of a text. *TESOL Quarterly*, 15, 169—181.
- Kelch, K. (1985) Modified input as an aid to comprehension. *Studies in Second Language Acquisition*, 7, 81—90.
- Kintsch, W. and J.M. Keenan (1973) Reading rate and retention as a function of the number of propositions in the base structure of sentences. *Cognitive Psychology*, 5, 257—274.
- Kintsch, W. and D. Vipond (1979) Reading comprehension and readability in educational practice and psychological theory. In L.—G. Nilsson (Ed.) *Perspectives on Memory Research*. Hillsdale, NJ: Lawrence Erlbaum Associates.
- Kintsch, W. and J.C. Yarbrough (1982) Role of rhetorical structure in text comprehension. *Journal of Educational Psychology*, 74, 828—834.
- Klare, G.R. (1984) Readability. In P.D. Pearson (Ed.) *Handbook of Reading Research* New York: Longman.
- Klare, G.R. (1985) Matching reading materials to readers: The role of readability estimates in conjunction with other information about comprehensibility. In T.L. Harris and E.J. Cooper (Eds.) *Reading, Thinking, and Concept Development: Strategies for the Classroom* NY: College Entrance Examination Board.
- Krashen, S.D. (1980) The input hypothesis. In J.E. Alatis (Ed.) *Current Issues in Bilingual Education*. Washington, DC: Georgetown University Press.
- Lively, B.A. and S.L. Pressey (1923) A method for measuring the “vocabulary burden” of textbooks. *Educational Administration and Supervision*, 98, 389—398.

- Long, M.H. (1981) Input, interaction and second language acquisition. In H. Winitz (Ed.) *Native Language and Foreign Language Acquisition* (Annals of the New York Academy of Sciences, No. 379). New York: New York Academy of Sciences.
- Long, M.H. (1985) Input and second language acquisition. In S.M. Gass and C.G. Madden (Ed.) *Input in Second Language Acquisition* Rowley, MA: Newbury House Publishers.
- Mandler, J.M. (1978) A code in the node: The use of a story schema in retrieval. *Discourse Processes*, 1, 14—35.
- Manzo, A. (1979) Readability: A postscript. *Elementary English*, 47, 962—965.
- McCall, W.A. and L.M. Crabbs (1925) *Standard Test Lessons in Reading: Teacher's Manual for all Books*. New York: Bureau of Publications, Columbia University Teachers College. (Revised 1950; revised 1961)
- Meyer, B.J.F. (1977a) The structure of prose: Effects on learning and memory and implications for educational practice. In R.C. Anderson, R.J. Spiro and W.E. Montague (Eds.) *Schooling and the Acquisition of Knowledge* Hillsdale, NJ: Lawrence Erlbaum Associates.
- Meyer, B.J.F. (1977b) What is remembered from prose: A function of passage structure. In R.O. Freedle (Ed.) *Discourse Production and Comprehension* Norwood, NJ: Ablex Publishing Corporation.
- Meyer, B.J.F., D.M. Brandt and G.J. Bluth (1980) Use of top-level structure in text: Key for reading comprehension of ninth-grade students. *Reading Research Quarterly*, 16, 72—103.
- Meyer, B.J.F. and R.O. Freedle (1984) Effects of discourse type on recall. *American Educational Research Journal*, 21, 121—143.
- Nuttall, C. *Teaching Reading Skills in a Foreign Language*. London: Heinemann.
- Nguyen, L.T. and A.B. Henkin (1982) A readability formula for Vietnamese. *Journal of Reading*, 26, 243—251.
- Park, Y. (1974) *An Analysis of some Structural Variables of the Korean Language and the Development of a Readability Formula for Korean Textbooks*. Unpublished doctoral dissertation, Florida State University.
- Pearson, P.D. (1974—75) The effects of grammatical complexity on children's comprehension, recall, and conception of certain semantic relations. *Reading Research Quarterly*, 10, 155—192.

- Pica, T. (1986, December) *Negotiated interaction as an aid to second language production*. Paper presented at the Ninth Annual Meeting of the American Association of Applied Linguistics, New York.
- Silberstein, S. (1987) Comments on the Chapter by Patricia L. Carrell. In J. Devine, P.L. Carrell and D. Eskey (Eds.) *Research in Reading in English as a Second Language*. Washington, DC: TESOL.
- Steffensen, M.S., C. Joag-dev and R.C. Anderson (1979) A cross-cultural perspective on reading comprehension. *Reading Research Quarterly*, 15, 10—29.
- Stevens, K.C. (1980) Readability formulae and McCall-Crabbs *Standard Test Lessons in Reading*. *Reading Teacher*, 33, 413—415.
- Strother, J.B. and J.M. Ulijn (1987) Does syntactic rewriting affect EST (English for Science and Technology) text comprehension? In J. Devine, P.L. Carrell and D. Eskey (Eds.) *Research in Reading in English as a Second Language*. Washington, DC: TESOL.
- Swales, J. (1986, November). *The concept of discourse community and its relevance to advanced ESL*. Paper presented at the Sixth Midwest Regional TESOL Conference, Ann Arbor, Michigan.
- Thorndike, E.L. (1921) *The Teacher's Word Book*. New York: Teacher's College, Columbia University.
- Urquhart, A.H. (1984) The effect of rhetorical ordering on readability. In J.C. Alderson and A.H. Urquhart (Eds.) *Reading in a Foreign Language* London: Longman.
- Vogel, M and C. Washburne (1928) An objective method of determining grade placement of children's reading material. *Elementary School Journal*, 28, 373—381.
- Widdowson, H.G. (1978) *Teaching Language as Communication*. London: Oxford.
- Widdowson, H.G. (1979) *Explorations in Applied Linguistics*. New York: Oxford.
- Zipf, G.K. (1935) *The Psycho-biology of Language*. Boston: Houghton-Mifflin.