

**LINGUISTIC COMPLEXITY AND TEXT COMPREHENSION - Readability
Issues Reconsidered** by Davison and Green (eds.)

by H.I. Bohn

(Universidade Federal de Santa Catarina)

The research presented by Davison and Green in **LINGUISTIC COMPLEXITY AND TEXT COMPREHENSION** addresses some of the main issues on readability, linguistic complexity and language understanding. The book is divided into ten chapters and an Introduction. In the introduction the editors justify the reappraisal of readability issues, summarize the content of the papers and put them in context. The main points discussed in this collection of ten articles are the definition and assumptions underlying formulas, problems and weaknesses of readability formulas, and new insights that recent research in linguistics, in the psychology of language and theories of learning have brought to the concept of language complexity, to language understanding and reading comprehension. The review will follow the thematic order in which the articles appear in the book.

1. Definition and assumptions underlying readability formulas

In the first paper of the collection, B. Bruce and A. Rubin ('Readability Formulas: Matching Toll and Task', 5-22) give the reader a good general overview of what readability formulas are, their usefulness and rules for the application of formulas. Readers familiar with readability issues will find the content of this article and the second one on 'Conceptual and Empirical Bases of Readability Formulas', by

Anderson and Davison (p. 23-54), quite basic and might want to skim over them swiftly.

Readability formulas were first developed in the 1920's. They sprang from the need to match reading materials with specific audiences and they have been widely used by teachers and administrators in the educational system, especially in the United States, to select reading materials for the classroom. More recently they found further applications in the preparation of public documents to adequate them to the reading level of specific audiences.

Formulas are based on the assumption that all reading problems can be related to certain text features, and that after these features have been identified, they can be entered into an equation and then numerical estimates of readability for a specific text can be established. The level of readability found will correspond to specific reading levels in the educational school system, or in the case of public documents to the reading level of the audience the documents are directed.

The most common features entering into readability equations are word difficulty - based on word frequency and word length (information drawn from a list of 3,000 words); and sentence difficulty - based on sentence length (average number of words per sentence).

One of the most widely used measures of readability is the Dale and Chall (1948) formula and it is based on the following equation:

$$\text{READING LEVEL} = 16 \text{ (% uncommon words)} + .05 \text{ (average number of words per sentence)}.$$

The data are drawn from three 100 word selections of a book or document under analysis.

The concept that text complexity is a physical feature or set of features which can objectively be taken from the text and plotted into an equation stems from the structural point of view of language description. This is a naive interpretation of textual structure and a naive way to look at the psychological and neurological baggage of the reader and the reading process itself. Under this perspective formulas contradict current knowledge about the reading process and they don't include in the measures important textual and personal variables that influence the understanding of a text.

In terms of the text the formulas overlook important syntactic complexity variables, discourse characteristics, information density of sentences, inferential needs required to understand the text, complexity of ideas, rhetorical structure, dialectal and style differences, etc. On the part

of the reader the formulas don't consider factors such as motivation, interest, purpose of reading, cultural background, circumstances in which the passage is read and general world knowledge. Bruce and Rubin caution the reader on the use of readability formulas and suggest six basic rules to follow in their application. In the next sections of the review some of the shortcomings above mentioned will be retaken and examined in detail, and some alternative proposals to look at linguistic complexity will be highlighted.

2. Problems with readability formulas

2.1 Conceptual problems

R.C. Anderson and A. Davison raise in their article on 'Conceptual and Empirical Bases of Readability Formulas' (p.23-54) further questions on the assumptions underlying readability formulas. On the conceptual level they discuss the inappropriate statistical model used in readability formulas and refer to the ill-founded theoretical underpinnings of the reading process underlying them.

What the authors do object in the statistical model by the proponents of formulas is the contradiction between the methodology of data analysis in the construction of formulas and the use made of the findings of this analysis. In the first instance the data are aggregated by grade while in the application of the formulas the results of this analysis are normally used to select books, texts for individuals or for small groups. For example, if a text has been chosen for a group of readers in which we have learners of the 2nd and 6th reading levels, the average fourth reading level would not be suitable for individual students in that group. On the other hand, when considering the more up-to-date interactive reading model in which text characteristics and characteristics of readers interact, again an aggregate model of statistical analysis does not seem adequate because of its inability to deal with individual reader's interactions. Additionally researchers have not developed models that can capture the interactions of text characteristics and neither have they accessed the interactions between specific text features and readers characteristics.

2.2 Empirical problems

Readability formulas also run into problems at the empirical level. One of the common elements entering into the equation of readability formulas is word difficulty. Research conducted by Anderson and Davison, has shown that even when significant more difficult words (longer or less frequent words) are introduced in a text, this only carries a maximum of 4% increase in its difficulty. Studies also show that derivatives and compounds, in spite of being longer than

primitive words, are normally easier to interpret, contrary to the predictions of readability formulas. Word difficulty is quite an obscure issue and it seems it will stay like that while we don't have more basic research on how individuals store and access the lexicon.

Another element included in formulas is sentence length. Research findings have largely documented that sentence length is not by itself a complexity element in comprehension. Longer sentences are often easier to understand than shorter ones. Anderson and Davison present the following sentences to exemplify their point:

1. I moved the switch.
2. The lights went off.
3. I moved the switch because the lights went off.
4. The lights went off because I moved the switch.

Sentences (3) and (4) are both easier to interpret than the two shorter sentences. This may signal to the fact that syntax is a more pertinent variable for understanding a text than sentence length. Left branching, for example, has proven to be systematically more difficult to interpret than right branching.

Right branching - It surprised me (that the cookies were brown).

Left branching - (That the cookies were brown)
surprised me.

One explanation for this difficulty is the way we process information and because of memory overloading. The bottleneck of linguistic processing in reading may be due less to sentence length than to factors such as memory overloading, reader's characteristics and text characteristics. Research has shown, for example, that interestingness "accounted for over thirty times as much variance in sentence recall as readability" (p.45). What makes a text difficult for certain readers is a problem of further investigation and it is unlikely that traditional readability formulas can incorporate the results of this research.

3. New insights into text complexity and comprehension

The information for these new insights come basically from three different sources: from the characteristics of the reader himself, from the characteristics of the text and from insights of linguistic descriptions and theories.

3.1. E.L.Baker, N.R.Attwood and T.M.Duffy ('Cognitive Approaches to Assessing the Readability of Text 55-84) retake the definition of reading and show how a bottom up process (in opposition to top down) which assumes that the

text contains all the causally related factors necessary for comprehension is an inadequate way to look at reading. It is the researchers claim that comprehension is "based on the information goals of the reader, her or his knowledge of the subject matter, and the representation of the information in the text" (p.57). Readers use their schemas on world knowledge and their expectations on text organization to interpret texts. To support their claim the authors present research on the reading performance of three groups of technicians with different training and working experience (experts and experienced technicians of a specific radar system, novice trainees, and expert technicians of another radar system). The subjects read two versions of a text. The original text was taken from a technical manual in the field of the first group of subjects and the adapted version was elaborated following the Kindaïd Readability Formula equation rules. The most significant findings of the study are related to the results of the third group of technicians who produced verbal summaries, on the revised version of the text, superior to both the experienced and the non-experienced technicians. It seems that world knowledge and the activation of specific context schemas can be held responsible for these results. The novice technicians might not have had the background knowledge necessary to call up such schemas during the reading of the text while the highly trained and experienced technicians might have been distracted by the unnecessary details of the adapted version. The lower density of semantic information per sentence in the revised version may also have contributed to the superior performance of the group.

This seems to indicate that an active process of meaning generation is involved in reading, and that it is based on the representation of the information in the text and on the world knowledge of the reader. It is, of course, not easy for formula proponents to include in their equations of readability the new insights of the reading process in which preexisting world knowledge of the reader, content and text schemas and goals for reading play an important role in understanding.

3.2 Formulas do not only fall short in including in their equations factors related to the reader. V. Charow's study ('Readability Vs. Comprehensibility: A Case Study in Improving a Real Document', (84-114) is an example of how other variables, such as general organizational rules, highlighting and use of appropriate vocabulary, can significantly contribute to the understanding of a real-world document directed to adult readers. Charow presented his subjects three versions of a recall letter directed to car buyers. Subjects read the original version of the letter, an adapted version of the letter for 11th grade readers according to a readability formula and a "guideline" letter that was produced following careful guidelines of organization, language choice and layout.

Readers answered a multiple choice questionnaire and an opinion questionnaire. In both surveys, contrary to the predictions of readability formulas, the guideline version produced the best results. The author suggests that the following factors used in the guideline document helped to improve the readability of the document: provision of a context, the presentation of the information in logical order, the informative headings, avoidance of impersonal or difficult grammatical constructions, absence of noun strings, avoidance or explanation of technical and unfamiliar terms, leaving out unnecessary words and using layout to make the letter legible (Cf.p.97-101). Charrow does not speculate on the reasons why the guideline version produced the best results. It seems that the different factors used in the guideline version trickered knowledge and text schemas that made the comprehension of the text easier for the readers.

3.3. Interestingness and motivation were pointed out at the beginning of the review as variables not included in the equation of readability formulas. G.M.Green and M.S.Olsen ('Preferences for and Comprehension of Original and Readability Adapted Materials, 115-140) carried out research controlling these two variables. Their findings seem to clearly indicate that original texts are significantly preferred by readers, especially by poor and older readers. Their research also showed that when adaptations were made from 2 reading levels up, the adapted materials were not significantly easier for children to understand than the original texts.

Motivation, style, challenge and satisfaction of solving problems are all important elements lost in the adaptation process. The authors battle off the adapters by saying that a child exposed to the adapted texts "may be seriously handicapped in understanding texts written in styles at variance with the prose of the homogenized texts that have been his primary model of written text".(p.139). The authors have harsh words for the trade industry and education policy makers who heavily rely on the criteria of readability formulas for the selection of reading materials with relevance for comprehensibility of texts and children's preference.

3.4 S.Kemper ('Inferential Complexity and the Readability of Tests, 141-166), in chapter 6, addresses the issue of the processes involved in reading and how linguistic factors can influence these processes making understanding more difficult. The author believes that readers approach texts segmenting sentences into clauses and identifying the clauses as referring to actions, physical states and mental states - who does what to whom and why - and, then, building a coherent event chain underlying the text. The process of identifying the clauses as referring to certain actions, physical states or mental states can be made more difficult

by a variety of linguistic characteristics. It is Kemper's claim that inference is one of the complicating syntactic devices that makes this assignment more problematic and therefore a source of comprehension difficulty for readers. Her findings indicate that when the Flesch readability formula was used in adapting materials, readers did not find the texts more difficult in a continuum from 7th to 11th levels of readability. But when inferential load was increased in the text, speed and accuracy of answers were affected. It seems that inference adds an extra load to working memory making the clause segmentation and the assignment of the clause to actions, physical or mental states to form the information chain more difficult. This is similar to the bottleneck phenomenon referred to by Anderson and Davison in chapter 2.

3.5. S.Crain and D.Shankweiler ('Syntactic Complexity and Reading Acquisition', 167-192) present the readability problem in an acquisitional perspective. They believe that, at least to a certain extent, reading is a bottom up process and is hierarchically organized, information flowing in a unidirectional and vertical ("bottom up") fashion so that lower levels serve as input to higher levels and not the reverse" (p.167-8). This could mean that if there is a working memory overload in lower levels, higher processing would automatically be affected. In this perspective Crain and Shankweiler introduce their two hypotheses. The Structural Deficit Hypothesis (SDH), and the Processing Deficit Hypothesis (PDH), to explain the reading difficulty of young children. The SDH claims that children have a structural deficit when they start reading and their problems in understanding texts can be related to the complexities of syntactic structures. The PDH claims that we can assume that, when children start to read, the linguistic structures have already been acquired and the difficulties generated in understanding a text stem from a processing difficulty directly related to working memory. The research on restrictive relative clauses presented by the author indicates that there is more evidence supporting the second hypothesis because children from an age as early as three years old are able to understand language involving restrictive relative clauses in spoken language. It seems, therefore, that working memory masquerades the upper levels of processing, i.e., difficulties are not generated by syntactic difficulties or structural deficits, but by a processing bottleneck that restricts the reader's access to higher linguistic processing systems. Attentional capacity and other maturation factors seem to be influential variables that distinguish good from poor readers, and not syntactic complexity.

3.6. L.Frazier ('The Study of Linguistic Complexity', 193-22) proposes that linguistic complexity can best be

understood if the relations between complexity data and specific grammatical representations and processing routes can be established. In this sense, a theory of language comprehension would be the best metric of text complexity because it would "characterize the complex interplay of different factors that contribute to complexity and reveal the reasons why some materials are more difficult to understand than others." (p.194). Frazier compares the language comprehension process to the teamwork of a group of experts working on the solution of a crossword puzzle. The solution of the puzzle can be delayed if the contribution of one of the members of the team is necessary input for the solution of the task of another expert. This is what seems to happen in a modular model of language processing - lexical structures are, for example, constrained by phonological rules and lexical items are categorized after they have been fitted into a syntactic structure. According to Frazier, the main tasks in the puzzle of language comprehension are the syntactic, lexical and semantic processes, and complexity can be defined in terms of the difficulty in assigning the input sentence a syntactic structure, in making a lexical analysis and give it a semantic interpretation. Then, Frazier goes on discussing the three modules involved in sentence processing.

In syntactic processing, ambiguity seems to be one of the main sources of difficulty. Ambiguous input introduces an extra load on working memory because all meaning possibilities have to be considered for sentence interpretation. For example:

1. John told the girl that Bill liked the story.
2. John told the girl that Bill liked the eggplant.

Sentence number (1) can have two meanings:

- a. John told (the girl that Bill liked) the story.
- b. John told the girl (that Bill liked the story).

The reader can only assign the correct interpretation to sentence (1), (a) or (b), after the correct parsing of the sentence has been performed.

The second process in language understanding is lexical processing. Because of limited research in the area it is not clear how the lexicon is recognized and processed in language understanding. Frazier suggests that there is not enough evidence to support "a single mental lexicon containing dictionary entries subsuming all different types (phonological, syntactic, and semantic) of information about a lexical item" (p.219). It might be important to pay attention to modularity within the lexicon, i.e., the vocabulary of phonological representation being quite different from the vocabulary of syntactic representation.

Finally, Frazier examines the semantic module. Semantic processing difficulty can only partially be related to the meaning complexity of words as outlined in dictionaries, or

phrase types. There are indications that one of the main sources of language understanding difficulty is context. It seems, therefore, that linguistic processing difficulties could be explained by the nonlinguistic difficulties in processing world knowledge rather than by difference of expression. Smith in her paper, similarly to Frazier, attributes the difference of reader's performance in reading, not to linguistic expression but to the task differences involved in comprehension.

3.7. J.H.Randall's ('Of Butchers and Bakers and Candlestick-Makers: The Problem of Morphology in Understanding Words', 223-246) research on word difficulty takes the reader beyond the word length and word frequency factors normally associated with readability formulas. Randal claims that there are word features and morphological processes that can be facilitating for the understanding of a text for some readers and have a negative effect for others. While the reader has not acquired specific morphological mechanisms, their presence in the text has a complicating effect on understanding, but after their mastery they may help the reader to assign to the word the right meaning, the correct syntactic function and give it the right semantic interpretation. Furthermore, Randal's research quite clearly shows that the knowledge of morphological processes makes the interpretation of unknown morphologically complex words easier than morphologically simple words for which there are no clues for the syntactic relationships of the element in the sentence. It seems that the relationship between word properties, such as length, frequency in the system, derivational processes and comprehension is quite a complex one and cannot be naively simplified as readability formula proponents have suggested.

3.8. The last article of the collection, written by the senior researcher of the University of Texas at Austin, Carlotta S. Smith "Factors of Linguistic Complexity and Performance", 247-280), addresses the question of linguistic complexity and performance. According to Smith a definition of linguistic complexity can only be achieved if the analysis is free from performance variables, i.e., complexity should be uniquely defined in relation to the processing of information and how meaning is related to the real world. Smith resources to the Government and Binding (GB) linguistic theory to anchor her discussions on a sound linguistic theoretical basis. The GB theory sees language in a modular organization. The modules interact among themselves and allow researchers to assign a structural description to each sentence at the phonological, syntactic and semantic levels. According to Smith "the various levels are all relevant to the generation and interpretation of a sentence, including material that is not phonetically realized. This suggests that factors besides those of surface structure are relevant to the processing of a sentence" (p.249). Within the GB theory Smith can reconize

four different levels of linguistic complexity: surface, syntactic, complexity, interpretative, and phonological complexity. The three first types are discussed in Smith's research.

Surface complexity is determined by AMOUNT (number of linguistic units in a sentence), DENSITY (the ways linguistic material is distributed in a sentence (NP, PP, etc) and AMBIGUITY (when a sentence can have multiple bracketing).

Interpretative complexity is related to three main determinants:

- the empty element - ex. (Mary want + past (e (to visit the zoo)))
- the (e) element, subject of the second S has to be supplied by the reader;
- semantic scope - a negative can, for example, refer to a specific constituent of a construction or it can refer to the full sentence;
- discontinuous constituent - such as the auxiliary in being.

Finally, systematic complexity refers to the constraints in the linguistic system. There are certain words or phrases in languages which add to the complexity of a sentence because they do constrain the application of the rule for affirmative sentences:

- + I lifted a finger to help her, in opposition to I didn't lift a finger to help her.

Smith researched the influence of the different levels of complexity on two tasks performed by children: a toy moving task and an imitation task. She hypothesized that the performance in these tasks would be sensitive to two different complexities: toy moving would be especially sensitive to interpretative complexity and imitation to surface complexity.

Her findings suggest that in the moving task, as hypothesized, her subjects (young children from 3.6 to 6.2 years old) were especially sensitive to interpretative complexity in opposition to surface structure and syntactic complexity. For the imitation experiment children were especially sensitive to surface structure complexity, constraints on the usage of nouns and verbs for ex., may have influenced significantly the results. It seems therefore that sentence complexity cannot be accounted for by a simple metric as advocated by formulas since complexity "may arise from the interrelation of different components of sentence structure." (Davison and Green, p. 4),

and also that "the relation between performance and linguistic complexity" (in this case defined as surface, interpretative, and systematic) "depends on the type of performance involved", i.e., different tasks may be affected by different factors of linguistic complexity.

4. Summary and conclusion

The content of the ten papers of the collection is vast and is not constrained by readability and language processing issues. The reader should be prepared to integrate in the discussions language acquisition principles, assumption underlying language description models such as government and binding, research design methodologies and statistic models of data analysis. Some of the articles are of a more general interest and aim to introduce the reader to the basic issues of language complexity, but the impetus of the research lies in the question of how people understand language and how certain linguistics forms, text and reader's characteristics can influence language understanding. The findings of the research are often not conclusive enough to fully answer all the questions raised but they do bring significant new empirical evidence of how language is understood.

Newcomers to the reading profession might find some of the content of the articles (Frazier and Smith, for example) quite demanding and hard to follow because reference is made to linguistic theories, research design and methodology that only more advanced readers might want to maze through and find enjoyable. Practitioners might find the book a little disappointing because the core content of the collection is not how to do things in the reading classroom but what learners do when they try to understand language. Davison and Green present the reading profession with a provocative book. It should be especially appealing to those interested in the question of how language works and how language is understood, and especially enjoyable and insightful for those who are unhappy with the manipulation of external variables to explain the reading process and performance, and who are willing to contribute to a better understanding of how linguistic, textual and reader's characteristics interact in language understanding.

Davison, Alice and Green, Georgia M. (eds.) **LINGUISTIC COMPLEXITY AND TEXT COMPREHENSION - Readability issues reconsidered.** New Jersey: Lawrence Erlbaum Associates, Inc., Publishers, 1988. Pp.viii + 291