16 TESL Reporter 32,1 (1999), pp. 16-27

The Role of Discourse Elements in Determining the Readability of Texts

Ramin Akbari, Tarbiat Modaress University, Iran
Mahmoud Reza Atae, Isfahan University, Iran
Hamidah Marefat, Tehran University, Iran

Introduction

In the realm of ELT/FLT, the concepts of grading and sequencing reading passages are of high importance, since for most second/foreign language learners, reading is the only means of contact with (and sometimes the sole purpose for learning) another language.

Subjective measures of reading difficulty have almost always proved inadequate and imprecise, leading the researchers to look for new ways of calculating readability. Different models and concepts of readability have been proposed, such as the Flesch Formula, the Dale/Chall Formula, and the Fair-Jenkins-Paterson Formula, to name a few. Almost all the available formulas and models rely on the syntactic or morphological characteristics of the text elements (words, phrases, sentences), while the semantic or "functional" aspects of the texts are ignored. This paper tries to look at the concept of readability from a discourse perspective and attempts to investigate the relationship between the formal measures of reading difficulty and their discourse counterparts.

Review of Literature

Readability Formulas

No matter what the motive of the readers, the text they are to read in their L2 must be of appropriate difficulty for them. Neglect of this principle will lead to boredom or frustration on the part of the reader.

To match a text's difficulty level with the reader's perceived level of competence, a number of researchers have developed what are termed as "readability formulas." Readability formulas can be defined as mathematical equations used for the determination or the prediction of the level of reading competence necessary for the comprehension of a particular piece of writing.

A readability formula, according to Klare (1963), uses counts of language variables in a piece of writing in order to provide an index of probable difficulty for the reader. It is a predictive device in the sense that no actual participation by the reader is needed.

Most of the readability formulas rely on linguistic criteria for the prediction of text difficulty. Semantic difficulty and syntactic complexity are two such measures. Semantic difficulty refers to the ease of recognition and comprehension of individual words within a reading passage. Semantic difficulty is measured either by the length of words (i.e., the number of syllables or letters in a word) or by the absence of a word on a word list (based on frequency counts).

Syntactic complexity, on the other hand, refers to the difficulty of the structure of language, and is usually analyzed on the sentence level. Most research studies have related syntactic complexity to sentence length. Thus, the mean number of words per sentence has been the most common measure of syntactic complexity.

Many readability formulas have been developed over the years. Out of these numerous formulas, some have grown in popularity and are more frequently used than the others. This popularity is partly due to the ease of application of these formulas, and partly due to their relative precision.

One of the earliest developed readability formulas is that of Flesch (Klare, 1984). The formula, which was designed for general (adult) reading materials, uses average sentence length in words, the number of personal references, and the number of affixes, as predictors of reading difficulty of a text. In subsequent formula, Flesch added "the number of syllables per 100 words" as another predictor of readability.

Dale and Chall are two other researchers who developed a practical readability formula. (Klare, 1984; Mirzaee, 1991). Their readability measure uses average sentence length and the relative number of words out of Dale's list of 3,000 words, as estimates of a text's difficulty level.

The Fox Index is a widely used readability formula (Klare, 1988). Average sentence length and the percentage of words of three or more syllables are the two elements on which the readability formula relies.

Another popular readability formula is that of Fry (1968). In this user-friendly formula, Fry makes use of the number of syllables per 100 words and the number of words per sentence as measures of readability. The user simply enters the count of these variables in a graph and reads the readability grade score directly from it.

Each of these has been used with some success, particularly with native speakers, but they also have been criticized to some extent for their narrowness in determining readability.

Discourse Elements

The concepts of "theme" and "rheme" have received relatively wide attention in the literature related to "Functional" or systemic grammar. The development of Discourse Analysis and Pragmatics has also added to the importance of these two concepts (see, for example: Brown & Yule, 1983; Cook, 1989, 1994; Eggins, 1994; Ghadessy, 1995; Halliday & Hasan, 1990; McCarthy, 1991).

However, the most extensive treatment of the concepts of theme and rheme is that of Halliday (1985). Halliday defines theme as "the element which serves as the point of departure of the message: it is that with which the clause is concerned" (p. 38). Whatever is not the theme is the rheme of the message. It becomes clear from these definitions that a message consists of two major elements: theme + rheme.

In English, theme is identified through word order; it is usually the element that comes first in the clause. In declarative sentences, if the theme and the grammatical subject of the clause coincide, then we label the theme as unmarked. Lack of conflation of these two elements will lead to a marked theme in a declarative sentence. Sentence 1 is an example of an unmarked theme; sentence 2 is an instance of marked theme.

1. I caught the first ball.
Theme

2. Today, I learned that Mary had a little lamb.
Theme

Themes can also be multiple or simple. Simple themes are made up of one element or "two or more elements forming a single complete element" (Halliday, 1991, p. 41). Sentence 3 is an example of a simple theme:

3. The Walrus and the Carpenter were both walking.
Theme

A multiple theme, on the other hand, appears when the first element in the clause does not function as subject or complement. As a result, the subject, complement, or adjunct next following would be regarded as a part of the theme. In other words, in a multiple-theme "part of the clause functioning as theme has a further internal structure of its own." (Halliday, 1985, p. 53).

The internal structure of a multiple-theme is the result of the interaction of three semantic processes or meta-functions: Ideational, Interpersonal, and Textual.

Ideational meaning of the clause deals with its representational aspects. Here we treat the clause as the representation of experience. Interpersonal meaning is meaning as a form of action, and textual meaning looks at the relevance of a clause to its context.

A theme must always have an ideational element. However, the presence of interpersonal and textual meaning is not obligatory. Nevertheless, if a theme contains all three types of meaning at the same time, their order will be Textual + Interpersonal + Ideational. The ideational component in the theme is an entity that acts as subject, adjunct or complement, which is sometimes referred to as "topical theme".

Thematic structure is not just limited to clause elements; it can also be detected at clause and text levels. An example of theme at clause level is "predicated theme" or what is termed in traditional grammar as "cleft-sentence". The function of such themes is the "explicit formulation of contrast" (p. 60), as it is indicated in the following sentence:

> It was Mary Magdalene, not Mary the Mother of Jesus, who had been the real, if secret object of Mariolatry cults through the ages.

The choice of theme in clauses of a text is not haphazard; it is in fact one of the organizing principles of any type of discourse, what Halliday (1985) refers to as the method of development of a text.

Core vs. Non-Core Words

Vocabulary has been a relatively neglected variable in foreign language instruction (Celce-Murcia and Rosenweig, 1979; Seal, 1991). The demise of Audiolingualism and the rejection of its linguistic foundations did not result in the long-expected revival of lexical studies. It does not mean that no attention has been paid to the role of vocabulary in L2 instruction; there have been a number of valuable studies dealing with the role of L2 lexical items in foreign language instruction (for a review, see Carter & McCarthy, 1988; and Schmitt & McCarthy, 1997). However, the lion's share of research has gone to studying the L2 syntax.

Early L2 lexical studies were mainly in the form of word lists and frequency counts (Thorndike & Lorge, 1944; West, 1959). Few serious attempts have been made to classify word categories in L2 through semantic or functional criteria. A relevant study, however, is a paper by Carter (1988). Carter classifies vocabulary items into two general groups: Core and non-core. "The term core vocabulary is used to describe those elements in the lexical network of a language which are unmarked. That is, they usually constitute the most normal basis of simple words available to the language user" (p. 9). According to Carter, core words have clear synonyms, high collocation frequency, are

used for defining words, lack connotations, do not belong to a specific domain, and ar usually superordinate terms. Words that do not possess these qualities are non-core.

The Study

Aims

The study reported here aimed to find out the extent to which discourse elements ca determine readability of texts. So many readability formulas have been developed t determine the appropriacy of passages for an intended population. Attempts were mad in this study to see if the discourse elements could be taken as reliable criteria for rankin the texts according to their levels of difficulty. Significant correlations between th rankings gained by the existing readability formulas, the reliability and validity of whic have already been established, and the rankings gained by the use of the criteria base on discourse elements may show that the new profile is also dependable.

Research Questions

The following research questions were addressed in this study:

1. Will there be a significant relationship between the different rankings of the tex based on the Flesch formula and the discourse elements?
2. Will the rankings of the texts based on the Flesch formula highly correlate wi the ranking of the same texts based on advanced students' performance.
3. Will the ranking of texts based on the Flesch formula correlate highly with t ranking of the same texts based on intermediate subjects' performance?
4. Do the ranking of the texts based on the discourse elements significan correlate with the advanced students' performance, resulting in the conclusi that the ranking of the texts based on discourse elements are at least as relial (for advanced students) as readability formulas?
5. Do the ranking of the texts based on the discourse elements significan correlate with the intermediate students' performance, resulting in the conclusi that the ranking of the texts based on discourse elements are as reliable (intermediate students) as readability formulas?

Subjects

A total of 208 subjects participated in this study. All the subjects were English ma students studying at an Iranian state university. One group of 136 subjects took a reti TOEFL and 72 others took another retired TOEFL, the listening parts of which had b

omitted. For the purposes of this study, the subjects' performance on the Reading Comprehension texts were taken for data analysis. Of course, the classification of subjects into advanced, and intermediate groupings were based on their performance on the whole tests. The advanced subjects were those whose score was one standard deviation above the mean and the intermediate ones were those who score was between .5 standard deviation above the mean or 0.5 standard deviation below the mean. From among the subjects selected in this way, only those who had answered all the comprehension questions of the texts of each TOEFL were chosen. Altogether, the data related to 65 subjects entered the process of data analysis: 26 from the advanced level and 39 from the intermediate level. The lowest proficiency group, whose scores were less than 0.5 standard deviation below the mean, were excluded from the study because they failed to answer all the comprehension questions.

Procedures

Data collection: The ten texts, five from TOEFL 1, and five from TOEFL 2, were analyzed from different viewpoints.

First, the Flesch grade level of difficulty of each text was determined.

Second, each text was analyzed for the number of core and non-core words it contained. The framework used for this purpose was that suggested by Carter (1988).

Third, each text was analyzed for the thematic structure it contained; the number of simple themes, multiple themes, ellipted themes, clauses as theme, and the independent clauses was counted.

Fourth, the number of marked and unmarked topical themes were counted for each text.

Fifth, the number of the correct answers given by subjects for each text was counted once for the advanced subjects and once for intermediate ones.

Data Processing

First, for the simple themes, multiple themes, ellipted themes, unmarked and marked topical themes, the proportion of their frequency over the total number of independent clauses was calculated.

Second, all the data gathered through the first stage were used to rank the texts. Each text was given a rank based on several criteria including: The Flesch grade level; the number of simple themes; the number of clauses as themes; the number of unmarked themes; the number of marked themes; the number of core words; the number of non-core words; the number of correct responses given by advanced students; the number of correct responses given by intermediate students; and the number of independent clauses.

Table 1 shows the ranking given for each text based on the results of all these analyses:

Table 1

The Ranking of Texts According to all Criteria

Texts	simple theme	multiple theme	ellipted theme	clause as theme	unmarked theme	marked theme	core words	non-core words	Flesch Grade level	Advanced students	e students
1	9	5	2	7	5.5	6	7	10	7	7	10
2	1	8	8	9	2	9	4	5	8	10	9
3	6	7	9	2	3	7	6	2	5	1	5
4	2	9	10	3	8	4	5	7	2	5	7
5	4.5	3.5	6	4	7	5	2	1	3	3	1
6	10	1	5	6	5.5	3	10	9	10	9	8
7	7	2	4	10	1	8	3	3	6	8	6
8	3	10	7	1	10	1	9	6	1	6	4
9	8	6	2	5	9	2	1	4	4	4	3
10	4.5	3.5	2	8	4	10	8	8	9	2	2

Results and Discussion

Addressing the first question, the rank order correlations between the Flesch grade level and the discourse elements were calculated. The results are shown in Table 2.

Table 2

The Rank Order Correlation Between Flesch and Discourse Elements

	Simple theme	Multiple theme	Ellipted theme	Clause as theme	Unmarked theme	Marked theme	Core words	Non-core words
Flesch	.407	.638	.423	.721	.644	.566	.33	.442

To test these correlations for significance, the obtained values were transformed into *t*-values and a null hypothesis was formulated (Table 3).

Table 3

The *t*-values for the Correlations Between Flesch and Discourse Elements

	Simple theme	Multiple theme	Ellipted theme	Clause as theme	Unmarked theme	Marked theme	Core words	Non-core words
Flesch	1.26	2.312*	1.32	2.944*	2.383*	1.929*	.988	1.395

As can be seen, there is a significant correlation (t_{crit} =1.86, p=.05, df=8) between the ranking of texts based on the Flesch formula and that based on some of the discourse elements including multiple theme, clause as theme, unmarked theme, and marked theme. In other words, the rankings produced by the use of multiple theme, clause as theme, unmarked theme, and marked theme are as reliable as those produced by the use of Flesch.

Research question 2 addresses the reliability of the Flesch formula itself: First, the texts were ranked according to the subjects' performance. The attempt was made to find out the extent to which this ranking correlates with the ranking produced by the Flesch formula. The performance of the advanced and intermediate subjects were dealt with separately.

Tables 4 and 5 show the rank order correlation and *t*-values between Flesch and the performance of the subjects for the two levels of proficiency.

Table 4

The Rank Order Correlation Between Flesch and the Performance of Advanced and Intermediate Subjects

	Advanced	Intermediate
Flesch	.345	.381

Table 5

The *t*-values for the Correlations Between Flesch and the Performance of Advanced and Intermediate Subjects

	Advanced	Intermediate
Flesch	1.413	1.168

Since the necessary value of *t* for significance at the .05 level with 8 degrees of freedom is 1.86, we cannot reject our null hypothesis. In other words, that which Flesch predicts to be difficult is not so for advanced students, neither is it difficult for the intermediate ones; though, of course, there is a trend in the expected direction.

To answer questions 4 & 5, the rank order correlations between the rankings of texts based on the performance of the advanced students and the rankings based on discourse elements were calculated. The results are shown in Table 6.

Table 6

The Rank Order Correlations Between the Advanced Students' Performance and the Discourse Elements

	Simple	Multiple	Ellipted	Clause	Unmarked	Marked	Core	Non-core
Advanced Students	.06	.109	.006	.667	.133	.054	.475	.503

As can be seen from Table 7, a good determinant of the level of difficulty of texts for advanced students is clause as theme; core words and non-core words also show a trend in this direction.

Table 7

The *t*-values for the Correlations Between the Advanced Students' Performance and the Discourse Elements

	Simple	Multiple	Ellipted	Clause	Unmarked	Marked	Core	Non-core
Advanced Students	.172	.311	.017	2.691*	.378	.154	1.53	1.646

As far as the intermediate subjects are concerned, the rank order correlations between the rankings of the texts based on their performance and the discourse elements are calculated and later adapted to *t*-values. Tables 8 and 9 show the results.

Table 8

The Rank Order Correlations Between the Intermediate Students' Performance and the Discourse Elements

	Simple theme	Multiple theme	Ellipted theme	Clause as theme	Unmarked theme	Marked theme	Core words	Non-core words
Intermediate	.145	.072	.3	.303	.322	.127	.503	.539

Table 9

The *t*-values for the Correlations Between the Intermediate Subjects' Performance and the Discourse Elements

	Simple theme	Multiple theme	Ellipted theme	Clause as theme	Unmarked theme	Marked theme	Core words	Non-core words
Intermediate	.417	.206	.891	.857	.962	.363	1.646	1.818

The results show that the best determinant (though not statistically significant) for the readability of texts for intermediate students is the non-core words.

Conclusion

The results of this study indicate that certain discourse elements can be reliable indicators of the readability of texts. If texts are ranked based on the number of multiple themes, clauses as themes, unmarked and marked structures they contain, the results will

highly correlate with the ranks given to texts based on the Flesch Formula. However, the Flesch Formula ranking of texts is different from that obtained by the performance of the advanced or intermediate students. But the rankings based on the clauses used as themes significantly correlate with rankings based on the advanced subjects' performance. And for the intermediate subjects, the ranking based on non-core words are better determinants of text difficulty.

In summary, readability is the product of many text and learner-related variables. The more variables are taken into consideration in determining readability, the more likely such a measure will possess predictive validity. The measure of discourse-level elements described in this study seem equally as predictive of student performance as traditional syntactic/morphological ones, and deserve inclusion in further studies on the determination of what is more or less difficult; what is readable, and what is not.

References

Beverly, Z. & Samuels, S. J. (Eds.). (1988). *Readability: Its Past, Present, and Future.* Delware: International Reading Association.

Brown, G., & Yule, G. (1983). *Discourse Analysis.* Cambridge: Cambridge University Press.

Carter, R. (1988). Front pages: Lexis, style and newspaper reports. In Ghadessy. (Ed.). *Registers of Written English: Situational Factors and Linguistic Features.* London: Printer Publishers.

Carter, R. A., & McCarthy, J. J. (1989). *Vocabulary and Language Teaching.* London: Longman.

Celce-Murcia, M. (Ed). (1991). *Teaching English as a Second or Foreign Language.* Massachusetts: Heinle & Heinle Publishers.

Celce-Murcia, M., & McIntosh, L. (Eds.). (1979). *Teaching English as a Second or Foreign Language.* New York: Newbury House.

Celce-Murcia, M., & Rosenweig, F. (1979). Teaching vocabulary in the ESL classroom. In Celce-Murcia, M. & McIntosh, L. (Eds.). *Teaching English as a Second or Foreign Language.* New York: Newbury House.

Cook, G. (1989), *Discourse.* Oxford: Oxford University Press.

Cook, G. (1994). *Discourse and Literature.* Oxford: Oxford University Press.

Eggins, S. (1994). *An Introduction to Systemic Functional Linguistics.* London: Printer Publishers.

Fry, E. (1968). A readability formula that saves time. *Journal of Reading, 11*, 513–516.

Ghadessy, M. (Ed.). (1988). *Registers of Written English: Situational Factors and Linguistic Features.* London: Printer Publishers.

Halliday, M. A. K. (1985). *An Introduction to Functional Grammar.* London: Edward Arnold.

Halliday, M. A. K., & Hasan, R. (1990). *Language, Context and Text: Aspects of Language in a Social-Semiotic Perspective*. Oxford: Oxford University Press.

Klare, G. (1963). *The Measurement of Readability*. Iowa: Iowa State University Press.

Klare, G. (1984). Readability. In Pearson, P. D. (Ed.). *Handbook of Reading Research.* New York: Longman.

Klare, G. (1988). The formative years. In Beverly, Z. & Samuels, S. J. (Eds.). *Readability: Its Past, Present, and Future*. Delware: International Reading Association.

McCarthy, M. (1991). *Discourse Analysis for Language Teachers.* Cambridge: Cambridge University Press.

Mirzaee, A. (1991). *Reliability of Readability Formulas and Graphs*. Unpublished M.A. Thesis, Tehran University: Tehran.

Pearson, P. D. (Ed.). (1984). *Handbook of Reading Research.* New York: Longman.

Schmitt, N., & McCarthy, M. (Ed.). (1997). *Vocabulary Description, Acquisition and Pedagogy*. Cambridge: Cambridge University Press

Seal, B. D. (1991). Vocabulary learning and teaching. In Celce-Murcia, M. (Ed.). *Teaching English as a Second or Foreign Language*. Massachusetts: Heinle & Heinle Publishers.

Thorndike, E. L., & Lorge, I. (1944). *The Teacher's Word Book of 3,000 Words.* New York: Teacher's College, Columbia University.

West, M. (1959). *A General Service List of English Words*. London: William Clowes & Sons, Ltd.

About the Authors

Ramin Akbari is a Lecturer at Tarbiat Modaress University and a Ph.D. candidate in TEFL at Isfahan University, Iran. His research interests include: the role of intelligence in learning foreign languages and second language learning strategies.

M.R. Atae is a Ph.D. candidate in TEFL at Isfahan University, Iran. He has taught EFL/EAP programs and is at present working on a nationwide EAP evaluation project. His research interests include EFL reading, EAP, and discourse analysis.

Hamideh Marefat is a Lecturer at Tehran University English Department and a Ph.D. candidate at Isfahan University, Iran. She is currently working in the area of psycholinguistics and her work investigates the way foreign language learners' mental comprehension system parses sentences.