AN ANALYSIS OF READABILITY OF TWO NATIONAL DAILIES, TIMES OF INDIA & THE HINDU

Dr. Pramila B. Kunnur
Associate Professor& HOD
Dept. of Mass Communication and Journalism
Maharaja's College
University of Mysore
Mysore, Karnataka-570005

Profile of news papers

Times of India		The Hindu	
Type	Daily newspaper	Type	Daily newspaper
Format	Broadsheet	Format	Broadsheet
Owner(s)	The Times Group	Owner(s)	The Hindu Group, and Kasturi and Sons Limited
Publisher	Bennett, Coleman & Co. td.	Publisher	N. Ravi
Editor in chief	Jaideep Bose	Editor	Suresh Nambath[1]
Founded	3 November 1838	Founder(s)	G. Subramania Iyer
Language	English	Language	English
Headquarters	Mumbai, Maharashtra, India	Founded	20 September 1878;
Country	India	Country	India
Circulation	2,640,770 daily[1] (as of January–June 2019)	Circulation	1,404,901 daily[2] (as of January–June 2019)
Sister newspapers	The Economic Times	Headquarters	Chennai, Tamil Nadu, India
	Navbharat Times	Sister Publications	The Hindu Tamil
	Maharashtra Times		The Hindu Business Line
	Ei Samay		Sportstar
	Mumbai Mirror		Frontline
	Bhubaneswar Times		
ISSN	09718257	ISSN	0971751X
OCLC number	23379369	OCLC number	13119119
Website	timesofindia.com	Website	www.thehindu.com
			www.hindutamil.in

Readability

Dale and Chall (1948) proposed a comprehensive definition of readability: "In the broadest sense, readability is the sum total (includingthe interactions) of all those elements within a given piece of printed material that affectsthe success a group of readers have with it) The success is the extent to which they understand it, read it at an optimum speed, and find it interesting." (2)11

According to G. Harry McLaughlin (1968) in "Proposals for British Readability Measures" states that one of the least ambiguous published definitions of readability is that given by English and English in their Dictionary of Psychological Terms. 'Readability', they say, 'is the quality of a written or printed communication that makes it easy for any given class of persons to understand its meaning, or that induces them to continue reading'. This means that readability is that combination of qualities which makes a text comprehensible or compelling or both, Variables Affecting Readability12, 13, 14

Irving Lorge, who was interested in psychological studies of language and human learning, wanted a simple formula for predicting the difficulty of children's books in terms of grade score. With the help of Mc Call –Crabbs standard Test lessons in Reading (1926) and Thorndike. McCall Reading Scale, Large developed for formula for children's reading. This was also found useful for adult material also. He could identify three elements that affect the readability.

Research Methodology

The present study compares the two newspapers ''The Hindu'' and ''The Times of India'' with content analysis was still a popular method, used in about one-third of all published articles (Roger D.Wimmer & Joseph R. Dominick 9th edition 2011:156). The present study being a comparison of two national dailies in general without focusing a specific category of the content – content analysis becomes the choice of the researcher.

Focus of the Study

In this study the focus is on the readability of the news reports of identified categories from "'The Hindu'" and ''The Times of India''. The readability of the news reports are measured in terms of Reading Ease and Reading Grade Level (after Grade level) using the readability formulas available. These scores obtained for both the news papers are compared so as to know about the styles of the two newspapers.

Also, the focus is on the basic measurable language factors that affect the readability viz., Average Sentence Length (ASL), and Average Word Length (ASW). Again these measures are obtained are compared with readability scores obtained so that the researcher can conclusively say something concrete about the signature styles of the two news papers.

Variables used in the study

Newspaper scores on the writing style variables were determined by use of the software developed by Long (1989).The variables were:

1) Average sentence length, or the number of words pre sentence.
2) Average syllables per word,
3) Grade-level readability.
4) Flesch's human interest measure, which as noted earlier is determined by such elements as number of quotes and pronouns (Flesch, 1962).
5) Flesch's readability measure, which takes into consideration such elements as the length of words and length of sentences.

The variables affecting readability fall under two categories viz., text variables and design variables. As for the purpose of the present study only text variables are considered.

Some other text variables which have effect on readability are image words, abstraction, direct and indirect, stationary, types of narration and sentences, phrases and clauses.

Readability factors are those that affect the level of proper reading and understanding of a text. There falls can broadly be clarified as

- Reader factors (related to the readers age for his reading ability)
- Text Factors
1. Certain aspects of words (word length, frequency, vocabulary load, using unusual or abstract works.
2. Average sentence length is an important feature that affects readability to a text and is widely used and studies in readability research.

Clarity of ideas mentioned in the text, **topography, metaphor and simile** usually affect this readability.

Flesch (1962) determined that readability depended on two factors: the difficulty of reading and a "human interest" measure. Reading difficulty deals with (1).length of words and (2). Length of sentences.

Human interest, on the other hand, assumes that people will find material easier to read and will be more motivated to read it if it contains **(1). personal references and (2). direct speech**. Researchers have attempted to apply Flesch's readability notions to journalistic writing styles.

In order to find reliable answers for the above research questions the independent and dependent variables involved in the study are identified and are measured using a appropriate tools and techniques.

Dependant variables

Readability of the news reports

Independent variables

- Length of sentences (measured in terms of number of words contained in a sentence)
- Length of words (measured in terms of number of syllables contained in a word)

Operational definitions

1. Readability
2. Reading Grade Level
3. Average Sentence Length
4. Word length
5. Text difficulty

Readability: Readability is defined as "the ease of understanding or comprehension due to the style of writing "(George Klare, 1963).

Reading grade level: The www.readabilityscore.com where the online version of the readability test administering software is available defines readability as "The number of years of education a person has had". E.g.: A passage of grade level of around 8 is appropriate for general public and a passage with grade level and 10-12 is for highly school graduates (All this based on USA education system)

Average Sentence Length (ASL): The length of sentence is an important linguistic variable that affect the readability. The most commonly used definition of average sentence length is "the average sentence length is the number of words contained on 'a sentence'. It is calculated by dividing total number of words in a passage by the total number of sentences in the passage.

Word Length (ASW): The word length refers to the number of syllables contained in a word. It is obtained by dividing the total number of syllables in a passage by the total number of words in it.

Text difficulty: The term text difficulty, a concept in linguistics is also known as text accessibility. Sometimes text difficulty is used synonymously with readability. The factors affecting text difficulty are: (1) Book & print texts (2) Vocabulary (3) Themes & ideas (4) Sentence complexity (5) language of literary text. (6) Content (9) Text Structures (Victoria Davids, 2011).

Hypothesis of the study

In order to achieve the objectives of the study the following hypothesis were formulated:

1) The overall readability of the news reports of ''The Hindu'' does not differ significantly from the overall readability of ''The Times of India''
2) The overall grade level of the news reports of ''The Hindu'' does not differ significantly from the overall grade level of the news reports of ''The Times of India''.
3) The hard news reports of ''The Hindu'' and ''The Times of India'' do not differ significantly in their average sentence length (measured in terms of the average number of words contained in a sentence).
4) The soft news reports of ''The Hindu'' and ''The Times of India'' do not differ significantly in their average sentence length (measured in terms of the number of words contained in a sentence).
5) The hard news and soft news reports of '' The Times of India'' do not differ in their word length (measured in terms of the average no of syllables per word)
6) The soft news reports of ''The Hindu'' and 'The Times of India' do not differ significantly in their word length (measured in terms of the average number of syllables per word).
7) The hard news and soft news reports of ''The Hindu'' do not differ significantly in their average sentence length (measured in terms of the number of words per sentence).
8) The hard news and soft news reports of ''The Times of India'' do not differ in their word length (measure in terms of the average number of syllables per word).
9) The readability of the hard news reports of ''The Hindu'' does not differ significantly from readability of the hard news reports of ''The Times of India''.
10) The average grade levels of the soft news reports of ''The Hindu'' and ''The Times of India'' do not differ significantly.

Sentence length and readability: The 'easy guide'66 for its writers in bureau around the world relates the sentence length to the readability as shown below.

Table showing the number of words with readability standards

Average Sentence Length	Readability
8 words or less	Very easy to read
11 words	Easy to read
14 words	Fairly Easy to read
17 words	Standard
21 words	Fairly difficult to read
25 words	Difficult to read
29 words or more	Very difficult to read

Research Design

The research titled "Measuring the Readability of Newspapers: A Comparative Study of Two National English Dailies" requires the measurement of readability of the selected news reports of the two national dailies of India – ''The Hindu'' and ''The Times of India'' and comparing them with each other. The different steps involved in this research are

Flow Chart of Research Design In Measuring The Readability of Newspapers

Sample for the study

Universe

The two dimensions that are used to determine the appropriate universe for a content analysis are: (i) The topic (ii) The time period

Here in this study the news content of ''The Hindu'' and '' The Times of India'' from July 2014 to June 2015 can be considered the Universe to be studied.

Sample

Considering the vastness of the data (Two national dailies published for twelve months) it was decided to choose a representative sample of the data. This study does not focus on any specific given category of content of the newspaper. Therefore, it was decided to study the two prominent genres of the content

Hard News: These are the reports where the focus is on the fact and the opinion of the journalist is kept outside the report. The journalist remains as objective as possible while reporting the event or information.

Soft News: These are the reports where the journalists' opinion about the news (event/information) can be noticed either overtly or covertly. In this category, the perception of the journalist can be observed.

The above two categories have further been classified in the 16 types of ''The Hindu'' and "The Times of India'' (06 types under Hard news and 10 types under soft news). Thus the sample for the study can be illustrated as given below:

Codification of News Categorization:

To facilitate the instant & easy retrieval of a specific news item from among the 200 samples culled out from 48 issues of 'The Hindu' and ' The Times of India', it was decided to assign a numerical code for each one of the news. Accordingly the following codification system was developed and has been used throughout the study.

Every one of the news item is given a six digit numerical code.

➢ Of the six digits the FIRST digit will represent the name of the news paper as follows :

    1 – 'The Hindu';
    2 – 'The Times of India'

➢ The SECOND digit represents one of the new broad categories of the news items as stated below:

1) Hard news
2) Soft news
    o 'The Hindu' / Hard news
    o 'The Times of India' / Hard news
    o 'The Hindu' / Soft news
    o 'The Times of India' / Soft news

➢ The THIRD and FOURTH digits stand for the sub categories of hard and soft news items as detailed below :

Hard News	Soft News
Political	Editorial
Economic	Column
Crime	Books / Film Reviews
Accident	Interview
Rallies	Science & Technology

Election	Agriculture& Environment
	Art & Culture
	Health
	Women and Children
	Sports
	Editorial

Sampling Technique

The present study involves the analysis of newspapers published across a period of twelve months. The general rule in any research is to make the sample larger (within reason) to make findings reliable. Therefore, the researcher decided to take 48 issues of ''The Hindu'' and 48 issues of ''The Times of India''.

Newspaper Clippings date

Date of the issues	Month and year
1,2,3,4	July 2018
5,6,7,8	August 2018
9,10,11,12	September 2018
13,14,15,16	October 2018
17,18,19,20	November 2018
21,22,23,24	December 2018
25,26,27,28	January 2019
1,2,3,4	February 2019
5,6,7,8	March 2019
9,10,11,12	April 2019
13,14,15,16	May 2019
17,18,19,20	June 2019

The 48 issues included for analysis was distributed among the 12 months at 4 issues a month keeping the dates of publication distributed across the twelve months in a sequence at random.

Tools and Techniques used for Data Collection

Measurement of Readability

According to G. Harry McLaughlin[69] (1968) in "Proposals for British Readability Measures" states that one of the least ambiguous published definitions of readability is that given by English and English in their Dictionary of Psychological Terms. 'Readability', they say, 'is the quality of a written or printed communication that makes it easy for any given class of persons to understand its meaning, or that induces them to continue reading'. This means that readability is that combination of qualities which makes a text comprehensible or compelling or both, Variables Affecting Readability

- Typographical,
- Motivational and
- Logical

Visibility

Typographical determinants of readability include visibility and legibility. the most important variable for visibility is the size of the type.

The findings may be summarized in a little rule-of-thumb which I have concocted: divide a child's age by two and that gives the number of printed lines per inch with which his perceptual system can cope satisfactorily.

Legibility

A grapheme may be roughly defined as a letter or combination of letters representing a single sound. The efficiency with which graphemes and combinations of graphemes in a text can be recognized at speed is termed the text's legibility for which visibility is, of course, a prerequisite.

Flesch-Kincaid Grade Level formula, Gunning Fog Index formula of Robert Gunning 1952 (Fog Index) and Mc Laughlings (1969) SMOG Index (Simple Measure of Gobbled Gook) are some of the formulas used in this study to assess the readability of news reports in terms of grade levels. This is to compare them with Reading Ease score of texts and the measures of linguistic variables (ASL and ASW).

In order to verify the hypotheses framed in the present study appropriate data are collected using suitable tools and techniques. And the major variable being readability and several readability formulas available were considered and finally it was decided to use Flesch-Kincaid Reading Ease formula throughout this study.

In order to measure the readability of the news reports taken for analysis the following tools and techniques are used:

- Flesch-Kincaid Reading Ease formula
- Gunning-Fog Index
- SMOG index
- Text Statistics
  - Average sentence length (ASL). This is obtained by dividing the total number of words a text by the total number of words.
    Average Sentence Length (ASL) $= {}^{\text{total no of words}}/_{\text{no of sentences}}$
  - Word Length (Average Number of syllables per word) – This is obtained by dividing the total numbers syllables by the total number of words in the text.

    Average number of Syllable per Word (ASW) $= {}^{\text{Total no of Syllable}}/_{\text{Total no of words}}$

Cloze Readability Procedure

Flesch Kincaid Reading Ease Formula

The original Flesch reading ease formula was modified and the formula given below has been used here to calculate the reading ease score.

Reading Ease Score = 1.599 nosw -1.015 SL-31.517
 Where,
Nosw = Number of one-syllable words per 100 words
 SL = average sentence length in words
This formula correlates better than 0.90 with the original Flesch-Reading Ease Formula and 0.70 with 75% comprehension of 100 word samplings of the Mc Call Crabbs readings lessons.70

The Flesch reading Ease score can be interpreted the following way

Flesch-ReadingEase score	Style
90-100	Very easy
80-90	Easy
70-80	Fairly Easy
60-70	Standard
50-60	Fairly Difficult
30-50	Difficult
0-30	Very Difficult

FOG Index

In "The technique of clear writing", Gunning (1952) published his own readability formula developed for adults, the FOG Index which became popular because of the ease of use. It uses two variables (i) average sentence length and (ii) the number of words with more than two syllables for each 100 words.

Grade level = 4 (average sentence length + hard words) Where,
Hard words = number of words of more than two syllables (for each 100 words).

Here the use of 90% correct score with Mc Call Crabbs reading tests made this formula a higher grade criterion than other formulas except for Mc Laughlin's SMOG formula.

SMOG Index

The SMOG formula was published by G. Harry Mc Laughlin in 1969. He counted the number of words of more than two syllables (poly syllables count) in 30 sentences and came out with his new formula.

SMOG Grade = 3 + (Square root of poly syllable count), this formula was validated against the MC Call-Crabbs passages and used a 100 percent correct- score criteria.

Cloze Readability Procedure

Cloze readability procedure can be used for various purposes and is designed accordingly. To measure the comprehension difficulties of text material, the following is the procedure widely used:

o Step I: An unseen passage (texts not previously read by the test takers) is chosen depending on the prior knowledge and the level of knowledge/ skill to be assessed.
o Step II: In cloze readability procedure generally every nth (preferably 5th) in the word / phrase / clause / sentence is deleted from a passage and they are replaced by underlined blank spaces of a uniform length. The test takers have to fill in the blanks with appropriate words.
o Step III: The deleted words are replaced by underlined blank spaces of a uniform length.
o Step IV: The tests are photocopied
o Step V: The subjects are instructed to fill in each blank with the word they think was deleted to from that blank.
o Step VI: A response is scored correct when it exactly matches the word.
o Step VII: *The difficulty of a passage is the Mean of the subjects' percentage scores on the test (The proportion of subjects using the correct word / phrase / clause / sentence in a blank is used as a measure of the difficulty of the word / phrase / clause / sentence / deleted)71 (John R. Bormuth, (1967).

Statistical Tools and Techniques used for data analysis

The following statistical tools and techniques were used for data analysis:

Mean
Standard deviation

F-test
T-test
Univariate analysis of variance (ANOVA)

The mean scores and the standard deviation were used at the outset for simple comparison and interpretation of the data. At the next level for the comparison of the measures obtained for significance 'F' test and 'T' tests were used. Univariate analysis   of variance was done to study the trend and type of relationship between the variables.

The variables affecting readability fall under two categories viz., text variables and design variables. As for the purpose of the study only text variables are considered.

Limitations of the Study

Most of the Readability Formulas taken into account only the text variables and they don't pay attention to the other variables such as readers variables, idea density etc.,

The University of Minnesota Library cautions those who use readability formulae that they "do not account for all variables" and observes that such formulae are limited because Giles (1990, p. 131-138):

o They measure only prose. They do not account for charts or graphs and their relation to the text.
o Readability formulae do not account for grammar ot other mechanical errors that might interfere with readability.
o Formulae that measure readability by sentence lenghth can mislead writers into deleting words that are needed for clarity and connecting ideas to one other.
o Readability formulae can't account for organizational problems.
o Readability formulae don't measure the effects of layout or other design elements.
o Kearl (1949, p. 348) also noted that because such formulae deal primarily with writing style, they omit potential impacts of such elements as content, format and organization, each of which is important to readability.

Major findings
The following are the major findings of the study
1. "The overall readability of the news reports of 'The Hindu' does not differ significantly from the overall readability of 'The Times of India'."
2. "The overall grade level of the news reports of 'The Hindu' differ significantly from the overall graded level of the news reports of 'The Times of India'."
3. The hard news reports of 'The Hindu' differ significantly from the hard news reports of Times of India in their average sentence length measured in terms of the average number of words.
4. The soft news reports of 'The Hindu' & 'The Times of India' differ significantly in their average sentence length (measured in terms of words contained in a sentence.
5. The hard and soft news reports of 'The Times of India' differ significantly in their average sentence length measured in terms of the average number of words contained in a sentence.
6. "The hard news and soft news reports of 'The Times of India' do not differ in their word length measured in terms of the average number syllables per word." 126
7. "The hard news and soft news reports of 'The Hindu' do not differ in their sentence length measured in terms of the number syllables per words per sentence."
8. "The readability of the hard news reports of 'The Hindu' does not differ significantly from readability of the hard news reports of 'The Times of India'"
9. "The readability of the soft news reports of TH does not differ significantly from the readability of the soft news reports of the TOI."
10. "The average grade levels of the soft news reports of 'The Hindu' and 'The Times of India' do not differ significantly"