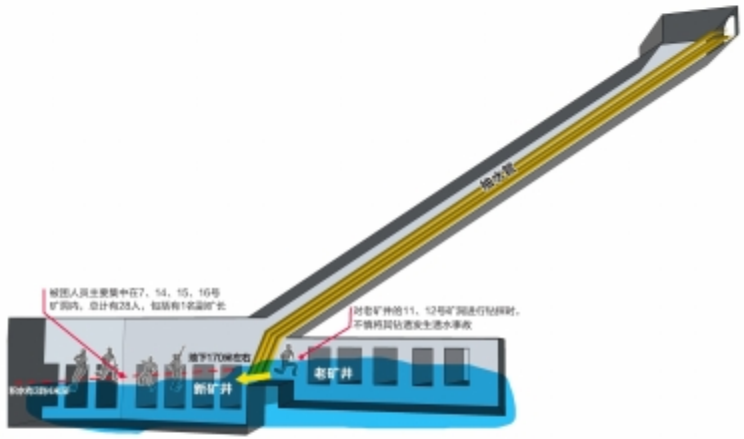
Measurement of Readability

主讲教师:王积龙

11月21日，救援人员在紧张救援。当日11时许，四川省内江市威远县八田煤矿发生一起透水事故。新华社记者江宏景摄新华社记者江宏景摄



威远县小河镇八田煤矿透水事故示意图

新华网快讯：１３时２０分，记者在威远煤矿救援现场看到最后一名矿工被救升井，２９名矿工全部获救。他们没有外伤但身体虚弱，被送到威远县人民医院和内江市第二人民医院接受检查治疗。

- 四川内江一煤矿发生透水事故
- 四川威远煤矿透水事故：29人全部幸存

近代哲学的认识论转向使知识或认识如何可能成为形而上学的中心问题。

英国经验论者以典型的科学方式处理知识如何可能的问题，而在这同时，笛卡尔以让人耳目一新的近代怀疑论为知识找到了一条来自内省的根据。但心、物二元对立这一主观与客观的形而下分离，或者说主观形式在此仅仅被当作科学可及的一"物"受到打量所带来的困扰表明，必须向前推进一步。

康德的革命以诉诸"先验"主观形式的方式解决了"科学"之知在"先验"与"经验"对立处的有效性问题。但形而上学知识却并不在"先天"的认识能力所可以达到的范围之内。"理性"试图在"现象世界"之外认识"自在之物"，必然陷于"二律背反"。康德的先验唯心主义终于只是在实证经验可及的科学"现象世界"之内坚持了自己的立场。而在所谓"超验"的"理念"领域，康德放弃了先天主观形式的有效性。

在自然科学的兴起作为主导时代特征的时代，为"科学"这一主观形式寻找合法性基础的工作的确已变

是可以在经验处证实的命题。但所谓"经验证实"在"命题"这里，仍然不能不是"同义反复"。"经验"得以证实的"命题"，总是在"经验"处具有逻辑自洽性，因而是"同义反复"的"命题"。

奎因在《经验主义的两个教条》中指出分析命题与综合命题无实质区别。他的理由是所有的命题都会随着经验领域的扩展发生变化。但感觉经验并没有提供"确定性"。我们所经验到的具有"确定性"的"经验"总是某种逻辑经验，亦即总是暗含某种逻辑命题或以某种逻辑命题为背景的"经验"。因此，逻辑实证主义者沿袭休谟并精致地加以区分的分析命题与综合命题，说到底不外是以不同方式表达其"同义反复"的命题。所有"命题"都奠基于"A = A"这一基本逻辑形式。古希腊哲学所标举的"思想"与"存在"的同一性，在当代知识体系谨守不渝的逻辑形式(A = A)这里，找到了自己从未中断的传承世系。

现代哲学的语言学转向正如此前发生的种种转向，在根本上并未转到古希腊哲学所开创的形而上学传统本来意指的方向之外。科学语言和日常语言的确只有在

Readability

- Readability is defined as reading ease, especially as it results from a writing style. Extensive research has shown that easy-reading text improves comprehension, retention, reading speed, and reading persistence. Ease-of-reading is the result of the interaction between the text and the reader.
- In the reader, those features affecting readability are 1. prior knowledge, 2. reading skill, 3. interest, and 4. motivation. In the text, those features are 1. content, 2. style, 3. design, and 4. structure. The design can include the medium, layout, illustrations, reading and navigation aids, typeface, and color. Correct use of type size, line spacing, column width, text-color-background contrast and white space make text easy to read.

Readability Formula

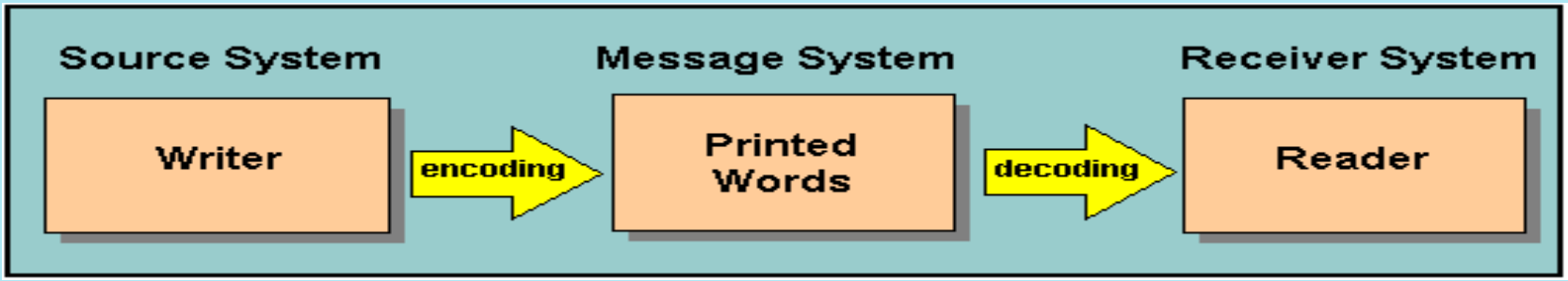
- **Readability formulas**, are formula for evaluating the readability of text, usually by counting syllables, words, and sentences. Readability tests are often used as an alternative to conducting an actual statistical survey of human readers of the subject text (a readability survey). Word processing applications often have readability tests in-built, which can be deployed on documents in-editing.

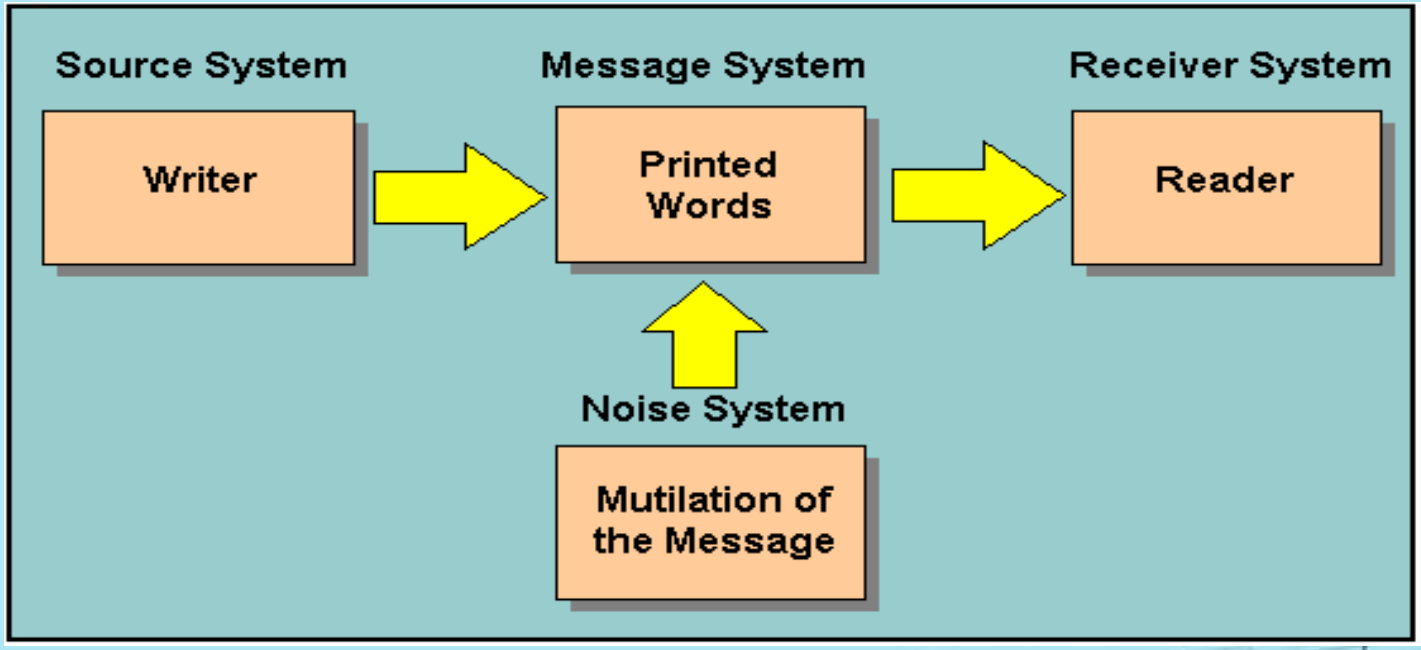
- The application of a useful readability test protocol will give a rough indication of a work's readability, with accuracy increasing when finding the average readability of a large number of works. The tests generate a score based on characteristics such as statistical average word length (which is a used as a proxy for semantic difficulty) and sentence length (as a proxy for syntactic complexity) of the work.

- Some readability formulas refer to a list of words graded for difficulty. These formulas attempt to overcome the fact that some words, like "television", are well known to younger children, but have many syllables. In practice, however, the utility of simple word and sentence length measures make them more popular for readability formulas.[citation needed] Scores are compared with scales based on judged linguistic difficulty or reading grade level. Many readability formulas measure word length in syllables rather than letters, but only SMOG has a computerized readability program incorporating an accurate syllable counter.

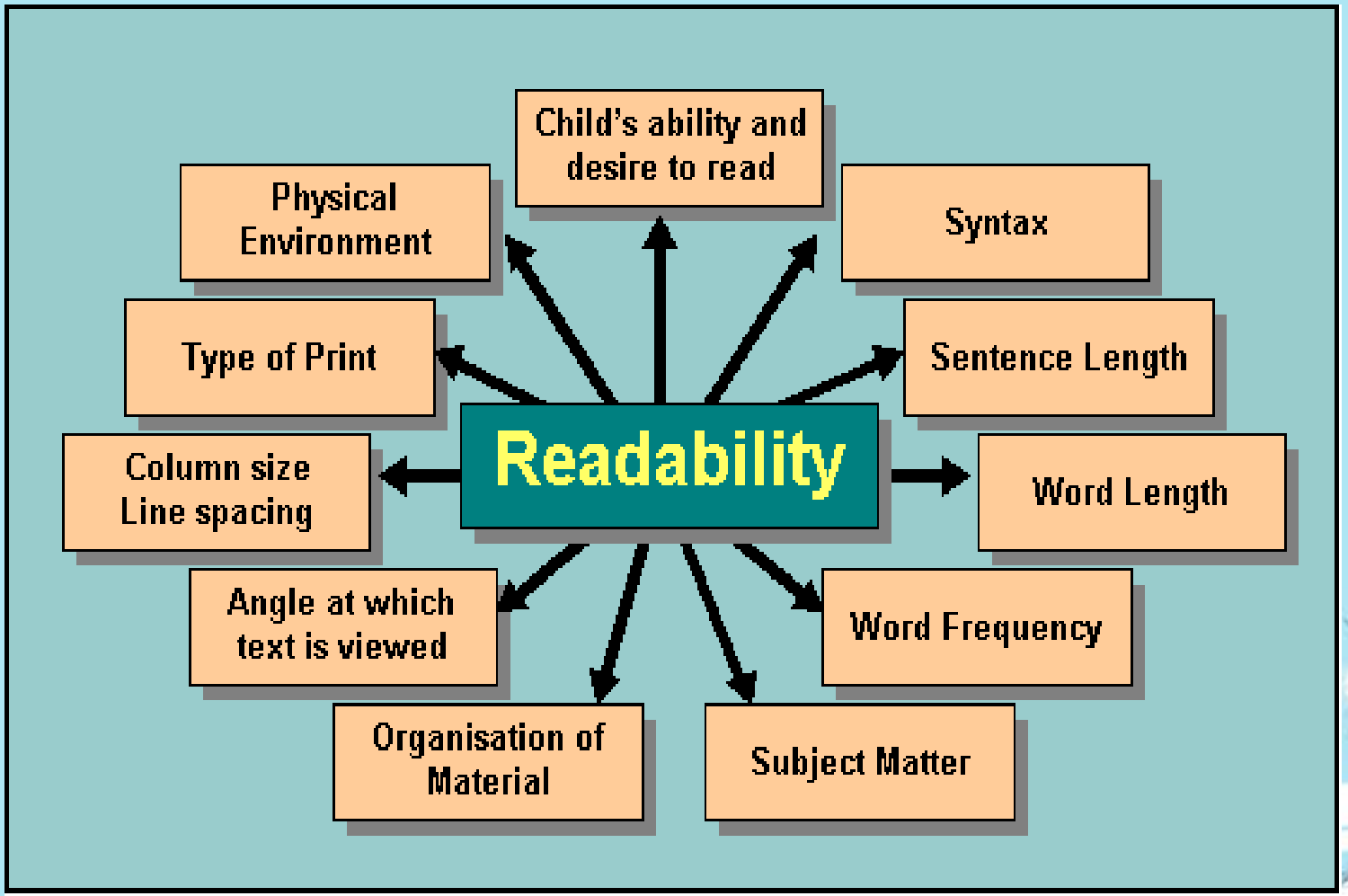
- Since readability formulas do not take the meanings of words into account, they are not considered definitive measures of readability.

Readability Formula





Readability Formula



Flesch Formula

Flesch Reading Ease

Flesch Reading Ease Formula is considered as one of the oldest and most accurate readability formulas. Rudolph Flesch, an author, writing consultant, and a supporter of the Plain English Movement, developed this formula in 1948. Raised in Austria, Rudolph Flesch studied law and earned a Ph.D. in English from the Columbia University. Flesch, through his writings and speeches, advocated a return to phonics. In his article, A New Readability Yardstick, published in the Journal of Applied Psychology in 1948, Flesch proposed the Flesch Reading Ease Readability Formula.

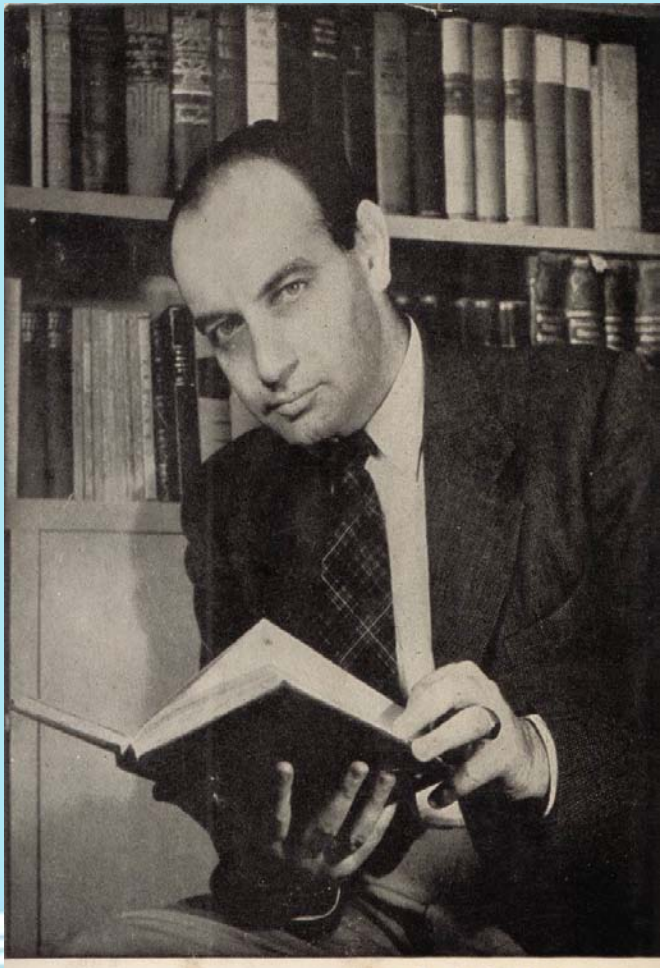


Photo by Annie M. Graf

Flesch Formula

- **Flesch Reading Ease**

The Flesch Reading Ease Readability Formula

The specific mathematical formula is:

RE = 206.835 – (1.015 x ASL) – (84.6 x ASW)

RE = Readability Ease

ASL = Average Sentence Length (i.e., the number of words divided by the number of sentences)

ASW = Average number of syllables per word (i.e., the number of syllables divided by the number of words)

The output, i.e., RE is a number ranging from 0 to 100. The higher the number, the easier the text is to read.

• Scores between 90.0 and 100.0 are considered easily understandable by an average 5th grader.

• Scores between 60.0 and 70.0 are considered easily understood by 8th and 9th graders.

• Scores between 0.0 and 30.0 are considered easily understood by college graduates.

Flesch Formula

- **Flesch Reading Ease**

The Flesch Reading Ease Readability Formula
If we were to draw a conclusion from the Flesch Reading Ease Formula, then the best text should contain shorter sentences and words. The score between 60 and 70 is largely considered acceptable. The following table is also helpful to assess the ease of readability in a document:

RE = 206.835 – (1.015 x ASL) – (84.6 x ASW)

90-100 : Very Easy	Grade 5
80-89 : Easy	Grade 6
70-79 : Fairly Easy	Grade 7
60-69 : Standard	Grade 8 & Grade 9
50-59 : Fairly Difficult	Rank 10 to Rank 12
30-49 : Difficult	Undergraduate
0-29 : Very Confusing	Graduate

Flesch Formula

- **Flesch Reading Ease**

Style	Flesch Reading Ease Score	Average Sentence Length in Words	Average No. of Syll. Per 100 Words	Type of Magazine	Estimated School Grade Completed	Estimated Percent of U.S. Adults
Very Easy	90 to 100	8 or less	123 or less	Comics	4th grade	93
Easy	80 to 90	11	131	Pulp fiction	5th grade	91
Fairly Easy	70 to 80	14	139	Slick fiction	6th grade	88
Standard	60 to 70	17	147	Digests	7th or 8th grades	83
Fairly Difficult	50 to 60	21	155	Quality	Some high school	54
Difficult	30 to 50	25	167	Academic	High school or some college	33
Very Difficult	0 to 30	29 or more	192 or more	Scientific	College	4.5

Flesch Formula

- **Flesch Reading Ease**

Formula B for predicting "**human interest**":
HI = 3.635 pw + .314ps.
Scores computed by this formula, too, have a range from 0 to 100.
A score of 100 has the same meaning as in Formula A. It indicates reading matter with enough human interest to suit the reading skills and habits of a barely "functionally literate" person.
A score of 0, however, means here simply that the passage contains neither persnal words! nor! personal senences!

Flesch Formula

- **Flesch Reading Ease**

Here's the breakdown,

Score Notes

90.0–100.0 easily understandable by an average 11-year old student [citation needed]

60–70 easily understandable by 13- to 15-year old students [citation needed]

0–30 best understood by college graduates [citation needed]

Reader's Digest magazine has a readability index of about 65, Time magazine scores about 52, an average year 7 student's (eleven years) written assignment has a readability test of 60-70 (and a reading grade level of 6-7) and the Harvard Law Review has a general readability score in the low 30s. The highest (easiest) readability score possible is 121 (every sentence consisting of only one-syllable words); theoretically there is no lower bound on the score -- this sentence, for example, taken as a reading passage unto itself, has a readability score of ~21.9. This paragraph has a readability score of ~53.93.

Many government agencies require documents or forms to meet specific readability levels.[citation needed] The U.S. Department of Defense uses the Reading Ease test as the standard test of readability for its documents and forms.[citation needed]

Use of this scale is so ubiquitous that it is bundled with popular word processing programs and services such as KWord, Lotus WordPro, Microsoft Word, and Google Docs.

Flesch Formula

- **Flesch Reading Ease**

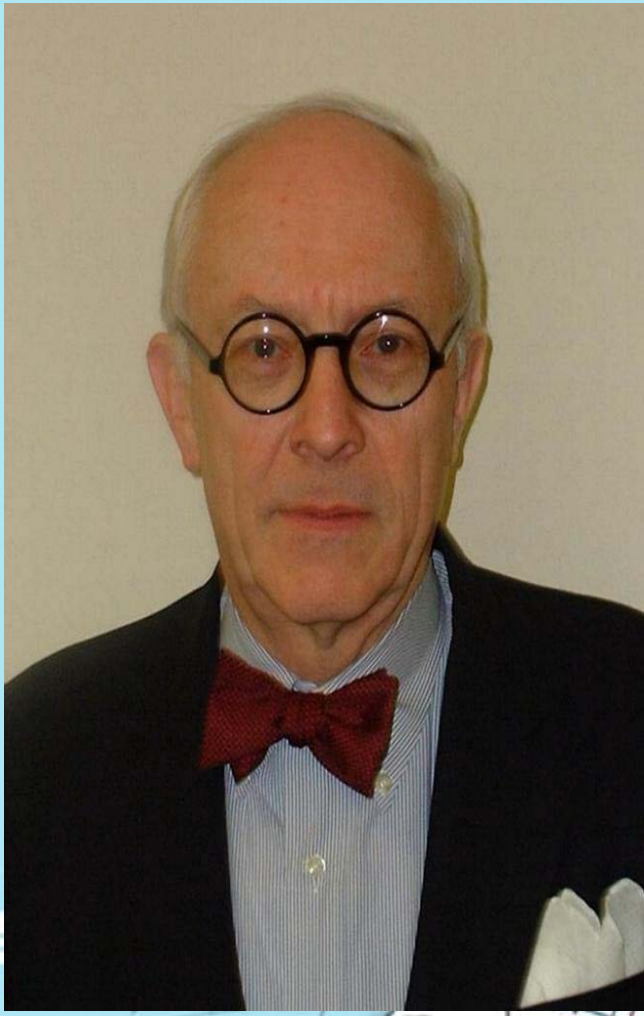
The Flesch Reading Ease Readability Formula
If we were to draw a conclusion from the Flesch Reading Ease Formula, then the best text should contain shorter sentences and words. The score between 60 and 70 is largely considered acceptable. The following table is also helpful to assess the ease of readability in a document:

RE = 206.835 – (1.015 x ASL) – (84.6 x ASW)

90-100 : Very Easy	Grade 5
80-89 : Easy	Grade 6
70-79 : Fairly Easy	Grade 7
60-69 : Standard	Grade 8 & Grade 9
50-59 : Fairly Difficult	Rank 10 to Rank 12
30-49 : Difficult	Undergraduate
0-29 : Very Confusing	Graduate

Flesch Formula

The Gunning's Fog Index (or FOG) Readability Formula
The Gunning Fog Index Readability Formula, or simply called FOG Index, is attributed to American textbook publisher, Robert Gunning, who was a graduate from Ohio State University. Gunning observed that most high school graduates were unable to read. Much of this reading problem was a writing problem. His opinion was that newspapers and business documents were full of "fog" and unnecessary complexity.
Gunning realized the problem quite early and became the first to take the new readability research into the workplace. Gunning founded the first consulting firm specializing in readability in 1944. He spent the next few years testing and working with more than 60 large city daily newspapers and popular magazines, helping writers and editors write to their audience.

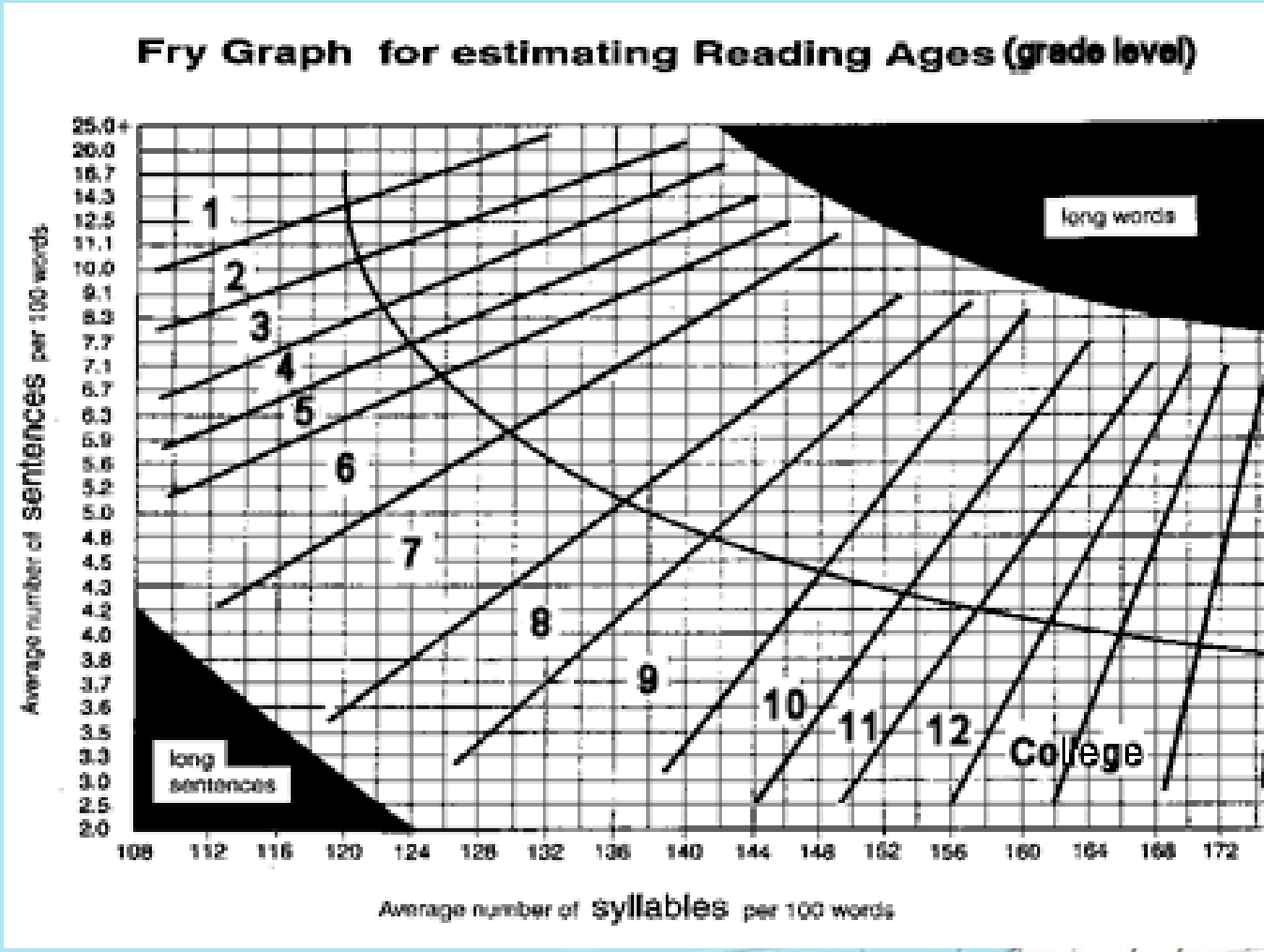


Flesch Formula

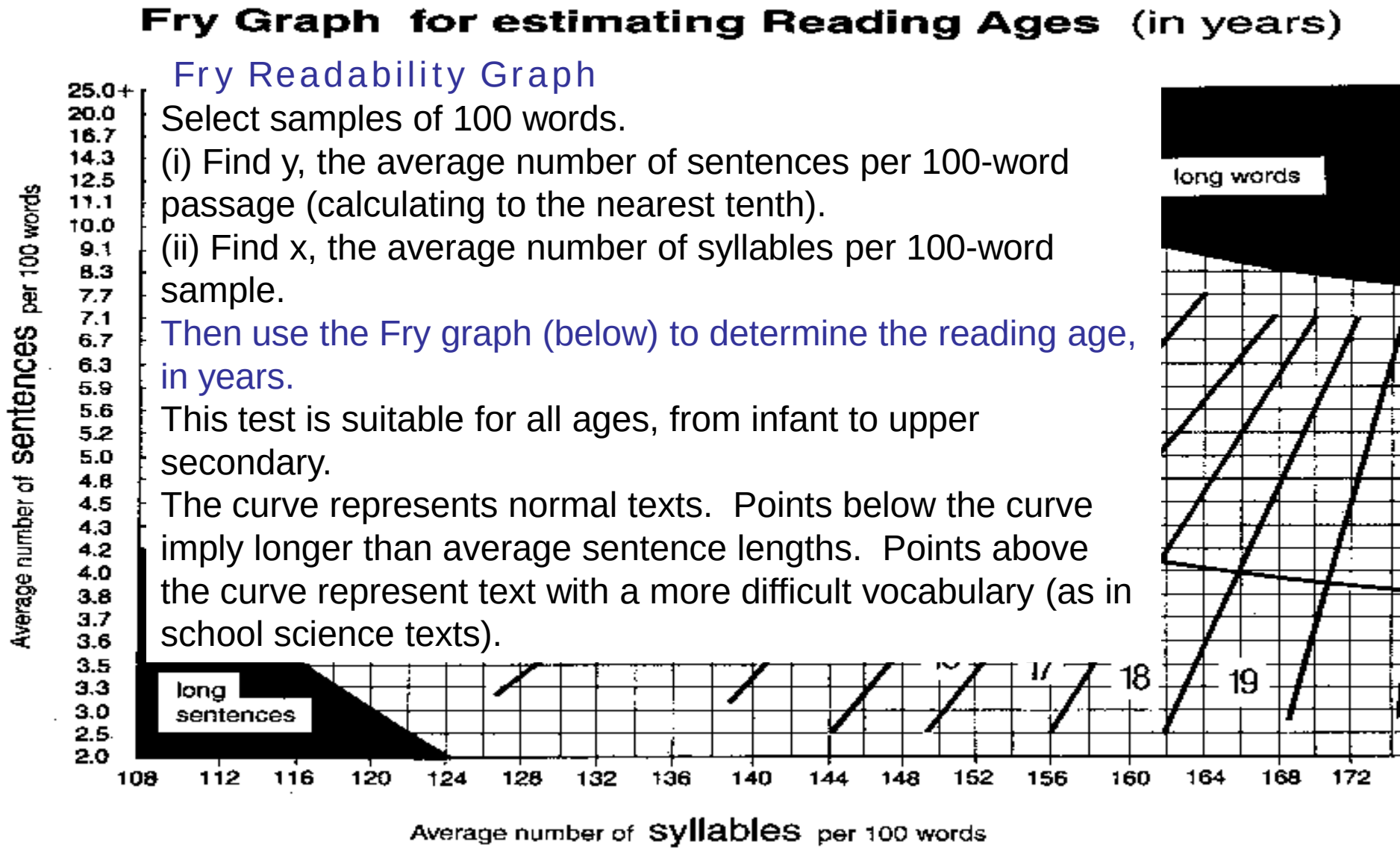
The Gunning's Fog Index (or FOG) Readability Formula

Step 1: Take a sample passage of at least 100-words and count the number of exact words and sentences.
Step 2: Divide the total number of words in the sample by the number of sentences to arrive at the Average Sentence Length (ASL).
Step 3: Count the number of words of three or more syllables that are NOT (i) proper nouns, (ii) combinations of easy words or hyphenated words, or (iii) two-syllable verbs made into three with -es and -ed endings.
Step 4: Divide this number by the number or words in the sample passage. For example, 25 long words divided by 100 words gives you 25 Percent Hard Words (PHW).
Step 5: Add the ASL from Step 2 and the PHW from Step 4.
Step 6: Multiply the result by 0.4.
The mathematical formula is:
Grade Level = 0.4 (ASL + PHW)
where,
ASL = Average Sentence Length (i.e., number of words divided by the number of sentences)
PHW = Percentage of Hard Words
The underlying message of The Gunning Fog Index formula is that short sentences written in Plain English achieve a better score than long sentences written in complicated language.
The ideal score for readability with the Fog index is 7 or 8. Anything above 12 is too hard for most people to read. For instance, The Bible, Shakespeare and Mark Twain have Fog Indexes of around 6. The leading magazines, like Time, Newsweek, and the Wall Street Journal average around 11.

Fry Graph


Fry Graph for estimating Reading Ages (grade level)
Average number of sentences per 100 words
25.0+
20.0
16.7
14.3
12.5
11.1
10.0
9.1
8.3
7.7
7.1
6.7
6.3
5.9
5.6
5.2
5.0
4.8
4.5
4.3
4.2
4.0
3.8
3.7
3.6
3.5
3.3
3.0
2.5
2.0
1
2
3
4
5
6
7
8
9
10
11
12
College
long words
long sentences
108
112
116
120
124
128
132
136
140
144
148
152
156
160
164
168
172
Average number of syllables per 100 words

Fry Graph



Fry Readability Graph

Select samples of 100 words.

(i) Find y, the average number of sentences per 100-word passage (calculating to the nearest tenth).

(ii) Find x, the average number of syllables per 100-word sample.

Then use the Fry graph (below) to determine the reading age, in years.

This test is suitable for all ages, from infant to upper secondary.

The curve represents normal texts. Points below the curve imply longer than average sentence lengths. Points above the curve represent text with a more difficult vocabulary (as in school science texts).

Methods of assessing Reading Age

- Subjective assessment has been shown to be inaccurate, with teachers (perhaps because of their reading competence and familiarity with the subject) usually under-estimating the difficulty of the text (by up to 8 years).
- There are four main methods

of objective assessment:

Methods of assessing Reading Age

- A. Question and answer technique
- Pupils of different ages are given the text to read. They are then questioned to gauge the level of comprehension and hence determine the reading age. This is usually unrealistic for practising teachers.

- B. Sentence completion (the 'cloze' technique)
- Sentences are taken from the text and every nth word is deleted. Often, n=5. These sentence completion exercises are then given to the pupils to test comprehension and gauge the reading age. Graham [7] and Mobley [8] have given details of how cloze tests can be applied to science texts. This method is also time-consuming.

Methods of assessing Reading Age

- C. Comparison of text with a standard word list
- The percentage of words not included in the Dale word list is determined and the reading age calculated from this. Well-known examples are the Dale-Chall [9] and Spache [10] tests. Again, this method is tedious.
- D. Calculations involving the sentence length and number of syllables
- Objective measures of readability are now widely used. They are formulae (or graphs) which are based on an enormous amount of research evidence.

Flesch Formula

- **Example** *Your Baby*

Sample I:

A happy, useful life--tthat's what you want for your baby, isn't it? And because a healthy mind and body are so necessary to happiness and long life, you must do all you can to get your baby off to a good start.
There is much you can do while he is still a baby to lay the foundation for good health and good health habits.
Many things affect your baby's health. One was the state of your own health during pregnancy, and the special care your doctor gave you before the baby was born. Other things important to your child's health are food, clothes, baths, sleep, and habit training. A baby needs a clean, happy place to live, and he must be kept from having any sickness that can be prevented.

Flesch Formula

- **Example** *Your Baby*

Sample 2:

Diphtheria used to kill many babies. Today no child need die of diphtheria. It is one of the diseases for which we have very good treatment and almost sure prevention. But your baby will not be safe from this disease unless he has been protected by immunization.

The way to protect your baby is simple. Physicians usually give injections of three doses of toxoid, three to four weeks apart, generally beginning when a baby is about six months old. Your doctor will tell you that your baby should have this protection before his first birthday.

Six months after the last injection of toxoid, the physician may test your baby to see if another dose of toxoid is necessary. Before the child enters school an extra shot of toxoid is often given.

Flesch Formula

- **Example** *Your Baby*

Sample 3:

The germs that cause tuberculosis can enter the baby's body through his mouth or be breathed in through his nose. These germs come to him on spray or moisture which the person with active tuberculosis breathes or coughs out. Germ-filled spray from the mouth or nose may light on the baby's food, his dishes, his toys. The baby's hands may carry germs from soiled objects to his mouth. Kissing is one way of spreading TB as well as other germs.

Tuberculosis of the bones or joints or of certain organs of the body besides the lungs can come to the bottle-fed baby in milk which has not been ?pasteurized or boiled.

Processing Data

- Effect of Ads context

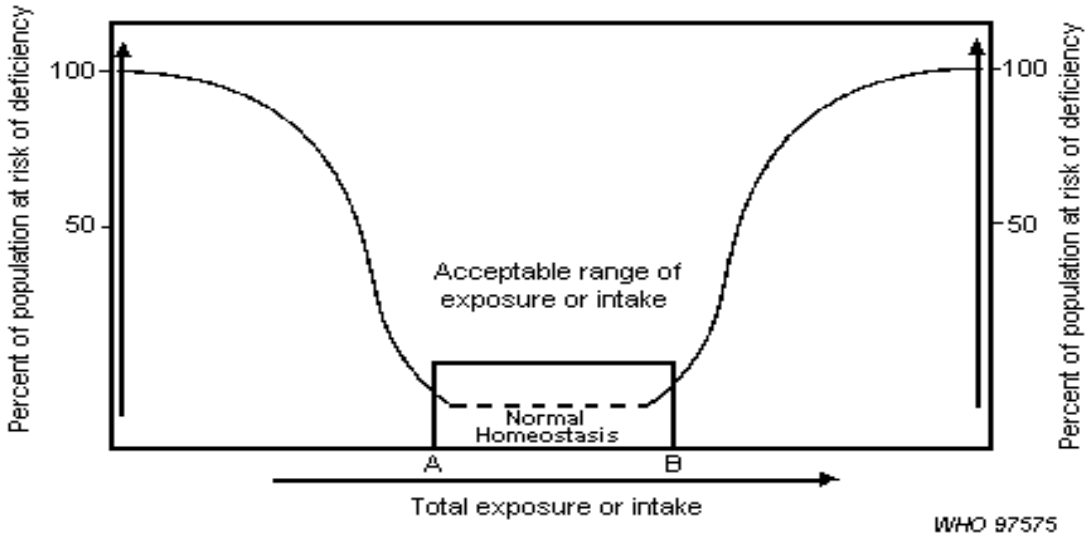


Fig. 1. Percent of population subjected to deficiency and toxicity effects according to exposure/intake. As intake drops below A risk for deficiency increases; at extremely low exposures or intakes all subjects will manifest deficiency. As exposure or intakes increase beyond B a progressively larger proportion of subjects will exhibit effects of toxicity.

- New Approaches to Readability