# AN ATTEMPT TO ESTABLISH LEVELS OF READING DIFFICULTY IN TERMS OF SYNTACTIC COMPLEXITY

Keith Johnson - University of Papua New Guinea

(Received December 1971)

Consider this statement.

1) Kekawewechetushamikowanowow.

Assuming that the reader does not speak Cree Indian it is certainly semantically unintelligible. It is also syntactically unintelligible in that it is not possible to establish the grammatical relationships between the basic meaning units of the sentence. Since we cannot determine the basic units of the utterance it may be said to be morphologically unintelligible, and since we would be unlikely to pronounce it in such a way that it would be recognisable to a native speaker of the language, it may also be regarded as phonologically unintelligible.

Careful work with a Cree-English dictionary might lead to (2).

2) You will I wish together remain be you it man you

which is phonologically intelligible in that any English speaker can say it (though the stress and intonation patterns assigned to it would be similar to those for an unconnected list of items), morphologically intelligible in that he can chunk it into basic units, semantically intelligible at least in that he can define each of the lexical items, but it remains syntactically unintelligible. It is not possible to establish the meaningful relationships between these lexical items; and because it is syntactically unintelligible it is meaningless as a whole and no communication takes place.

Over the past forty years a great deal of research has been carried out on vocabulary, and the results, primarily in terms of word frequency counts, have been applied to the preparation of English language materials designed for second language (L2) speakers. However vocabulary control without syntactic control may be of little value.

(3) is taken from a Japanese primer which had a vocabulary restricted to the 500 most frequent headwords in English.

3) It is writing that makes what we read our own.

At upper secondary and tertiary levels, vocabulary control becomes less effective. All subject areas have their specialised vocabularies and, in the situation where English is the medium of instruction, teachers in the special subject areas can be relied upon to introduce their students to the specialised vocabulary required. Thus the L2 secondary school leaver who has had three years training at agricultural college as an agricultural extension officer understands the vocabulary of his profession well even though the specialised terms would have a low frequency count. His difficulty in reading the extension manual stems rather from the structural difficulty of the material. A similar problem arises at University in the selection of suitable texts for students. Lecturers and tutors can deal with vocabulary problems which arise, but where meaning is obscured for the L2 student through the structural complexity of a passage, it is difficult to establish the problem in the first place particularly when the teacher is a fluent speaker of English, and harder to find any solution that will assist the student in future.

There is then a need for research into factors which make a passage more or less 'readable' structurally, and for a measure which will determine the structural difficulty of a text, so that its suitability may be assessed for the use of L2 readers with different levels of competence in English, in particular for L2 readers whose professional training would make any assessment in terms of vocabulary counts irrelevant. It is hoped that what follows is a step in this direction.

The computation of this readability rating involves the analysis of samples of text, between 300 to 350 words in length. The analysis has four basic stages. Each stage is outlined, and then some details are added rather to illustrate the types of decision that have been made and have to be made than to present a full description.

A. The average number of words per sentence

The length of sentences is not important in itself, but rather in relation to 'B' (below). As (3) showed, a short sentence may be extremely complex, while long sentences may be extremely simple syntactically. (Nevertheless it is true that long sentences provide scope for complexity, and it is a good rule of thumb to maintain a high proportion of short sentences in simplifying material.)

Some Details

In counting the number of words:

I Hyphenated words are counted as 2 words

II Contractions - "don't", "isn't" etc - 2 words

III Numbers - "1971" - 1 word

IV Compound words - "New York, " "Port Moresby" - 1 word

V Lists of items or actions are taken out of the sample

B. Density of Information

Density of information is measured in terms of the average number of 'decoding units' per sentence, where a decoding unit may be described, very loosely, as a single run through the phrase structure rules in a generative grammar. A recursion gives an additional decoding unit. To define such a statement rigorously a generally accepted transformational generative grammar of English would be required, and such a grammar is not available. Again I will simply illustrate the type of decision which has been made.

Some Details

4) He likes pretending to be ill

5) The old man was angry in the park today.

(4) is regarded as having three basic decoding units:

He likes x

He pretends x

He is ill

while (5) is regarded as having only one: 'in the park' and 'today' being regarded as modifiers within the basic structure, or as requiring a single passage through a set of (as yet undefined) phrase structure rules.

II Contrary to generally accepted transformational generative theory, single adjectives (i.e. 'old' in (5) above) are not counted as additional decoding units. Psycho-linguistic evidence on this matter is summed up by James Deese: "adding adjectives to the subject and predicate nouns of sentences scarcely makes those sentences more difficult to perceive, to remember or otherwise process. There is a discrepancy between generative theory and human ability in this case..." (1)

Abandoning T. G. theory in this instance had the intuitively satisfying effect of raising the readability rating of Ian Fleming's James Bond stories . (The books had been widely read and enjoyed by my L2 students in Africa, yet had a fairly low readability rating

until this adjustment was made.)

Nevertheless second, third, fourth etc adjectives continue to be counted as decoding units, as do all verb based modifiers (i.e. "freed slaves"). 'The', 'a', 'this', 'that', 'these', 'those', 'all the' etc. are not counted at all.

III Evidence from experimental psychology suggests that expansions of the auxiliary do not influence ability to learn sentences, and errors in recall do not tend towards a simplification of the auxiliary (2). Therefore the length or complexity of the auxiliary has not been regarded as a factor in determining the number of decoding units, though it is perhaps unsatisfactory that no distinction is made between (6) and (7) in this analysis of structural complexity.

6) He drives to work

7) He must have had to drive to work.

8) He must have had to have been driven to work.

'Have' is recursive in (8) (have had to have) which would justify an additional decoding unit being counted.

IV 'Relational markers' such as 'when','because', 'how', 'then', 'later', etc. are not counted as decoding units. It is considered that the relationships between clauses and sentences in discourse exist whether expressed in the surface structure or not. The fact that they are expressed, rather than deleted, should increase the readability of the passage rather than decreasing it.

C. Ratio of 'Given' Information to 'New' Information in the Sample

The terms are those used by M.A.K. Halliday, but they are not used here as he has defined them (3). It seems reasonable to claim as a universal of the discourse structure of language that new information is carried by given information (given meaning a point of contact with what is already known.) And that the ratio of given to new effects readability. The readability of a good novel, of books in an area of study in which one is expert etc., may well relate to these elements within what Halliday calls the thematic function; and the business of 'getting into' a novel may be that of establishing a given element in the text. In the same way there will be a large 'given' element in any conversation between friends or

long standing acquaintances. However, the relationship between given and new elements in a text is extremely complex and co-reference, paraphrase and synonymy are theoretical problems that are far from being solved. The solution presented here is extremely crude, but hopefully seems to be effective.

Some Details

I Any noun or adjective which has been used previously in the sample is counted as given, all others are counted as new information. In support of this decision it might be added that this measure gives an indication of the size of the writer's vocabulary: a limited vocabulary being less complex for the L2 reader in that it will tend to consist of high frequency words, where the person with a large vocabulary will more often use low frequency words.

II Deictics (this that these those), and all pronouns are included under given.

III It is here that the length of the sample is important. New elements necessarily predominate in the first few sentences, but later the given element tends to catch up. The average ratio has been approximately 1/1 in the samples analysed so far. The two extremes can be illustrated by a sample from 'Bleak House', where the ratio of given to new was 1 to 3 as one might expect from a man of Dicken's literary distinction, and a sample from booklets prepared for L2 readers by the Reserve Bank of Australia where the ratio was 2 to 1.

D. Structural Complexity

Structural complexity is measured in terms of the average number of 'complexities' per sentence. In early analyses, all optional transformations and all embedded clauses were regarded as complexities, but this has since been modified.

Some Details

I (9) represents the 'normal' word order of English:

9) We borrowed money from John at the bank yesterday

and would not be counted as having a unit of complexity. (10, (11) and (12) would each count as having one complexity.

10) That we borrowed money from John is true.

11) What we borrowed from John was money.

12) It was money that we borrowed from John.

The passive transformation has also been counted as a complexity (12)

(13) Money was borrowed from John

but this decision may need to be revised since there is no very clear justification for counting (14) as a complexity while not counting (15)

14) The door was opened

15) The door opened.

Questions and negatives are not counted as complexities, and passives may be treated similarly in future. However it is possible that a preference for the passive indicates other aspects of a writer's style which may justify a lowered readability rating.

All front position transformations longer than a three word prepositional phrase are counted as complexities, (16):

16) When we were broke, we borrowed money from John.

since all such transformations impose a strain on the short term memory, which would not be the case if the subordinate 's') were in final position.

II All centrally embedded clauses count as complexities: they too impose a burden on the short term memory as well as leading to potential confusion. Right branching embedded clauses do not affect the reader in this way, and the first such clause is not counted; nevertheless the second and subsequent right branching embedded clause are counted as complexities.

17) We, who were always broke, borrowed from John.

18) We borrowed money from John, who was rich.

19) We borrowed money from John, who, being rich, could afford it.

Thus (17) and (19) would be counted as having one complexity each, while (18) has none.

III In compound sentences the complexity count begins again after each conjunction.

Summary of the basic data required to establish a readability rating

1. Number of words in the sample (300-350)

2. Number of sentences in the sample.

A. Average number of words per sentence.

3. Number of 'decoding units' in the sample.

B. Average number of decoding units per sentence.

4. Number of given and new elements in the sample.

C. Ratio of $\frac{\text{New}}{\text{Given}}$ elements in the sample.

5. Number of structural complexities in the sample.

D. Average number of structural complexities per sentence.

The computation whichgives the structural readability rating is as follows

$$\frac{A}{(B \times C) \times D}$$

The following illustrates the method of obtaining a structural readability rating. The length of the sample is short and a nominal 1/1 ratio is used under C.

20) The point we want to make is that everyone who has had anything to do with the making of a motor car - from producing the raw materials to the final task of assembling the parts - must be paid wages by the people who own and manage the many factories engaged in the industry.

A - 53

B - 13

C - 1/1

D - 10

$$\frac{53}{(13 \times 1/1) \times 10} = 0.4$$

The readability rating for (20) is 0.4. (20) illustrates clearly what has been noted in many samples: the attempt to be precise almost invariably leads a writer into syntactic complexity. In fact investigations so far suggest that complex thoughts are regularly expressed through complex syntax, and that certain types of complexity of thought are expressed through the complexity of the syntactic relationships between semantic units and may not be expressed in any other way. This fact is of considerable importance for students doing tertiary level studies in a second language.

## Some Structural Complexity Ratings

Some ratings are discussed to illustrate the results of analyses so far.

Material designed for young children, i.e. Readers for beginners tend to have a readability rating of infinity or approaching infinity, since D (average number of structural complexities per sentence) is low, and often zero.

Analysis of simplified readers designed primarily for L2 readers in junior secondary schools (4) gave an average reading of 6.7, and it appears from these tests that a level of 4.0 at least is desirable in such materials.

An analysis was also undertaken of booklet material prepared by the Reserve Bank of Australia for use in Papua New Guinea, and specifically for use in seconday schools. The average readability rating for these samples was 5.3 and the material was regarded as being suitable apart from certain lapses (see (20) above).

Analysis of materials from the Department of Agriculture Stock and Fisheries Extension Manual before and after revision gave figures of 1.2 and 3.4 respectively.

Most recently an analysis has been carried out on samples taken from texts used by preliminary year students at the University of Papua New Guinea. One text, judged extremely difficult by students, had a rating of 0.7 (5); while a second text, judged easier by students, had a rating of 1.2 (6). A third text designed specifically for L2 readers, its vocabulary deliberately simplified (7) had a rating of 2.7 an improvement certainly but by no means a high level of readability.

As a final illustration, an article (8) which had a rating of 0.8 was rewritten to make it more accessible for students. The rewritten version, though much simplified conceptually still had the very low rating of 1.8. A further rewriting gave the figure of 6.3 but the content as well as the syntax of the passage changed in the process, many subtleties and implications of ideas being lost.

## Conclusion

This is an approach towards a measurement of structural complexity and a tentative one at that. The figures given above are intended as illustrations only since over the period during which they have been obtained the techniques of analysis have been changed and reconsidered continuously. Nevertheless, in spite of such changes, the figures gained have consistently confirmed practical experience and what might be termed 'informed

intuition[1]. In addition to the hope that measurement will prove possible in this neglected area , the analyses carried out so far have suggested a number of interesting hypotheses as to the nature of structural complexity which will be tested experimentally. Finally, it is clear from recent work that syntactic analysis by computer is quite feasable, and programmes designed to provide readability ratings of texts for L2 readers would be extremely valuable.

REFERENCES


1. Deese, James "Psycholinguistics" Allyn & Bacon Inc. Boston, 1970 p. 44.

2. Garrett, M. and J. Fodor "Psychological Theories and Linguistic Constructs" reporting on experiment by J. Mehler, in Dixon and Horton (Eds.) "Verbal Behavior and General Behavior Theory " Prentice-Hall, 1968. p. 471.

3. Halliday M.A.K., "Notes on Transitivity and Theme in English" Journal of Linguistics 3, 37-81; 199-244; and 4, 179-215 (1967-68).

4. The Rapid Reading Series published by Nelson.

5. Childe, Vere, Gordon: "Man Makes Himself" Fontana Library 1966. p 1 sample begins paragraph 2, 358 words.

6. Buchanan R. A. "Technology and Social Progress", Oxford, Pergamon Press 1965. P 1 sample begins paragraph 2, 380 words.

7. Bruton, J. S. "The Story of Western Science" Cambridge University Press 1966. P5 sample begins paragraph 2, 341 words.

8. Needham Joseph, "Science and Society in East and West" in "The Science of Science: Society in the Technological Age" Maurice Goldsmith and Alan Mackay (Eds.) Souvenir Press Ltd. 1964. p 127 sample begins paragraph 1, 328 words.