

# Applications of readability findings in teaching languages

Nguyen Thi Nhu Diep\* and Lam Thanh Nam  
Vietnam Aviation Academy, Vietnam

## ABSTRACT

*Readability or text readability, particularly in English, has been studied since the end of the 19th century with hundreds of thousands of published works. The research results on this issue are not only quite rich in resources but also universally applicable in many fields. Currently, it is an interesting problem and still being continuously studied by researchers, not only for English but also for many other languages. However, in Vietnam, the study of readability is still limited research to find the solutions to meet the social needs. Based on the findings and applicable results of readability studies, the paper, initially, presents an overview about readability and the linguistic factors having influence on readability. On the view from readability and applications, the paper investigates and analyses the linguistic factors on 3 main levels: Words, Sentences, and Texts by the illustrative case studies in English language, or Vietnamese language. On this basis, the paper suggests the practical applications of readability in teaching languages in general, particularly in the context of teaching English in Vietnam currently, to meet the demand for users and Vietnamese community in the era of technology.*

**Keywords:** readability, linguistic factors, applications, teaching languages

## 1. INTRODUCTION

Readability or text readability, especially in English, has been studied since the end of the 19th century with hundreds of thousands of published works and universally applicable in many fields. Currently, it is an interesting problem and still being continuously studied by researchers for English and many other languages. However, in Vietnam, the study of readability is still limited research to find the solutions to meet the social needs. Based on the findings of readability studies, the paper presents an overview about readability and the linguistic factors having influence on readability and suggests some practical applications of readability in teaching languages in general, English, and Vietnamese to meet the demand for users and Vietnamese community in this era of technology.

On the view from readability and applications, the paper investigates and analyzes the linguistic factors on 3 main levels: Words, Sentences, and Texts by the illustrative case studies in English and Vietnamese. With the paper's scope, this study is conducted by the following research questions: (1) what are the factors having influence on readability? (2) How are the linguistic factors

influencing readability applied in real-life scenarios? and (3) what can practical applications of readability be used in teaching languages in the current context of Vietnam?

## 2. LITERATURE REVIEW

### 2.1. Readability

Dale and Chall in 1949 [1] gave a common definition as follow: "the sum total (including the interactions) of all those elements within a given piece of printed material that affect the success of a group of readers have with it. The success is the extent to which they understand it at an optimal speed and find it interesting." In 2008, Fukun Xing, Dongyuan Cheng, & Jianzhong Pu presented: "readability is how easily written materials can be read and understood. Readability depends on many factors, including (a) the average length of sentence in a passage, (b) the number of new words a passage contains, (c) the grammatical complexity of the language used" [2].

Readability and Comprehensibility are closely related, but they are intrinsically different. Two concepts are not considered as synonyms because there is a remarkable distinction between them in

---

Corresponding author: Nguyễn Thị Như Diệp  
Email: [diepntn@vaa.edu.vn](mailto:diepntn@vaa.edu.vn)

many readability studies. Wray and Dahlia in 2013 [3] distinguish that readability is as a characteristic of the text itself and comprehensibility is understandability as an indication of the readers' ability to make meaning of the text. While the readability features, making a reading text easy, are content, style, design, and organization, the reader features, making a reading text easy, are prior knowledge, reading skill, interest, and motivation as Dubay [4] presents. Having the same view, Chiang et al. [5] and Jones [6] state that the readability level of a text is an indicator of the textual difficulty level of the text, it is fixed for a given text and is not varied by reader characteristics. Comprehensibility is affected by reader characteristics, such as the reader's background, prior knowledge, interest, and general reading ability. Thus, the subject of readability is texts, and the subject of comprehensibility is readers.

## 2.2. Factors Affecting Readability

Gray and Leary's in 1935 [7] identified 288 elements affecting English text readability, and these elements were classified into four main categories: (i) format or mechanical features, (ii) general features of organization, (iii) style of expression and presentation, and (iv) content. Within the scope of their study, they have identified 82 language elements that function as the "potential elements" affecting text readability by investigating the linguistic elements of style of expression and presentation alone. These elements are classified under three different units, namely word, sentence, and paragraph/passage.

Examining the same "structural elements" as Gray and Leary, Lorge in 1939 [8] examined the word frequency and added an additional variable, "a weighted index of word difficulty", he also believed that the "hard word" and the part of speech "prepositions" played as the key roles to measure syntactic complexity in English. He suggested the readability formula which adjusts weights and uses various combinations of two variables, such as (i) prepositional phrases and different hard words, (ii) average sentence length and different hard words, and (iii) the number of prepositional phrases and average sentence length.

Like many of the predecessors who tend to include two sorts of measurements of difficulty: one for vocabulary and the other for sentence structure.

Rudolf Flesch in 1948 [9] developed his original formula in his 1943 PhD dissertation, entitled "Marks of Readable Style: A Study in Adult Education", and completed his studies "The Art of Plain Talk" in 1946 and "A New Readability Yardstick" in 1948 to build the formulas to measure the readability of written materials in English. Flesch formulas have been widely used as ways of assessing textual difficulty; especially, the Reading Ease formula. This formula is counted by these two factors: *wl* stands for word length (the number of syllables per 100 words), and *sl* stands for sentence length (the average sentence length). In more details, a text can be determined the grade level automatically by the surface linguistic elements: (i) the number of words; (ii) the number of characters; (iii) the number of paragraphs; (iv) the number of sentences; (v) the average of sentences per paragraph; (vi) the average of words per sentence; and (vii) the average of characters per word. The Reading Ease formula was adapted for use by the United States Military, using the same factors with different weights in 1975, and this formula can be found to this day as a tool in the most popular word processing program in the world, Microsoft Word.

Thus, surveying, investigating, and analyzing readability focuses on linguistic elements on three main levels: words-sentences-texts.

## 3. METHODOLOGY

To study applications of readability in teaching languages, the paper uses the descriptive method and contractive method to describe the influence of the linguistic elements and factors on Vietnamese readability, compared with English, as well as states the paper findings.

The case study method of the paper is the illustrative case studies, the paper uses 3 diagrams texts as the illustrated examples to compare the surface linguistic elements affecting text readability between Vietnamese and English. In addition, Investigating and analyzing the linguistic factors on 3 main levels: Words, Sentences, and Texts is also to make clear for the illustrative case studies in the paper.

## 4. RESULTS AND DISCUSSIONS

It can be seen that readability is investigated and analyzed by the linguistic factors on 3 main levels: Words, Sentences, and Texts as mentioned in part II, Literature Review. These linguistic elements can

make a text more complicated or less difficult to understand. The writer needs to understand the text's level in order to encode his or her texts so that readers can decode the information easily, efficiently, and accurately, based on their proficiency levels. For example: compiling textbooks; creating exam questions for each proficiency level A1, A2, B1, B2, C1, C2; writing definitions in explanatory dictionaries, instructional materials, articles, contracts, laws, etc.

Within the scope of the study and based on the readability findings, the paper presents linguistic factors on three main levels: words, sentences, and texts from the illustrative case studies in English and Vietnamese. On this basis, the paper suggests some practical applications of readability in teaching languages in general, particularly in the context of teaching English in Vietnam currently.

At word level, language vocabularies are generally very large and it is thus reasonable to further divide

words into subclasses to make the word-formation rules and those of their usage more comprehensible. There are several ways to do so. For example, words can be further divided in terms of (1) their meanings; namely some words convey one meaning while others are polysemantic, in terms of (2) their origin, where they can be classified into cognates and borrowed words, (3) according to the frequency of their usage, where common, everyday words are used more often than words of slang, dialectal expressions, technical terms, and others. Words can also be divided (4) based on their word-forms into monosyllabic and polysyllabic words, or else into single and compound words, and nonetheless (5) according to their first letter, as in dictionaries [10].

In Vietnamese, however, there is another crucial way of word classification, which is based on words' lexical meanings together with their grammatical functions. It is called lexico-grammatical category. This can be seen in the Figure 1 [10]:

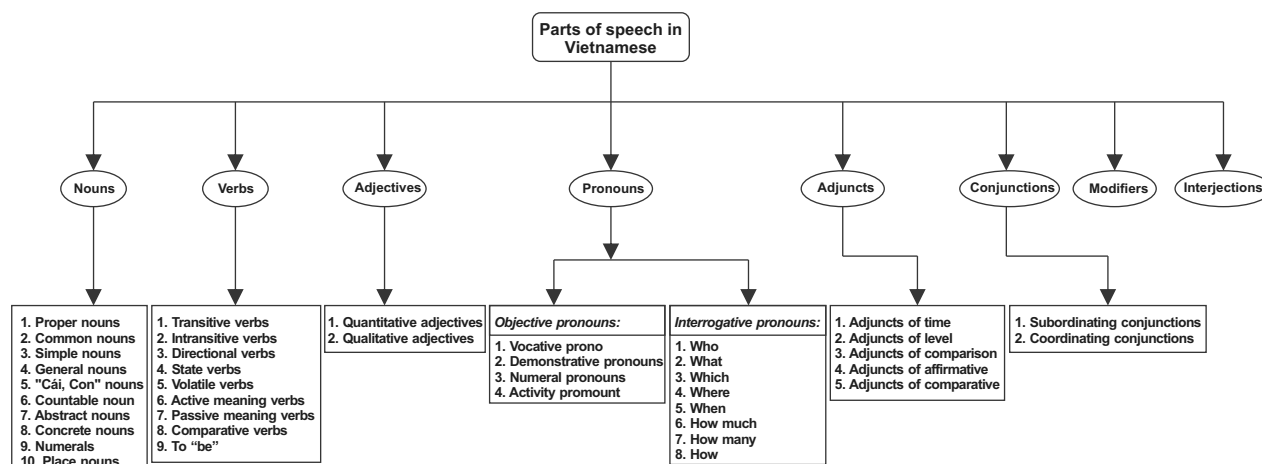


Figure 1. Word level- Part of Speech in Vietnamese language

In order to make the classification more effective and useful in forming sentences, lexical words and form words are divided further into two groups. Lexical words are categorized into nouns, verbs and adjectives, whereas form words are classified into adjuncts and conjunctions. In addition to these categories, we also make the use of pronouns, while modifiers, and interjections are the two categories that belong to both lexical words or form words and differ from the category of pronouns. To sum up, parts of speech in Vietnamese are categorized into eight main groups, of which former six groups are subdivided as above. And in Vietnamese, "in word level", part of speech is one of the linguistic factors having influence on Vietnamese readability.

In "sentence level", ambiguous sentences are always linguistic factors influencing text readability both Vietnamese and English because this type of sentence makes texts more difficult to comprehend. The example in the following diagram can be seen as an illustrative case study [11].

The ambiguity in sentence: "Jack saw the man with his glasses" arises from the connection of the prepositional phrase "with his glasses" with "Jack saw the man" – understood as "with his glasses, Jack saw the man: Jack is wearing glasses and looking at the man," or alternatively, with "the man" – understood as "Jack saw the man who is wearing glasses: Jack is looking at the man (who is) wearing glasses."

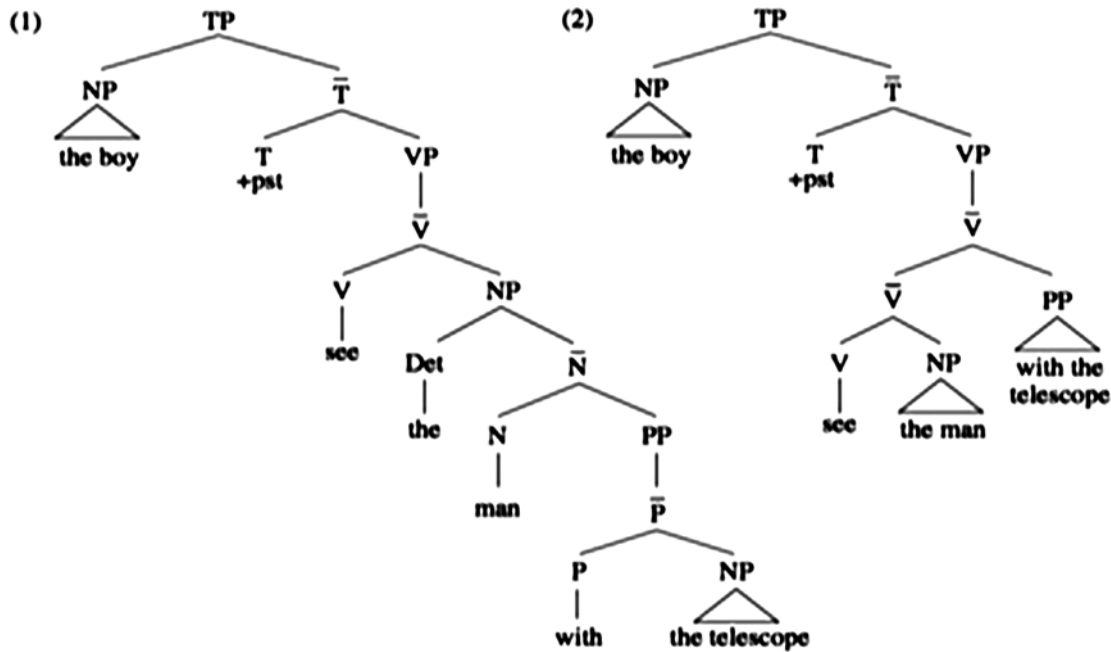


Figure 2. Sentence level- Ambiguous sentences in English language

In addition, the factors influencing readability are investigated in terms of text level. Diagram 3 [12] can be seen as an illustrative case study.

The text [12] is measured by Coh-Metrix Readability. The results show: (1) Narrativity: 17%; (2): Syntactic Simplicity: 7%; (3) Word Concreteness: 38%; (4) Referential Cohesion: 81%; and (5) Deep Cohesion: 58%. These linguistic factors are measured based on: (1) Narrativity: whether a text tells a story with characters, events, places and things that are familiar to the reader; (2) Syntactic Simplicity: the use of simple sentences structure that are easy to understand; (3) Word Concreteness: the use of words that can

be easily imagined; (4) Referential Cohesion: the presence of overlapping ideas and concept in a text; and (5) Deep Cohesion: the presence of explicit causal relationships. And the text readability as follows: “This text is low in narrativity which indicates that it is less story-like and may have less familiar words. Less story-like texts are usually harder to comprehensive. It is low in syntactic simplicity which means the sentences may have more clause and more words before the main verb. Complex syntax is harder to process. This text is high in both referential and deep cohesion, which may scaffold the reader, particular if the content is challenging.”

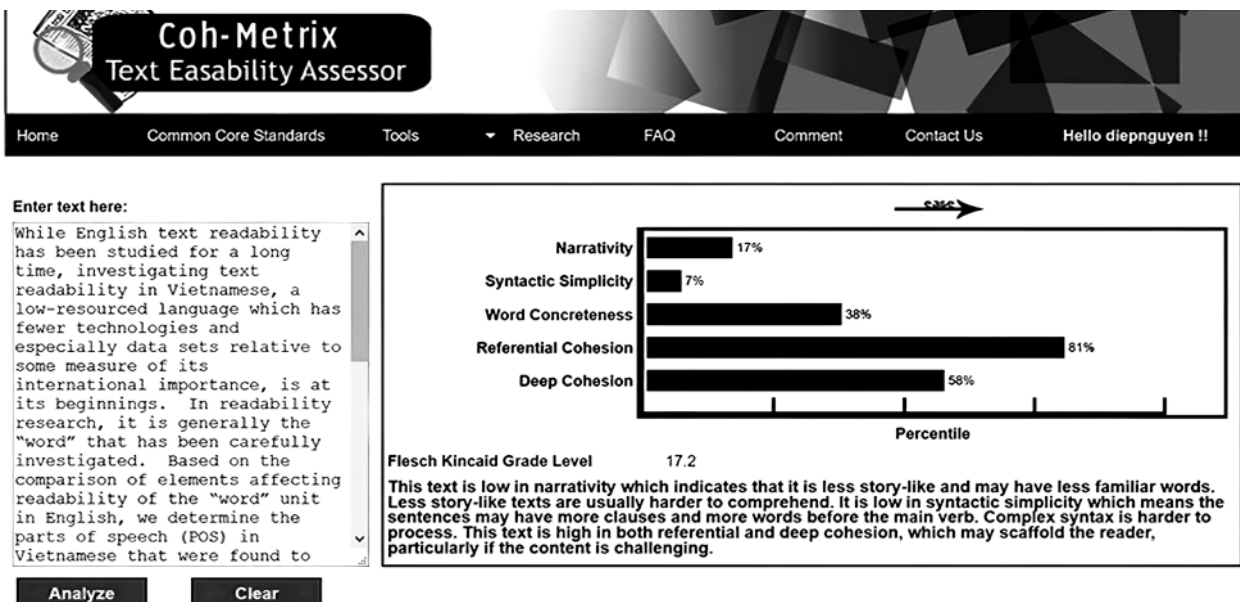


Figure 3. Text level- Cohesion and Readability

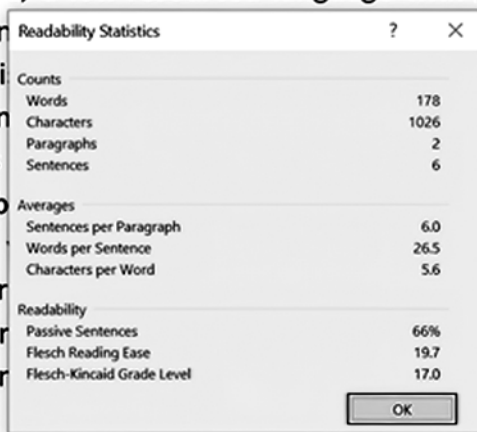
Readability plays an important role in teaching and learning languages in general and English in particular. Investing and analyzing the factors affecting readability in 3 levels: Words- Sentences- Texts; using the achievements of readability research, it saves time, effort, and money for

teachers with the applicable suggestions in the current context of Vietnam as follows:

- (1) Using available readability tools to category English materials with specific levels and grades as readability formular in Microsoft Word.

### AFFECTION OF THE PART OF SPEECH ELEMENTS IN VIETNAMESE TEXT READABILITY Điệp Thi Nhu NGUYỄN- An-Vinh LƯƠNG- Điền ĐINH

While English text readability has been studied for a long time, investigating text readability in Vietnamese, a low-resourced language with poor research technologies and data sets question readability research, it is Based on the comparison we determine the parts Vietnamese text readability different difficulty level their correlations. In terms editing documents, reform general and foreigners in identify those linguistic




is at its beginnings. It has been carefully investigated. The “word” unit in English, it were found to influence Vietnamese textbooks at the POS frequencies and initially assist users when s for native Vietnamese in with these findings we can “potential” POS affecting

Figure 4. Automatic Tool of Flesch Readability in Microsoft Word

- (2) Using available corpus to select appropriate materials for learners' levels in teaching and learning general English or specific one.

#### Earthquake in Indonesia – level 1



02-10-2018 07:00

Level 1
Level 2
Level 3


Sulawesi is an island in Indonesia. An **earthquake** hits near it. The earthquake makes a **tsunami**. It is 3 metres tall.

The tsunami moves into two cities. Around 600,000 people live there. More than 832 people die. Hospitals, hotels, a shopping centre, and thousands of homes are **destroyed**.

Difficult words: earthquake (when the ground moves), tsunami (a big wave started by an earthquake), destroy (break completely).

You can watch the original video in the Level 3 section.

#### Earthquake in Indonesia – level 2



02-10-2018 07:00

Level 1
Level 2
Level 3

A 7.5-**magnitude** earthquake hit near the Indonesian island of Sulawesi which **triggered** a 3-metre tsunami that **smashed into** two cities on the coast. These cities are home to 600,000 people.

The tsunami killed more than 832 people and destroyed hospitals, hotels, a shopping centre, and thousands of homes. The event affected the lives of as many as 1.6 million people.

Difficult words: magnitude (the size of power of something), trigger (start suddenly), smash (move into with a lot of force).

You can watch the original video in the Level 3 section.

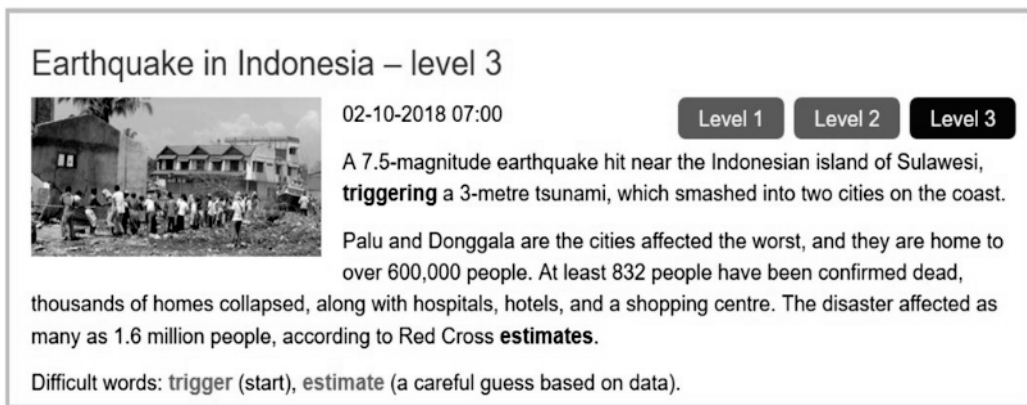


Figure 5. Examples of news in levels

It can be seen that the three written news are the same content, discussing the current news topic 'Earthquake in Indonesia.' However, they are composed at three different difficulty levels: Level 1 - Easy, Level 2 - Intermediate, and Level 3 - Difficult. These levels are categorized based on vocabulary and grammar complexity, progressively increasing in difficulty. This categorization aims to cater to specific reader needs: Level 1 is designed at a basic level for beginners in English reading, Level 2 is crafted for intermediate proficiency, and Level 3 represents a challenging difficulty level intended for advanced readers.

Using the automatic tool of Flesch Readability in Microsoft Word as presented in (1) (Using available readability tools to category English materials with specific levels and grades as readability formula in

Microsoft Word), the results show that the surface linguistic elements increase from level 1 to level 3 as follows: at level 1, the total number of words in the text is 48, this number is 110 at level 2 and reaches to 185 at level 3; the average sentence length in words is 6 at level 1, level 2 is 9.1, and 11.5 at level 3.

While 1000 basic words in English language is used to write at level 1, 2000 basic words for level 2, and level 3 is 3000 basic words. In addition, The sentence structure also progressively increases in complexity and difficulty in each of the levels, ranging from basic to advanced. For example, the simple present tense at level 1; adding the past tense in level 2; and the having present perfect tense which relative clauses and reducing relative clauses at level 3. This can be seen more detailed in the following table:

Table 1. Comparing readability of the texts in the same content at 3 different levels

	Level 1	Level 2	Level 3
<b>Surface linguistic elements (Using FLESCH)</b>	- Total number of words in the text: <b>48</b> - Average sentence length in words: <b>6</b>	- Total number of words in the text: <b>110</b> - Average sentence length in words: <b>9.1</b>	- Total number of words in the text: <b>185</b> - Average sentence length in words: <b>11.5</b>
<b>Vocabulary</b>	- Using 1000 basic words in English language - Single words: earthquake (when the ground moves) -> explained by the most popular words	- Using 2000 basic words in English language - Single words - Compound words A 7.5-magnitude earthquake (magnitude: the size of power of something)	- Using 3000 basic words in English language - Single words - Compound words A 7.5-magnitude earthquake -> Unexplained difficult word Estimate (a careful guess based on data)

<p><b>Grammar</b></p>	<ul style="list-style-type: none"> <li>- Simple present tense</li> <li>- Simple sentences</li> </ul> <p>The earthquake makes a tsunami</p>	<ul style="list-style-type: none"> <li>- Simple present tense, simple past tense</li> <li>- Complex – compound sentences with relative clause:</li> </ul> <p>A 7.5-magnitude earthquake hit near the Indonesian island of Sulawesi which triggered a 3-metre tsunami ...</p>	<ul style="list-style-type: none"> <li>- Simple present tense, simple past tense, present perfect tenses</li> <li>- Complex – compound sentences with reducing relative clause:</li> </ul> <p>A 7.5-magnitude earthquake hit near the Indonesian island of Sulawesi, triggering a 3-metre tsunami, which smashed into two cities on the coast.</p>
-----------------------	--	--	--

(3) Comparing and contrasting the source language and the target one from the view of readability in teaching translation Subject

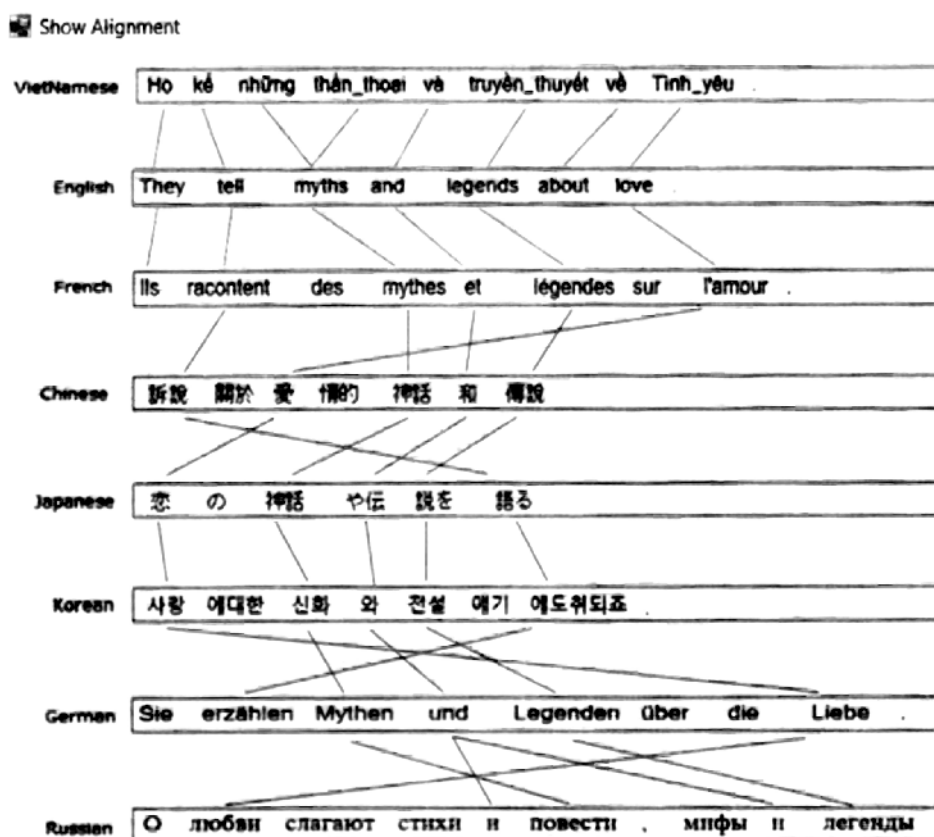


Figure 6. Alignment in 8 languages: Vietnamese, English, French, Chinese, Japanese, Korean, German, and Russian [13]

(4) Building own corpus for different levels for practical classes.

(5) Determining the criteria for rubrics quantitatively on Words- Sentences- Texts from the view of readability for assessment in teaching and learning languages.

**5. CONCLUSIONS**

Within the scope of a paper, the paper only presents an overview of readability and the linguistic factors having influence on readability and

suggests some practical applications of readability in teaching languages from the readability findings, so the details related to information processing and technology are mentioned in the cited publications. Besides suggesting the practical applications, we hope that the paper findings can support in teaching languages, it is also the foundation for us to investigate the other future studies on the view readability more deeply and comprehensively to meet the demand for users and Vietnamese community in this era of technology.

## REFERENCES

- [1] Dale, E., & Chall, J. S., "The Concept of Readability", *Elementary English*, 26(1), pp 19-26, 1949.
- [2] Fukun Xing, Dongyuan Cheng & Jianzhong Pu, A New Approach to Readability Study Based on Information Computing, *Proceedings - ALPIT 2008, 7th International Conference on Advanced Language Processing and Web Information Technology*, pp. 156–161, 2008.
- [3] Wray, D. & J. Dahlia., "Exploring the Readability of Assessment Tasks: The Influence of Text and Reader Factors", *REMIE: Multidisciplinary Journal of Educational Research*, 3(1), p. 73, 2013.
- [4] Dubai, H. W., "Smart Language: Readers, Readability, and the Grading of Text. Impact Information," *Costa Mesa: California*, pp. 4-6, 2007.
- [5] Chiang, W.-C., T.D. Englebrecht, T.J. Phillips Jr & Y. Wang., "Readability of financial accounting principles textbooks," *The Accounting Educators' Journal*, 18, p. 48, 2008.
- [6] Jones, M.J., "Methodological themes: Critical appraisal of the Cloze procedure's use in the accounting domain," *Accounting, Auditing and Accountability Journal*, 10(1), pp. 105–106, 1997.
- [7] Gray, S. W., & Leary, E. B., "What Makes a Book Readable", Chicago, Illinois: The University of Chicago Press, 1935.
- [8] Lorge, I., "Predicting Readability," *Teachers College Record*, 45(6), pp. 404–419, 1944.
- [9] Flesch, R. F. A., "New Readability Yardstick", *Journal of Applied Psychology*, 32(3), pp. 221-233, 1948.
- [10] Nguyễn Thị Như Điệp, "Các yếu tố ngôn ngữ ảnh hưởng đến độ khó của văn bản tiếng Việt (đối chiếu với tiếng Anh)," Doctoral Thesis, University of Social Sciences and Humanities, HCM city, 2021.
- [11] Fromkin V, Rodman R & Hyams N., "An Introduction to Language", 11th Edition, Cengage Learning, 2019.
- [12] Diep Thi Nhu Nguyen, An-Vinh Luong, Dien Dinh, "Affection of the part of speech elements in Vietnamese text readability," *Acta Linguistica Asiatica*, 9(1). doi:10.4312/ala.9.1.105-118, 2019.
- [13] Hoàng Khuê, Nguyễn Thị Như Điệp, Đinh Điền, Nguyễn Thanh Thủy, "Ứng dụng kho ngữ liệu song song đa ngữ trong việc giảng dạy tiếng Việt cho người nước ngoài", *Proceedings of the Workshop on 'Vietnamese Studies and Vietnamese Language Research*, 2018.

## Ứng dụng các kết quả nghiên cứu độ khó của văn bản trong việc giảng dạy ngôn ngữ

Nguyễn Thị Như Điệp và Lâm Thành Nam

### TÓM TẮT

Độ khó của văn bản, đặc biệt trong ngôn ngữ Anh, đã được nghiên cứu từ cuối thế kỉ thứ 19 với hàng ngàn công trình được công bố. Kết quả nghiên cứu về vấn đề này không chỉ tạo ra nguồn tài nguyên phong phú mà còn được ứng dụng toàn cầu trong nhiều lĩnh vực. Hiện nay, vấn đề này luôn gây hứng thú và các nhà nghiên cứu vẫn tiếp tục nghiên cứu không chỉ trong ngôn ngữ Anh mà còn trong nhiều ngôn ngữ khác. Tuy nhiên, tại Việt Nam, nghiên cứu về Độ khó của văn bản vẫn còn hạn chế nhằm tìm ra các giải pháp đáp ứng cho nhu cầu của xã hội. Dựa trên các kết quả nghiên cứu về độ khó của văn bản, bài viết trình bày tổng quan về độ khó của văn bản và các yếu tố ngôn ngữ có ảnh hưởng đến độ khó của văn bản. Từ góc nhìn của độ khó của văn bản và các kết quả nghiên cứu được ứng dụng của độ khó của văn bản, bài viết trình bày và phân tích các yếu tố ngôn ngữ có ảnh hưởng đến độ khó của văn bản ở 3 cấp độ chính: Từ, Câu và Văn bản, được thể hiện bằng các nghiên cứu điển hình minh họa trong ngôn ngữ là tiếng Anh hoặc tiếng Việt. Dựa trên nền tảng này, bài viết đề xuất một số ứng dụng thực tế của độ khó của văn bản trong việc giảng dạy ngôn ngữ nói chung, đặc biệt là việc giảng dạy tiếng Anh tại Việt Nam hiện nay, để đáp ứng nhu cầu của người sử dụng và cộng đồng người Việt trong thời đại công nghệ.

**Từ khóa:** độ khó của văn bản, yếu tố ngôn ngữ, các ứng dụng, giảng dạy ngôn ngữ

Received: 05/12/2023

Revised: 23/12/2023

Accepted for publication: 24/12/2023