Readability Calculation by Microcomputer

Tom Rich

Often when preparing scripts, support materials and other materials for audiovisual programs, questions arise about the readability of the language. How do the writers know that the material is written at the right level? Often materials seem to be prepared on the basis of intuition or experience. Certainly, changes are made after piloting, but the changes are often superficial. Seldom are alterations made in the language level of the material.

Readability formulas which could help to determine the actual difficulty of the language, often seem too complex and cumbersome to calculate. Recently, however, microcomputers have been able to adapt the readability measures making them more useful to the curriculum expert and the audiovisual practioner. Microcomputers seem to eliminate the tedium from the mathematics of readability evaluations.

This article will describe a program adapted from a longer one described in an article by Michael Schuyler, "A Readability Formula Program for Use on Microcomputers," in the March, 1982 *Journal of Reading*. It had been redesigned to run on a PET microcomputer. Because of memory limitations some of the features of the original program have been eliminated. The program can be loaded from disk or tape and, if desired, a printer can be used for the final output.

Purpose

The program gives basic statistical information on the sample passage including: words, sentences, three-syllable words, syllables per 100 words and sentences per 100 words. A total of six different formulas for readability are computed and the data for plotting a Fry Readability Graph is given.

The passage to be evaluated can be any length but commonly passages of 100 words are used. In order to get a better evaluation of material from a book, it is suggested you use several passages of approximately 100 words from various sections of the book.

Methodology

A few comments must be made about how the material to be evaluated should be entered. The first thing you will see on the screen is an abbreviated set of instructions. After the instructions, the program will ask you to enter the name of the book or material you are evaluating and the number of lines of text in the first section to be entered. After typing each, press the RETURN key.

There is no need to count the exact number of words or sentences in the passage to be evaluated as the program does that for you. Simply count the number of lines, as they appear in the source material, in the first passage you are evaluating.

A question mark (?) will appear at the bottom of the screen when you are ready to

Tom Rich is the Media Coordinator for the Department of Education in Prince Edward Island.

enter the text. You will enter it one line at a time; a new line is indicated with a question mark. The lines correspond to the lines in the source material and not to the screen on the computer. Simply type the line in continuously without hitting RETURN even if the edge of the screen is reached. The line will automatically wrap around to the next line on the screen. Don't worry about words being broken up at the end of the line. The only caution is that one line from the book cannot take up more than two lines on the computer screen. If it does, you will have to break it into two parts.

Each line of text entered must end with either a slash (/) or the normal punctuation (. ? !) If the line in the source passage does not correspond to the end of a sentence then space after the last word and type a slash before hitting RETURN. Words at the end of a line cannot be broken with hyphens. If the source text has words broken at the end of a line, complete the word before adding the slash and hitting RETURN.

The actual typing of the material is a little different from normal typing. Other than punctuation at the end of the lines and sentences, absolutely no punctuation may be used including commas, colons, semi-colons, dashes, quotes or parentheses or any other symbol that is not a letter or a number. At the end of the sentence the normal punctuation is used. Note however, that you must leave a space between the last word and the ending punctuation. The following punctuation begins immediately with no space between the punctuation and the first word of the sentence. In addition, the typing is done automatically in all capitals. You do not have to use the shift key.

Here is an example. Each line from the source material begins at a question mark printed by the computer each time you hit RETURN.

> ? THIS IS SAMPLE TEXT .NOTE THE SPACE BEFORE THE PERIOD NOT AFTER/
> ? AND THE SLASH AT THE END OF THE LINE WHEN NOT A SENTENCE END .

The only time you will hit RETURN is when you have completed entering a line from the text. Before you do so, check the line for errors and make sure it ends with a slash or end punctuation as you cannot make corrections after hitting RETURN. The final line of the passage must end with a period, question mark or exclamation mark. If you must make corrections, use the DEL key at the upper right of the number pad. This will erase letters. To position the cursor to change letters, the cursor controls next to the DEL key may also be used.

When you hit RETURN, there will be a slight pause as the line is evaluated by the

program and then another question mark will appear for the second line to be entered. When you have entered the final line of text that sentence must have normal punctuation at the end — not a slash. You must also use the number of lines which you specified in the beginning of the program. The computer will keep returning with question marks until that number of lines has been entered.

When the last line has been entered the computer will automatically process the material and, after a slight pause, display the results. At that time, you will have the option of continuing with another passage from the same text or ending. If you answer YES the program will again ask how many lines of text and display the question marks to enter the copy. This must be material from the same source as the preceeding passage. If you wish to switch to evaluating different materials, you should answer NO and start the program again by typing RUN.

When you have completed the number of passages you wish to evaluate in the particular text, answer NO and the program will display the summary data table (See Figure 1). This is essentially the same as the one seen after each passage except that it notes how many passages have been evaluated and the total number of words, three syllable words, sentences and syllables for all the passages. The rest of the data is averaged across all the passages evaluated.

The statistics for the syllables per 100 words and sentences per 100 words are used in plotting a Fry Readability Graph (See Figure 2). Enter the sentence data for the
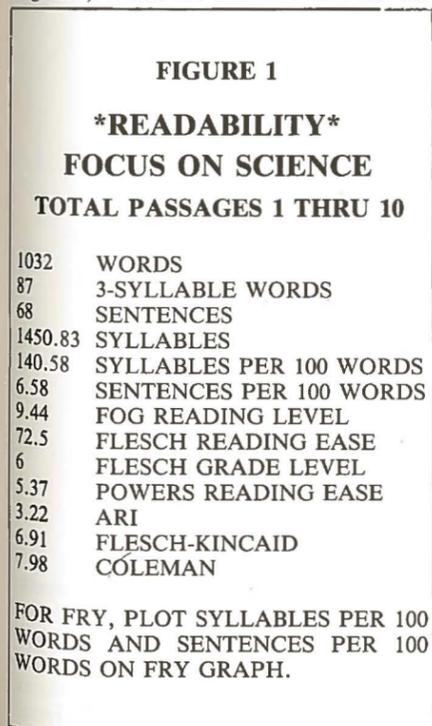
FIGURE 1

READABILITY
FOCUS ON SCIENCE
TOTAL PASSAGES 1 THRU 10

1032	WORDS
87	3-SYLLABLE WORDS
68	SENTENCES
1450.83	SYLLABLES
140.58	SYLLABLES PER 100 WORDS
6.58	SENTENCES PER 100 WORDS
9.44	FOG READING LEVEL
72.5	FLESCH READING EASE
6	FLESCH GRADE LEVEL
5.37	POWERS READING EASE
3.22	ARI
6.91	FLESCH-KINCAID
7.98	CÓLEMAN

FOR FRY, PLOT SYLLABLES PER 100 WORDS AND SENTENCES PER 100 WORDS ON FRY GRAPH.

graph on the vertical axis and enter the syllable data on the horizontal axis. The point where the two intersect is the approximate Fry grade level.

You will have the opportunity to get a paper copy of the final table. Answer YES to the question if you wish a printout and the computer's directions. When complete, the program will terminate. To start again and analyze material from another source, simply type RUN again. There is no need to load the program again.

Program Output

On the printed output of a sample run, you will see that the name of the document along with the total number of passages evaluated. As mentioned earlier, these passages can be of any length. The statistics given include the sum total of words, syllables, sentences and three syllable or longer words which are found in all the passages evaluated.

The program counts the number of words by looking for spaces. HELLO, PEI and 1982 are all counted as one word each. If P.E.I. was entered with periods, the program would count three sentences, one for each period, but only one word as no spaces occur. A comma or other punctuation after a word would not affect the word count but would cause one extra letter to be added to the punctuation. The slash at the end of a line simply tells the computer that the sentence continues on the next line.

The program estimates the number of syllables by dividing the total number of letters by a constant, 3.1127, and rounding the result to two places. The number of three syllable or longer words is estimated by checking for words with nine or more letters and three or more vowels.

A total of six formulas plus the Fry data (sentences and syllables per 100 words) are computed. The formulas are as follows: Fog, Flesch, Powers, ARI (Devereaux), Coleman, and Flesch-Kincaid. Following is a brief overview of each adapted from Schuyler's article. (1982). A partial bibliography of more detailed sources on readability is found in the Reference section.

Fry

The index is derived by plotting the average number of sentences per 100 words and syllables per 100 words on a graph. It provides an approximate grade level and reasons for the score. As a result, Fry scores are particularly helpful in suggesting ways to increase readability and in understanding why passages are rated as they are.

Fog

The formulas used in the Fog scores is as follows: $.4 * (T/W * 100 + W/S)$ when T equals three syllable words, W equals words,

and S equals sentences.

In Fog the percentage of the three syllable words expressed as a whole number is added to the average sentence length. This is multiplied by a constant, .4. This index tends to grade consistently high when compared to the other readability formulas.

Flesch

The formula for a Flesch count is as follows: $206.835 - .846 * L3 - 1.015 * (W/S)$, when L3 equals the syllables per 100 words, W equals words and S equals sentences.

This formula does not yield a direct grade level but provides a reading ease score beteen 0 and 100 based on the number of syllables, words and sentences. A high score indicates easy material. Gray (1975) has derived an appoximate grade level from this. He defines scores of 92-100 as grade 4: 81-90 as grade 5, and so on. No fractions of a grade level are used.

Powers

The formula for the Powers scores is as follows: $-2.209 + .0778 * (W/S) + .0455 * L3$. when L3 equals syllables per 100 words, W equals words and S equals sentences.

Although Powers uses the same variables as Flesch, this formula uses different constants and a different order of computation. The value derived relates directly to a grade level. The Powers index grades consistently lower than any of the others with the exception of ARI.

ARI or Devereaux

The ARI formula is as follows: $1.56 * WL + .19 * S1 = 6.49$, where W1 equals word length and SL equals sentence length.

This formula is the simplest formula of the readability formulas given. Although it correlates well with the others, it consistently grades lower. Schuyler found that when the other formulas gave levels of from 12 to 14, the ARI formulas often gave levels of from 5 to 6.

Flesch-Kincaid

The Flesch-Kincaid formula is computed as follows: $.39 * (W/S) + 11.8 * (Sy/W) - 15.59$, when W equals words, S equals sentences and Sy equals syllables.

This formula is a derivation of the Flesch formula which compute grade level directly. Unlike the Flesch scale, grade levels of less than four are also computed. For very simple material, negative values may be obtained.

Coleman

The Coleman formula is as follows: $-27.4004 * Cloze\% + 23.06395$. In the Coleman formula Cloze% equals $(141.8401 - .21459*(L1/W*100) + 1.079812*(S/(W/100)))/100$. Within the Cloze formula, L1 equals letters in the passage, S equals

sentences and W equals words.

This formula uses letters rather than syllables and correlates very highly with Fry. It does tend to grade somewhat higher than Fry but will also yield negative values for very easy material.

Schuyler (1982) points out that although all of these readability indices have a very high degree of correlation they may not be equally valid. A high degree of correlation among the scores should not be surprising as they are all using the same data and many of the same variables. However, it should be pointed out that the ARI index, closely followed by the Powers, consistently tend to score lower than the others while the Fog index tends to score consistently higher than the others.

A critical element in the evaluation of readability is sample size. There is some evidence (Coke and Rothkopf, 1970) to indicate that more than 10 per cent of the words in the material analyzed may have to be sampled in order to reduce the possibility of error to an acceptable level.

Limitations

This version of the program has certain limitations over the original Apple version. For instance, the Dale formula, which necessitates a 3,000 word list for comparisons, is not included. It simply overtaxed the memory of a 32k PET, although a means of modification are currently being considered.

Those who are interested in the Dale formula or the Apple version of the program should consult Schuyler (1982) for a complete program listing.

I will be happy to provide a PET version of the program to anyone who sends me a cassette audiotape (30).

References

Coleman, M. and Liau, T.L. A computer readability formula designed for machine scoring. *Journal of Applied Psychology.* 1975, *60*, 282-284.

Coke, E.U. and Rothkopf, E.Z. Note on a simple algorithm for a computer produced reading ease score. *Journal of Applied Psychology.* 1970, *54*, 208-210.

Flesch, R.F. A readability yardstick. *Journal of Applied Psychology.* 1948, *32*, 384-390.

Fry, E. Fry's readability graph: Clarifications, validity, and extension to level 17. *Journal of Reading.* 1977, *21*, 242-252.

Gray, W.B. How to measure readability. Philadelphia, Pa.: Dorrance and co., 1975.

Kincaid, J. Derivation of new readability formulas (Automated readability index, fog count, and Flesch reading ease formulas) for navy enlisted personnel. Arlington, Va., 1975. (ERIC Document Reproduction Service No. ED 108 134.)

Schuyler, M.R. A readability formula program for use on microcomputer. *Journal of Reading.* 1982, 560-591.

Smith, E. A Devereaux readability index. *Journal of Educational Research.* 1961, *54*, 289-303.

FIGURE 2
The Fry Readability Graph

Directions

Enter graph with average number of sentences per 100 words and average number of syllables per 100 words. Plot a dot where the two lines intersect. Area where dot is plotted will give you the approximate grade level.

This graph is taken from an article by Fry (1977) in the *Journal of Reading*.
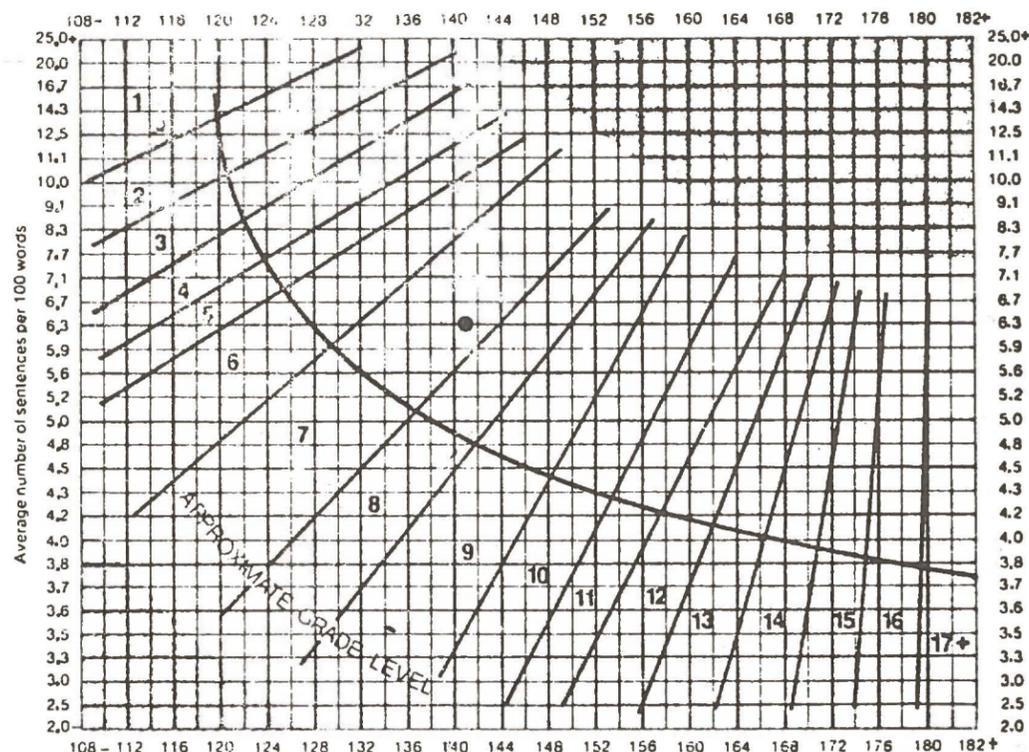
"...like the Arabs..."?

Denis Hlynka

The word is out. The information society is here. Those of us in AMTEC have, of course been expecting and even predicting this next stage in communications technology for some time. We always knew that our field would someday thrust itself to the forefront ever since Thomas Edison predicted for motion pictures a future ahead of its time:

Maybe I'm wrong, but I should say that in ten years textbooks as the principal medium of teaching will be as obsolete as the horses and carriages are now... Visual education, the imparting of exact information through the motion picture camera, will be a matter of course in all our schools.

(Colliers, Feb. 28, 1925, *V 75,* #8.)

Well, Edison was wrong. It didn't happen by 1935, nor for that matter by 1945, 1955, or 1965. It still hasn't happened. Similar predictions were made for other media which promised to revolutionize education: programmed instruction, television, games and simulations. Always the new media fell short of the promise.

And now the computer revolution is upon us, and somehow, this time, we have been caught standing on the corner, watching all the chips go by.

What are the implications for the computer/telecommunications revolution upon AMTEC? It seems to me that AMTEC has three obvious but equally possible alternatives. First, it can adapt immediately to new developments, strengthen its goals and resolves, and come out as a major leader in educational technology in Canada. Alternatively, AMTEC can maintain its rather low profile, status quo approach. Or, third, AMTEC may find itself anachronistic and unnecessary in a world of media specialization, and in the words of the poet, should "...fold its tents like the Arabs And as silently steal away."

Many of us perhaps would immediately reject this latter option, but if we are serious in examining AMTEC's future, I believe that we must allow all three options to stand as very real and viable possibilities.

The fact is that the study of the information society, far from being a timid field seeking a home, has already been claimed and welcomed by more than one eager group of adherents from engineers to futurists, while by and large, educational technologists have been by-passed. And, the fact is that we are moving to more specialized interest groups... a challenge which may be beyond the scope of AMTEC. Why attend an AMTEC '82 conference in Winnipeg when more specific options are available: a distance education conference (ICCE); a children and television conference; a videotex conference; a vocational computer conference; or a teleconferencing conference?

And if AMTEC doesn't survive through the '80s? It is a difficult question, not to be answered glibly. Unless the AMTEC membership is prepared to support fully a clearly defined set of objectives, we may indeed have no choice but to "fold our tents."

Parenthetically, it should be noted that the problems addressed here are not uniquely AMTEC's nor uniquely Canadian. Consistent rumors from the United States suggest that AECT, a much larger and healthier organization is likewise undergoing similar problems.

Which leads us to the major concern of this editorial... the future of the *Canadian Journal of Educational Communication.* Let us review what has been done so far.

First and most significant, Dr. Richard Lewis has taken our modest *Media Message* and given it a facelift coupled with a new name is a new philosophy. *Media Message* is now the *Canadian Journal of Educational Communication.* Its functions have expanded considerably. CJEC presents readers with profiles of educational media and communication technology happenings in Canada and elsewhere; it acts as the official organ of AMTEC; and it provides a channel for formal refereed professional and academic papers at the cutting edge of our field.

Second, under the editorship of Richard Lewis, the first four issues of CJEC have already appeared. Thus the first action steps have already been taken to give the journal a more significant role in AMTEC.

Third, a new editor of CJEC is about to take over. It is always difficult to predict precisely what impact a new editorial policy might have. Nevertheless, we think we can give you some glimpses into what you might expect from CJEC for the next two years:

Feature Articles. An attempt will be made to continue the recent trend of providing a mechanism for formal refereed papers as the backbone of the journal. The trend will be towards fewer but longer contributions within this section.

Profile. A common request from AMTEC members is to obtain more information on media organizations across the country. We hope to be able to profile such organizations on a regular basis within this column.

Update. A summary of current events and happenings across the country is the goal for this column.

ERIC. A review of significant current documents from the Educational Resources Information Exchange document service will be a regular feature.

Bibliographies. This section will feature basic print and media bibliographies on subjects of interest to educational technologists.

Fiction. This section is only in the experimental stages. Most media journals tend to be cognitive in approach and content. The premise of including a fiction section in this journal is based upon the assumption that there is indeed a body of work from the entertainment domain which approaches truth as a unique slice of life, a frozen moment in time. Should the concept be workable, CJEC will begin modestly by

Denis Hlynka is an Associate Professor of Educational Media and Technology in the Faculty of Education at the University of Manitoba. He is also the new editor of the Canadian *Journal of Educational Communication.*