

0.1 Traditional Readability Formulae and their creation processes

Most of the researchers who worked on early readability formulae also gave significant contributions to their field (typically Educational Psychology). So, I am also adding wikipedia pages where they exist, to give a perspective.

- (Thorndike, 1921) - describe a process of vocabulary list creation for schools. Wikipedia page of Edward Thorndike: https://en.wikipedia.org/wiki/Edward_Thorndike
- (Lively and Pressey, 1923) proposed a method to estimate the vocabulary burden of text books. Sidney L.Pressey's wikipedia page: https://en.wikipedia.org/wiki/Sidney_L._Pressey
- (Vogel and Washburne, 1928) proposed "Winnetka Formula", which perhaps for the first time considered syntactic aspects of text and validated their formula against reading test scores.
- (Waples and Tyler, 1931) studied adult reading instead of children and looked at the specific issues involved there.
 - Wikipedia page on Waples: https://en.wikipedia.org/wiki/Douglas_Waples
 - Wikipedia page on Tyler: https://en.wikipedia.org/wiki/Ralph_W._Tyler
- (Patty and Painter, 1931) proposed another approach to measure the vocabulary burden in textbooks, as an improvement over previous approaches like (Lively and Pressey, 1923).
- (Dale and Tyler, 1934) studied the factors influencing adult reading difficulty. Edgar Dale's Wikipedia page: https://en.wikipedia.org/wiki/Edgar_Dale
- (Gray and Leary, 1935) studied a wide range of factors that influence readability, with adult readers as their end-users. William Gray's wikipedia page: https://en.wikipedia.org/wiki/William_S._Gray
- (Lorge, 1939; Lorge, 1944; Thorndike and Lorge, 1944) worked on proposing simple, easy to use readability formulae to select reading passages for children.
- (Flesch, 1943; Flesch, 1948) studied the markers of readable style and proposed a readability formula which is still widely used. Wikipedia page on Rudolf Flesch: https://en.wikipedia.org/wiki/Rudolf_Flesch
- (Dale and Chall, 1948a; Dale and Chall, 1948b) also proposed another readability formula, and continued improving it. Wikipedia page on Chall: https://en.wikipedia.org/wiki/Jeanne_Chall
- (Farr et al., 1951) describes a simplification of the Flesch reading ease formula.
- (Klare, 1952) studied the measures of readability in written evaluation and also continued to work on various aspects of textual readability for another 4,5 decades. Some of his articles/books on this topic are: (Klare et al., 1955b; Klare et al., 1955a; Klare et al., 1957; Klare, 1963; Klare, 1969; Klare, 1974; Klare, 2000b; Klare, 2000a). Wikipedia page on George Klare: https://en.wikipedia.org/wiki/George_R._Klare

- (Kandel and Moles, 1958) proposed a readability formula for French, based on Flesch formula (the article referred here is written in French.)
- (Aukerman, 1965) wrote a report on the readability of secondary school literature textbooks.
- (Bormuth, 1966) proposed a new approach to readability, based on cloze-tests and studied a wide range of linguistic properties in the process. He also worked extensively on the topic.
- (Smith and Senter, 1967) described the ARI readability formula.
- (Björnsson, 1968) described the LIX readability formula for Swedish (This book is written in Swedish.)
- (Gunning, 1968) described an early version of the Gunning Fog Index.
- (McLaughlin, 1969) proposed the SMOG readability formula.
- (Jakobsen, 1971) described a Danish version for LIX. (This book is written in Danish.)
- (Caylor et al., 1973) described the FORCAST readability formula for Military communication use.
- (McLaughlin, 1974) compared some of the existing readability formulae.
- (Granowsky and Botel, 1974) described a new syntactic complexity formula.
- (Coleman and Liau, 1975) proposed another readability formula.
- (Kincaid et al., 1975) described the creation of readability formulae for Navy personnel.

0.2 Theoretical perspectives

0.3 User based experiments

0.4 corpora creation/description/analysis

- (Miller and Coleman, 1967) described a set of 36 prose passages calibrated for text complexity.

0.5 Related standardized tests and resources

- (McCall and Crabbs, 1926)
- (Taylor, 1953) - cloze test as a measure of readability.
- General Service List (West, 1953) and its enhanced versions. More details can be seen in the Wikipedia page of GSL: https://en.wikipedia.org/wiki/General_Service_List.

0.6 Feature engineering and Machine Learning

0.7 Evaluation

0.8 Applications of readability analysis

0.9 Various Theses

0.10 Readability related workshops, symposiums etc.

References

- [Aukerman1965] Robert C. Aukerman. 1965. Readability of secondary school literature textbooks: A first report. *The English Journal*, 54(6):533–540.
- [Björnsson1968] Carl-Hugo Björnsson. 1968. *Lesbarkeit durch Lix*. Pedagogiskt Centrum i Stockholm, Stockholm.
- [Bormuth1966] John R. Bormuth. 1966. Readability: A new approach. *Reading Research Quarterly*, 1(3):79–132.
- [Caylor et al.1973] John S. Caylor, Thomas G. Sticht, Lynn C. Fox, and J. Patrick Ford. 1973. Methodologies for determining reading requirements of military occupational specialties: Technical report no. 73-5. Technical report, Human Resources Research Organization, Alexandria, VA.
- [Coleman and Liau1975] Meri Coleman and T. L. Liau. 1975. A computer readability formula designed for machine scoring. *Journal of Applied Psychology*, 60:283–284.
- [Dale and Chall1948a] Edgar Dale and Jeanne S. Chall. 1948a. A formula for predicting readability. *Educational research bulletin; organ of the College of Education*, 27(1):11–28.
- [Dale and Chall1948b] Edgar Dale and Jeanne S. Chall. 1948b. A formula for predicting readability: Instructions. *Educational research bulletin; organ of the College of Education*, 27(2):37–54.
- [Dale and Tyler1934] Edgar Dale and Ralph W. Tyler. 1934. A study of the factors influencing the difficulty of reading materials for adults of limited reading ability. *The Library Quarterly*, 4:384–412.
- [Farr et al.1951] James N. Farr, James J. Jenkins, and Donald G. Paterson. 1951. Simplification of flesch reading ease formula. *Journal of applied psychology*, 35(35):333–337.
- [Flesch1943] Rudolf Franz Flesch. 1943. *Marks of Readable Style: A Study in Adult Education*. Ph.D. thesis, Columbia University.
- [Flesch1948] Rudolf Franz Flesch. 1948. A new readability yardstick. *Journal of Applied Psychology*, 32(3):221–233.
- [Granowsky and Botel1974] Alvin Granowsky and Morton Botel. 1974. Background for a new syntactic complexity formula. *The Reading Teacher*, 28 (1):31–35.
- [Gray and Leary1935] William S. Gray and Bernice E. Leary. 1935. *What makes a book readable: With special reference to adults of limited reading ability, an initial study*. The University of Chicago Press, Chicago, Illinois, USA.
- [Gunning1968] Robert Gunning. 1968. *The Technique of Clear Writing*. McGraw-Hill Book Company, New York, 2nd edition.
- [Jakobsen1971] G. Jakobsen. 1971. *Dansk LIX 70*. Landsforeningen af læsepædagoger. Dragør, København.
- [Kandel and Moles1958] L. Kandel and Moles A. Moles. 1958. Application de l’indice de flesch a la langue francaise. *Cahiers Etudes de Radio-Television*, 19:253–274.
- [Kincaid et al.1975] J. P. Kincaid, R. P. Jr. Fishburne, R. L. Rogers, and B. S. Chissom. 1975. Derivation of new readability formulas (Automated Readability Index, Fog Count and Flesch Reading Ease formula) for Navy enlisted personnel. Research Branch Report 8-75, Naval Technical Training Command, Millington, TN.

- [Klare et al.1955a] George R. Klare, James E. Mabry, and Levarl M. Gustafson. 1955a. The relationship of human interest to immediate retention and to acceptability of technical material. *The Journal of Applied Psychology*, 39(2):92–95.
- [Klare et al.1955b] George R. Klare, James E. Mabry, and Levarl M. Gustafson. 1955b. The relationship of style difficulty to immediate retention and to acceptability of technical material. *The Journal of Educational Psychology*, 46(5):287–295.
- [Klare et al.1957] George R. Klare, William H. Nichols, and Emir H. Shuford. 1957. The relationship of typographic arrangement to the learning of technical material. *The Journal of Applied Psychology*, 41(1):41–45.
- [Klare1952] George Klare. 1952. Measures of the readability of written communication: An evaluation. *The Journal of Educational Psychology*, 43(7):385–399.
- [Klare1963] George R. Klare. 1963. *The Measurement of Readability*. Iowa State University Press, Ames, Iowa.
- [Klare1969] George R. Klare. 1969. Automation of the flesch reading ease readability formula, with various options. *Reading Research Quarterly*, 4(4):550–559.
- [Klare1974] George R. Klare. 1974. Assessing readability. *Reading Research Quarterly*, 10(1):62–102.
- [Klare2000a] George R. Klare. 2000a. The measurement of readability: useful information for communicators. *ACM Journal of Computer Documentation*, 24(3):107–121. Reprint of the first chapter of Klare (1963).
- [Klare2000b] George R. Klare. 2000b. Readable computer documentation. *ACM Journal of Computer Documentation*, 24(3):148–168.
- [Lively and Pressey1923] Bertha A. Lively and Sidney L. Pressey. 1923. A method for measuring the "vocabulary burden" of textbooks. *Educational Administration and Supervision*, 9(7):389–398.
- [Lorge1939] Irving Lorge. 1939. Predicting reading difficulty of selections for children. *The Elementary English Review*, pages 229–233.
- [Lorge1944] Irving Lorge. 1944. Predicting readability. *Teachers College Record*, 45:404–419, March.
- [McCall and Crabbs1926] William Anderson McCall and Lelah Mae Crabbs. 1926. *Standard Test Lessons in Reading*. 5. Teachers College, Columbia University, Bureau of Publications.
- [McLaughlin1969] G. Harry McLaughlin. 1969. SMOG grading – a new readability formula. *Journal of Reading*, 12(8):639–646.
- [McLaughlin1974] G. Harry McLaughlin. 1974. Temptations of the flesch. *Instructional Science*, 2:367–384.
- [Miller and Coleman1967] G.R. Miller and E.B. Coleman. 1967. A set of thirty-six prose passages calibrated for complexity. *Journal of Verbal Learning and Verbal Behavior*, 6(6):851–854.
- [Patty and Painter1931] W.W. Patty and W.I. Painter. 1931. A technique for measuring the vocabulary burden of textbooks. *Journal of Educational Research*, 24:127–134.
- [Smith and Senter1967] E. A. Smith and R. J. Senter. 1967. Automated Readability Index. Technical Report AMRL-TR-66-220, Aerospace Medical Research Laboratories, Wright-Patterson Airforce Base, OH, November.
- [Taylor1953] W.L. Taylor. 1953. Cloze procedure: A new tool for measuring readability. *Journalism Quarterly*, 30:415–433.
- [Thorndike and Lorge1944] Edward L. Thorndike and Irving Lorge. 1944. *The Teacher's Word Book of 30,000 Words*. Bureau of Publications, Teachers College, Columbia University, New York.
- [Thorndike1921] Edward L. Thorndike. 1921. Word knowledge in the elementary school. *Teachers College Record*, 28(5):334–370.
- [Vogel and Washburne1928] Mabel Vogel and Carleton Washburne. 1928. An objective method of determining grade placement of children's reading material. *The Elementary School Journal*, 28:373–381.

[Waples and Tyler1931] Douglas Waples and Ralph W. Tyler. 1931. *What People Want to Read About: A Study of Group Interests and a Survey of Problems in Adult Reading*. American Library Association and University of Chicago Press.

[West1953] Michael West. 1953. *A General Service List of English Words*. Longmans, London.