Authoring WCAG2.0-Compliant Texts for the Web through Text Readability Visualization

Evelyn Eika¹ and Frode E. Sandnes^{1,2}

¹Oslo and Akershus University College of Applied Sciences, Norway ²Westerdals Oslo School of Art, Communication and Technology Evelyn.Eika@hioa.no, frodes@hioa.no

Abstract. Texts on the web need to be readable in order to be accessible to a wide audience. WCAG2.0 states that tests should not exceed the reading level of upper secondary education. Several readability measures have been proposed over the last century. However, these measures give an accumulated measure of the text and do not help pinpoint specific problems in the text. This paper proposes a text visualization approach that emphasizes readability issues in texts. The texts are visualized in the textual domain. The intention of the visualization approach is to draw the attention of the author towards the aspects of the text that potentially are hard to read, allowing the author to revise the text and consequently making the text more readable.

Keywords: WCAG2.0, readability, visualization, universal design, cognitive disability, dyslexia

1 Introduction

The Web Content Accessibility Guidelines (WCAG2.0) provide a set of minimum requirements that are intended to help make the web accessible to as many people as possible. One of the issues addressed by WCAG is the quality of the text. Criterion 3.1.5 [1] states that: "When text requires reading ability more advanced than the lower secondary education level … supplemental content … is available. (Level AAA)". How is one to realize this relatively abstract criterion in practice? As with many of the criteria in WCAG2.0, they are formulated as issues to be checked after the process of designing web content is finished, such as minimum contrast on web pages [2, 3]. The viewpoint taken in this paper is that the most efficient and cost effective way of designing high quality and accessible web content is to do incorporate WCAG-requirements formatively during the design process instead of after the design process.

There have been few major advances in the computer-assisted composition of text in recent decades. Perhaps the only exception is tools that support distributed and collaborative writing activities [4]. Most authors use word processors to compose their text. The WYSIWYG (what-you-see-is-what-you-get) editors [5] have become a convention since the graphical user interface became commonplace in the mid 90's. Some individuals still believe that the final presentation should be separate from the task of composition and will therefore use simple text editors to compose their text and later employ typesetting software such as Latex [6]. This is because a focus on the final presentation during the authoring process draws valuable attention from the composition. However, the text editing tools used in the Latex community are probably even less useful for text composition, and often focus on syntax highlighting for Latex commands, rudimentary spelling and grammar checks.

On the other hand, the main objective of the WYSIWYG editor is to represent the text as close as possible to the way it will appear in published form, being it in print, or in electronic form. WYSIWYG editors are becoming increasingly powerful by integrating tools that enhance the authoring process such as on the fly spell-checking, synonyms, antonyms, grammar checks and simple style issues. For example, Microsoft Word typically underlines erroneously spelled words with a red line, while passages with grammatical issues are marked in blue. One drawback of using the blue underlines is that it employs one visualization feature to represent a vast array of grammatical issues, while the red line only indicates one thing -a misspelled word. The blue line does not give a clue to what the particular issue is. The author therefore has to investigate the blue line further by, for instance, mouse-over to get a tool-tip.

The interest in measures of readability has been vast, and a number of readability measures have been proposed and debated over the last century [7]. The WCAG2.0 supplementary documentation also refers to one of these, namely, the Flesch-Kincaid metric. Although many of these methods demonstrate high correlations with actual readability, they are all aggregated measures and do not indicate what the problems are and where these problems are.

This paper proposes a different way of editing documents. The text is not represented in the ad-hoc manner as in a text-editor, or presentation centric as in the WYSIWYG editors. Instead, the text is presented according to commonly cited readability features.

2 Background

The literature on readability is plentiful, and much of the attention has been focused on readability indices, especially readability index validity [7] and readability index accuracy [8]. Researchers have experimented with various ways of optimizing the formulas for quantifying readability [9]. The most commonly cited measurements include Flesch-Kincaid, Gunning Fog and SMOG. A common feature of most of these formulas is the use of sentence length and word difficulty. More recently, readability is also seen in the context of disability, such as dyslexia [10] and multimodal access via screen readers.

Readability formulas are objective and can be computed automatically. They were originally developed for matching the appropriate texts to students with a given reading level. The reading indices should therefore also be suitable for determining readability on the web. However, the readability formulas have been criticized for not capturing real readability. Also, these numbers are not necessarily useful to authors. Qualitative methods have also been proposed such as levelling [11].

There are at least two ways of making text on the web more accessible to a broad audience: during composition or after composition. Several attempts have been made at automatically summarizing and simplifying existing texts. For example, Jing [12] used syntactic knowledge, context information and statistics to remove extraneous phrases from sentences. Chandrasekar et al. [13] experimented with finite state grammars and the Supertagging model for text simplification. More recent approaches are corpus based [14]. Simplification techniques have also been applied to other languages such as Spanish [15]. Common to all automatic approaches is their current inefficiency compared to humans.

It is our belief that the text becomes more readable if the quality of a text is ensured during composition, as it is harder for others to improve a text they have not written themselves, let alone depending on automatic summarizers and text simplifiers. One approach is to give authors tools that allow them to visualize flaws in their writing.

The visualization of text has been proposed to help navigate digital libraries, including galaxies visualization [16], mapping text to surfaces [17], principal components analysis [18], multiple views [19], and self-organizing maps [20]. The purpose is to visualize the relationship between different documents and not express aspects of the document contents per se. Such methods are thus seemingly not helpful to authors.

Other text visualization techniques aim at helping users understand the content of documents. Tag clouds [21] have become widely used where non-stop words in the text are displayed in a cloud like shape where the size of each word is related to its frequency. Variations on the tag cloud include spark clouds that allow the frequency of terms to be observed over time [22]. Ham et al.'s [23] phrase nets visually show how terms are linked via a user-selected relation, such as finding the family lines in the Bible. Other interesting text visualizations include Wattenberg and Viégas' [24] word tree, which is used to show various forms of a query phrase in a document. Different instances of a query prefix are shown at different levels, and text size is used to communicate importance. This can help a user quickly find a particular instance instead of browsing all instances of the prefix sequentially.

Chung et al. [25] used visualization to help deaf people understand news texts in Korean. They claim that deaf people have difficulty comprehending complex texts as they are more used to visual representations through sign language. Their solution involved identifying the various clauses of text and representing them visually. The clauses are split up and the relationships between the clauses are visualized using arrows.

Kim et al. [26, 27] used shaded dots to visualize the readability of text. At the beginning of a sentence, the dots were white and gradually become darker as the sentence becomes longer. Very long sentences end up with nearly black dots. Punctuation marks such as commas generally add to the readability and these symbols slow the further darkening of the pixels. Kim et al.'s approach makes it easy to get an overview of the difficulty of a text, and in particular, where the challenges approximately occur in the text. The approach therefore provides more useful information than the traditional readability formulas. However, since the visualisations do not show text, it is still not obvious how authors can easily make use of the visualisations to improve their writing. Oelke et al. [28] developed the VisRA tool for visualizing readability with the intention to help authors compose more readable texts. This tool provides a mixture of both text centric visualisations and other visualisation mechanisms such as correlation maps. Typically, each paragraph is marked with a colour that indicates its readability. In addition, a column on the right indicates which factors affect readability. The factors employed in VisRa are vocabulary difficulty (percentage of words not found in a list of 1000 most frequent words), word length (average characters per word), nominal forms (a combined measure comprising the noun/verb ratio and the number of gerunds and nominalized words ending with -ity, -ness, etc.), sentence length (words per sentence) and complexity of the sentence structures (the branching factor in the phrase structure tree of a sentence). VisRa illustrates problems at paragraph level but problems at sentence and word level go undetected. Another drawback is that the VisRa tool is not made available to the general research community.

In a similar attempt, Kamakar and Zhu [29] visualized text at paragraph level to assist authors using several indices including Flesch-Kincaid, Gunning Fog, SMOG, Coleman Liau and ARI. Also, three visualization techniques are used, colour-coded circles, colour-coded abbreviations and Chernoff faces. The colour-coded rings show the average readability indices for the paragraphs where red is the lowest readability, going through orange yellow and green being the highest readability. The colour-coded abbreviation view shows each of the indices with the same colour-coding. The Chernoff face visualization shows the five parameters as oval vs. round face, size of the eyes, orientation of the eyes, size of the mouth and orientation of the mouth. One potential problem with Kamakar and Zhu's approach is that it visualizes relatively abstract and aggregated entities – readability indices. It may be hard for the authors to transform this information into concrete text editing improvements. One study by Liu et al. [30] made use of faces. They created an emotion-detecting engine to detect the emotion in writing and then colours and emoticons to convey one of six emotions for a passage of text.

A different visualization approach proposed by Kamakar and Zhu [31] uses bar graphs to represent sentences in terms of word length, levels of grey to represent six levels of word difficulty and white blocks to indicate sentence clauses. Kamakar and Zhu's approach is among the closest to the one presented herein. The major difference between Kamakar and Zhu's method and our method is that they use a different representation of coloured square bars while we only use text. Another difference is that our visualization is simple with less noise, while theirs provide rich details that may divert the users' attention from the important issues.

3 Method

This section outlines the proposed text visualization technique. First, the textual attributes of interest are discussed, followed by the visual mechanisms employed.

Readability Attributes

This study adopted three key features from the readability research literature: sentence length, word complexity and prepositional phrases. Sentence length is commonly connected to readability, as long sentences are generally considered harder to read than shorter sentences. Word complexity is also highly cited as a factor that affects readability, in particular, syllable count where words with more than two syllables are considered hard words. Prepositional phrases, that is, phrases introduced by prepositions such as on, in, over, etc., are believe to add to the complexity of a sentence if many prepositional phrases are used within the same sentence.

In addition, we propose simply to use paragraph length as a feature of readability measured in number of characters. Long paragraphs are less tempting to read for the impatient reader, may seem overwhelming for less trained readers, and are more timeconsuming to navigate for individuals that rely on screen readers.

Visual Attributes

Visualizations are achieved by employing visual features that are noticeable to the reader. These typically include position, size, shape, orientation, colour and texture. Unlike other text visualization approaches that represent readability in other domains through various transformations [31], our approach operates in the textual domain.

The textual domain imposes certain typographical constraints such as how sequences of letters form words along horizontal lines going from left to right (in languages based on the Latin alphabet). Moreover, word orders are constrained by their respective sentences, and sentence orders are constrained by their respective passages. Therefore, the visual degrees of freedom include the following: (a) positional features, such as horizontal spacing, line breaks, vertical spacing, size, colour comprising text and background colour; and (b) textural features, such as typeface family, bold, italics, superscript, subscript, underline, strikethrough, etc.

Visualization Framework

Sentence length is central to a majority of readability studies [7]. Sentence length is thus chosen as an attribute. Often length is measured in the number of words per sentence. In this study, the number of characters is used as the unit of measure. In ordinary typeset passages of flowing prose, it is not immediately obvious how long sentences are. To determine the length, the passage must be scanned or read. In our framework, a sentence is represented on a single line up to 100 characters. If a sentence is longer than 100 characters, it is clipped at the end and ellipses (...) are used to signal that the sentence is clipped. Each sentence thus becomes a bar as in a bar graph, where the length of the line directly relates to the sentence length. A non-proportional font where each character has the same width is used to ensure that the representation of length is in the same scale throughout, as illustrated in the examples below.

A somewhat longer sentence.

A short phrase.

Sentences with several parts separated by commas are divided using a ling break and continued on the next line. That is, the first part to the first comma is placed on the first line, the next part after the comma to the subsequent comma is placed on the next line after indentation, and so forth. This is illustrated in the following example:

```
This is the first sentence,
it has more parts,
and even a final part.
```

Another attribute frequently employed in readability studies is word length, usually represented in terms of number of syllables. Several studies consider words with more than two syllables as difficult. We therefore emphasize words with more than two syllables in the text using upper case. In the previous example, the word *sentence* has two syllables (sen-tence), but is incorrectly detected as having three syllables and is therefore marked as a difficult word. The syllable count is approximated using a simple algorithm based around vowel counting. Although the syllable counting procedure is not perfect, it gives a sufficient indication for this application. The visualization is thus:

```
This is the first SENTENCE,
it has more parts,
and even a final part.
```

Prepositional phrases are also known to affect readability. The propositional phrases in all sentences with more than one propositional phrase are highlighted. This allows the author to easily spot sentences that may come across as hard to read.

Finally, we expect that paragraph length is one predictor of readability on the web. In typeset text, it is usually quite easy to spot the length of a paragraph, except when a paragraph spans several pages. Moreover, the advantage of perceiving paragraph length from the typeset text is lost with the proposed approach since lines are broken. We therefore introduce different backgrounds to indicate length. That is, the lines of the text for the first 100 words have an ordinary background, the lines for the subsequent 100 words is marked as long. The marking strength is increased for every 100 words; the limit of 100-word paragraph was based on [32].

Note that the word counts are used irrespective of lines since the layout proposed herein breaks up text utilizing more lines than the typeset text. Length marking paragraphs may help authors to know where to rephrase passages, cut text or reorganize the text into different paragraphs.

Alternative Views

To help authors focus on the most important issues in a text, the text visualizer allows other views besides the view showing the text in its correct chronological order. The sentence viewer lists the sentences in decreasing length, allowing authors to focus on assessing the readability of the longest sentences in the text. The word view lists all words with more than two syllables in decreasing order of syllable length and character lengths. Finally, the passage length view shows the paragraphs in decreasing order according to length.

4 **Results**

The first example illustrates the technique on a text for children at the level one (Flesch-Kincaid Reading Ease 87.6, Flesch-Kincaid Grade Level 3.2):

The layout suggests that this is an easy-to-read text. There are no long sentences and no complex prepositional phrases, but there are a couple of long words. However, these words are well known. The next example shows the same content at a higher reading level (Flesch-Kincaid Reading Ease 82.7, Flesch-Kincaid Grade Level 4.9):

```
A young man goes [TO] a SUPERMARKET [FOR] his first day [AT] work.
The SUPERMARKET MANAGER says hello [TO] him with a smile and then gives him a brush [TO] clean th...
"But I studied at UNIVERSITY."
THE MANAGER then says,
"On,
I'm sorry.
I didn't know that you studied at UNIVERSITY.
Give me the brush and I'll show you how to do it."
```

This passage is more difficult to read as it has fewer and longer sentences with a couple of more complex prepositional phrases. The following example shows the same content with an even higher reading level (Flesch-Kincaid Reading Ease 81.1, Flesch-Kincaid Grade Level 5):

```
A young man goes [TO] a SUPERMARKET [FOR] his first day [AT] work.

The SUPERMARKET MANAGER WELCOMES him with a smile and then gives him a brush to sweep the floor.

The young man looks very SUPERFISED and says,

"But I studied at UNIVERSITY."

THE MANAGER then says,

"Oh,

I'm sorry.

I didn't REALIZE that you studied at UNIVERSITY.

Give me the brush and I'll show you how to do it."
```

There is not much visible difference between the two and the readability scores are marginally different. The final example shows an extract from a hard-to-read disclaimer (Flesch-Kincaid Reading Ease -3.4, Flesch-Kincaid Grade Level 25.2):

```
LICENSOR agrees with *** that [FOR] this effort [TO] be SUCCESSFUL,
LICENSOR and *** need [TO] work TOGETHER [IN] COLLABORATION and COOPERATION [IN] their ...
IN PARTICULAR,
LICENSOR shall assist and consult with *** upon request [ON] any topic REASONABLY RELAT...
```

Clearly, the readability indices are off the scales. A visual inspection reveals that the paragraph itself is too long, and that the sentences are all too long with many prepositional phrases. Moreover, there are many difficult-to-read words. Note also that not all prepositional phrases are detected.

5 Discussion

The focus of the visualization approach is on sentence length and word difficulty. This is based on the assumption that long words and long sentences are difficult to read. Although this is often the case, it is not always true. It is very possible to write incomprehensible short sentences using short words. Some sentences can become easier to read if more words are used, inclusive of appropriate long words. Certain long words are frequently used and are thus well known. One may also argue that rhythm and variation in language make texts easier to read. Fortunately, the proposed approach makes such variations and rhythms visible.

In conclusion, the visualization is not intended to be used to eradicate all long words and long sentences, but rather to make the authors aware of their presence and allow them to deliberate their appropriateness. A potential drawback of the proposed strategy is that the limits are based on fixed pre-determined values. These values may not necessarily be correct for different writing styles and different genres.

6 Conclusions

A text-oriented visualisation approach was presented where the objective is to draw attention to aspects of writing which may reduce the readability of text. The approach is simple and thus easy to implement. However, although the features visualized are useful, the present approach does not capture other important aspects of text that affect readability. It is therefore unlikely that such a tool can be a complete solution. It could be one of many tools in the authors' toolbox. Future work should focus on evaluating the effectiveness of the visualisation and exploring how to automatically detect and visualize deeper attributes of readability, such as the use of transitional words and text coherence. For such purpose, it may be necessary to draw from natural language processing techniques. The approach presented herein may also be applicable to language learning [33] and teaching academic writing [34].

References

- 1. World Wide Web Consortium, Web content accessibility guidelines (WCAG) 2.0 (2008)
- Sandnes, F.E., Zhao, A.Q.: An Interactive Color Picker that Ensures WCAG2.0 Compliant Color Contrast Levels. Procedia Computer Science 67, 87–94 (2015)
- Sandnes, F.E., Zhao, A.Q.: A contrast colour selection scheme for WCAG2. 0-compliant web designs based on HSV-half-planes. In: Proceedings of System, Man and Cybernetics Conference SMC2015, pp. 1233-1237. IEEE (2015)
- Hart-Davidson, W., Spinuzzi, C., Zachry, M.: Visualizing writing activity as knowledge work: challenges & opportunities. In: Proceedings of the 24th annual ACM international conference on Design of communication, pp. 70-77. ACM (2006)
- DeRose, S.J., Durand, D.G., Mylonas, E., Renear, A.H.: What is text, really?. ACM SIGDOC Asterisk Journal of Computer Documentation 21, 1—24 (1997)
- 6. Lamport, L.: Latex. Addison-wesley (1994)
- Janan, D., Wray, D.: Reassessing the accuracy and use of readability formulae, Malaysian Journal of Learning and Instruction 11, 127—145 (2014)
- Pitler, E., Nenkova, A.: Revisiting readability: A unified framework for predicting text quality. In: Proceedings of the Conference on Empirical Methods in Natural Language Processing, pp. 186-195. ACL (2008)
- McLaughlin, G.H.: SMOG grading: A new readability formula. Journal of reading 12, 639– 646 (1969)
- Berget, G., Sandnes, F.E.: Do autocomplete functions reduce the impact of dyslexia on information-searching behavior? The case of Google. Journal of the Association for Information Science and Technology (2015)
- 11. Fry, E.B.: Text readability versus leveling. Reading Teacher 56, 286-292 (2002)
- 12. Jing, H.Y.: Sentence reduction for automatic text summarization. In: Proceedings of the sixth conference on Applied natural language processing, pp. 310-315. ACL (2000)
- Chandrasekar, R., Srinivas, B.: Automatic induction of rules for text simplification. Knowledge-Based Systems 10, 183—190 (1997)
- Chandrasekar, R., Doran, C., Srinivas, B.: Motivations and methods for text simplification. In: Proceedings of the 16th conference on Computational linguistics, pp. 1041-1044. ACL (1996)
- 15. Saggion, H., Martínez, E.G., Etayo, E., Anula, A., Bourg. L.: Text simplification in simplext. making text more accessible. Procesamiento del lenguaje natural 47, 341–342 (2011)
- Wise, J., Thomas, J.J., Pennock, K., Lantrip, D., Pottier, M., Schur, A., Crow, V.: Visualizing the non-visual: spatial analysis and interaction with information from text documents. In: Proceedings of Information Visualization, pp. 51-58. IEEE (1995)
- Rohrer, R.M., Ebert, D.S., Sibert, J.L.: The shape of Shakespeare: visualizing text using implicit surfaces. In: Proceedings of IEEE Symposium on Information Visualization, pp. 121-129. IEEE (1998)
- Booker, A., Condliff, M., Greaves, M., Holt, F.B., Kao, A., Pierce, D.J., Poteet, S., Wu. Y.J.J.: Visualizing text data sets. Computing in science & engineering 1, 26–35 (1999)
- Eler, D.M., Paulovich, F.V., Oliveira, M., Minghim, R.: Coordinated and multiple views for visualizing text collections. In: 12th International Conference on Information Visualisation, pp. 246-251. IEEE (2008)
- Henderson, J., Merlo, P., Petroff, I., Schneider, G.: Using syntactic analysis to increase efficiency in visualizing text collections. In: Proceedings of the 19th international conference on Computational linguistics, pp. 1-7. ACL (2002)

- 21. Bateman, S., Gutwin, C., Nacenta, M.: Seeing things in the clouds. In: Proceedings of the nineteenth ACM conference on Hypertext and hypermedia, p. 193. ACM (2008)
- Lee, B., Riche, N.H., Karlson, A.K., Carpendale, S.: Sparkclouds: Visualizing trends in tag clouds. IEEE Transactions on Visualization and Computer Graphics 16, 1182–1189 (2010)
- 23. Van Ham, F., Wattenberg, M., Viégas, F.B.: Mapping text with phrase nets. IEEE Transactions on Visualization and Computer Graphics 15, 1169--1176 (2009)
- 24. Wattenberg, M., Viégas, F.B.: The word tree, an interactive visual concordance. IEEE Transactions on Visualization and Computer Graphics 14, 1221-1228 (2008)
- Chung, J.W., Min, H.J. Kim, J., Park, J.C.: Enhancing readability of web documents by text augmentation for deaf people. In: Proceedings of the 3rd International Conference on Web Intelligence, Mining and Semantics, p. 30. ACM (2013)
- 26. Kim, H., Lee, D., Park J.W.: Textual visualization based on readability. In: SIGGRAPH Asia 2011, p. 9. ACM (2011)
- 27. Kim, H., Park, J.W., Seo, D.: Readability Visualization for Massive Text Data. International Journal of Multimedia & Ubiquitous Engineering 9 (2014)
- Oelke, D., Spretke, D., Stoffel, A., Keim, D.: Visual readability analysis: How to make your writings easier to read. IEEE Transactions on Visualization and Computer Graphics 18, 662--674 (2012)
- 29. Karmakar, S., Zhu, Y.: Visualizing multiple text readability indexes. In: 2010 International Conference on Education and Management Technology, pp. 133-137. IEEE (2010)
- Liu, H., Selker, T., Lieberman, H.: Visualizing the affective structure of a text document. In: CHI'03 extended abstracts on Human factors in computing systems, pp. 740-741. ACM (2003)
- Karmakar, S., Zhu, Y.: Visualizing text readability. In: 2010 6th International Conference on Advanced Information Management and Service, pp. 291-296. IEEE (2010)
- Markel, M., Vaccaro, M., Hewett, T.: Effects of paragraph length on attitudes toward technical writing. Technical communication, 454-456 (1992)
- Jian, H.-L., Sandnes, F.E., Law, K.M.Y, Huang, Y.-P., Huang, Y.-M.: The role of electronic pocket dictionaries as an English learning tool among Chinese students. Journal of Computer Assisted Learning 25, 503--514 (2009)
- Jian, H.-L., Sandnes, F.E., Huang, Y.-P., Cai, L., Law, K.M.Y.: On students' strategy-preferences for managing difficult course work. IEEE Transactions on Education 51, 157-165 (2008)