International Journal of Computing and Digital Systems
ISSN (2210-142X)
Int. J. Com. Dig. Sys. , No. (1) (Mar-2022)

https://dx.doi.org/10.12785/ijcds/110189

The Design and Evaluation of an Online Dictionary User Interface

Sayush Lal Shrestha [1] **and Pietro Murano**[1]

[1]*Department of Computer Science, Oslo Metropolitan University, Oslo, Norway*

Received 3 Jun. 2021, Revised 3 Feb. 2022, Accepted 8 Mar. 2022, Published 31 Mar. 2022

Abstract: This paper presents a new user interface design for an online dictionary. Online dictionary user interfaces have not been given much research attention over the years and this paper addresses this important gap in knowledge. The dictionary chosen for the redesign was the Norwegian Online Dictionary. The design aimed to improve usability and universal design aspects. The prototype was firstly evaluated using the Wave and SortSite automated tools. The prototype was evolved until these tools showed no errors to be present. Then, an empirical experiment with 20 human participants was conducted as a comparison of the new user interface and the Norwegian Online Dictionary user interface. Ten participants were native Norwegian speakers and the remaining ten participants were nonnative Norwegian speakers with a minimum of A2 Norwegian language level. Errors and user experience data were collected and statistically analysed using t-tests. User experience was elicited by means of a bespoke questionnaire. Statistically significant results were observed where most p-values were to the 0.0001 level, indicating strong significance. The analysis suggests the new prototype user interface incurred significantly fewer errors and significantly higher perceptions of positive user experience when compared with the Norwegian Online Dictionary. The new design could be easily used as a basis for future online dictionary designs or for improving current online dictionaries.

Keywords: Online dictionary, electronic dictionary, usability, universal design, user experience

1. Introduction

The online world would ideally be composed of web sites that are always usable and universally designed. If this were the case, it would mean that more diverse users with various requirements would be able to access and enjoy web content with fewer problems.

One such area where there is a need to improve the usability and universal design, concerns online dictionaries. Online dictionaries are used by millions of users every month. Dictionary.com [1] reports that from 2014 more than 70 million users access their dictionary and Merriam-Webster's [2] online dictionary reports they have more than 40 million users per month. Also, in Norway, the online dictionary approved by the Language Council of Norway [3] saw in 2020 around 37 million searches [4].

This clearly shows that use of such web sites is large. However, several of these well-known dictionaries show themselves to be poor or lacking in usability and in being universally designed. We used the Wave Evaluation Tool [5] as a Chrome extension to assess the home pages and a one-word search page of Dictionary.com, Merriam-Webster.com, Macmillan Dictionary [6] and Cambridge Dictionary [7]. The aim was not to present detailed results of the Wave analysis, but purely to show that each of these dictionaries had various errors and other issues flagged up by the Wave tool and connected to accessibility.

The above indicates clearly two main things. The first, is that online dictionaries are rather popular. The second, is that well known online dictionaries have some accessibility and usability problems.

This paper will therefore present the results of a re-designed user interface and empirical evaluation of the redesign. Practically, the Norwegian Online Dictionary [8] was used as a test case to demonstrate what could and should be achieved by designers of online dictionaries. The Norwegian Online Dictionary was chosen because the authors are based in Norway and so had strong familiarity with the dictionary. This dictionary was also chosen because improving its design could make a contribution to Norwegian society whilst benefiting the wider world society for online dictionaries.

Overall, the approach used in this research was essentially a form of empirical experimental work. Whilst quantitative data was collected in the evaluation via an empirical experiment, we felt it important to also collect qualitative data for the user experience aspects. Fig. 1 presents a block diagram illustrating the high-level overall

E-mail address: lalsayush@gmail.com, pietro.murano@oslomet.no

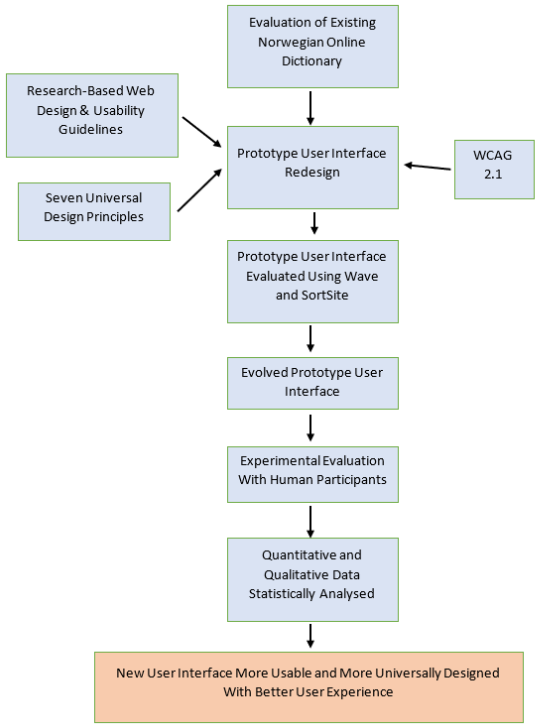

Figure 1. Overall High-Level Process Followed

process followed for the new user interface.

The rest of this paper is therefore structured as follows: The next section will discuss some relevant literature. Then the prototype developed and prototype evaluation will be presented. These will be followed by a series of results and a final discussion and conclusion.

2. RELATED WORKS

A comprehensive series of studies involving the usability and universal design of Online dictionaries, to our knowledge has not been done before. Therefore, the literature on specifically the usability of online dictionaries is rather meagre.

However, there are related works to this area regarding various findings concerning web page design which would apply to online dictionaries and their usability.

In a useful study by Ling and van Schaik [9], the authors investigated font type and text line length. They conducted two experiments where in one a visual search task was used and in the other information retrieval featured. Overall, they found that font type did not affect task performance. However, the Arial font was preferred by participants. Line length aspects achieved significance, where the first experiment suggested that longer line lengths were superior for scanning. Both experiments suggested that line lengths that were shorter achieved improved results for subjective aspects.

In another study by Duchnicky and Kolers [10], text line length, 'character density' and window height were investigated. Amongst several results that they attained, they found that lines of text in full-width and two-thirds width screens obtained a 25% faster reading time when compared with text lines appearing on a one-third wide screen. They also found that reading text in a very narrow column format was significantly slower than in a wider column format. Furthermore, in a more recent study [11] the researchers also had data to suggest that reading text on the screen in shorter lines was slower.

Colour is also a factor which can affect how users use and perceive a web page. Ling and van Schaik [12] found that in a navigation bar context, using a higher contrast between the foreground text and the background colour led to performance gains (search speed) and better user experience.

Concerning online dictionaries Fuertes-Olivera [13] had expectations that in a teaching and learning context they would become more usable over time. The author also expected users to somewhat change in relation to using such dictionaries.

Also of interest are the viewpoints of Heid and Zimmermann [14]. They suggested that in their experience some of the electronic dictionaries were rather complicated to use. They were also doubtful about users being able to deal with the complexities surrounding these kinds of dictionaries. The authors further suggested that electronic dictionaries could be evaluated like any other type of software (Note: An electronic dictionary may not necessarily be an online entity).

As stated above, there are very few actual evaluations of online dictionaries which have been published. The closest we were able to find was the work by Heid and Zimmermann [14]. We suggest that our work presented in this paper is extremely novel, because to our knowledge there have been no empirical evaluations of online dictionaries. Further, to our knowledge, there have been no published evaluation studies of the Norwegian Online Dictionary [8]. Therefore, in the next sections we present our suggested user interface redesign of the Norwegian Online Dictionary [8] along with an empirical evaluation comparing the current user interface and the redesigned user interface. In our evaluation we collected both performance data and subjective user experience data.

3. NEW USER INTERFACE

Before the new user interface was designed, we aimed to get a thorough understanding of the current user interface and the kinds of tasks that could be done. We also used the Wave Evaluation Tool [5] and SortSite [15] on the current Norwegian Online Dictionary [8] (version of 2019). We specifically tested the home page, some word searches and a word search where both versions of Norwegian were displayed at the same time. The tools revealed numerous

errors and problems.

The new user interface, represented by a small prototype that could potentially represent a whole dictionary was designed by following the WCAG 2.1 guidelines [16], the Research-Based Web Design & Usability Guidelines [17] and the Seven Universal Design Principles [18]. With the help of these guidelines the final prototype of the dictionary has keyboard compatibility, can accommodate screen readers, specified headings, labels, sections, contrast customization, page zoom customization, browser and device compatibility. We also note explicitly that the aim of the redesign was purely about the user interface and not the language content of the dictionary.

Some of the Research-Based Web Design & Usability Guidelines [17] were chosen as a guide to our design because the guidelines have been used by government agencies and private entities [17]. This suggests the guidelines have been useful in the past and therefore we decided to use these guidelines which have the indication of being tried and tested in real-life projects for real-life web sites. Since online dictionaries are used by people around the world and have been in existence for some time, these guidelines fitted the context of this work. While several Universal Design aspects are clearly implied in different ways in the guidelines, we also decided to make use of the more explicit Seven Universal Design Principles [18] to help guide our redesign. These helped to focus on ensuring the redesign was as universally designed as possible.

To design the new prototype three pages were selected containing the word search 'tur' in Bokmål, 'prøve' in Bokmål and 'språk' in both Bokmål and Nynorsk ('Begge'). These three words were selected as content for the prototype because they are familiar and simple Norwegian words. Our prototype outputted exactly the same content as the original dictionary, but within a redesigned user interface.

Figure 2 shows the overall final user interface design. The figure specifically shows the results of a search for the Norwegian Bokmål word 'Tur'. (Note: The content of the Norwegian Online Dictionary is copyrighted to the University of Bergen and the Norwegian Language Council - https://www.uib.no/ub/fagressurser/spesialsamlingene/142334/lisens-bokm%C3%A5lsordboka-og-nynorskordboka).

The basic operation of the dictionary was via a search box (See Figure 2). Users are able to type in a word, search for it and when this is found, the details are displayed, as shown in Figure 2 for the word 'Tur'. Furthermore, users are also able to specify which version of Norwegian is to be used. The example in Figure 2 shows Bokmål is the selected version of Norwegian.

The protype was developed using Bootstrap 4.1.3. For the layout and responsiveness, class, container, row, section and 'col-sm' were used. For the search form, 'form-control'

Figure 2. Overall Appearance of the Redesigned User Interface

was used. While for the button(s) 'btn' was used. The Class table was used to display the table. Also, HTML, CSS, PHP and JavaScript were used to achieve a final working prototype.

Finally, the redesigned user interface was then subjected to rigorous evaluation. The evaluations done are described in the next section.

4. EVALUATION

We conducted two main types of evaluation of the new user interface. The first was with the Wave Evaluation Tool [5] and SortSite [15]. The prototype was evolved to ensure that each of these tools showed no errors for our prototype.

Once the automated tools showed the prototype to be error free, we then conducted an empirical experiment with human participants.

A. *Experiment*

An empirical experiment was designed to compare the user interface of the current Norwegian Online Dictionary [8] (version of 2019) against the newly designed prototype user interface.

B. *Participants*

For this study it was useful to have participants with some basic experience and familiarity with the current Norwegian online dictionary. This was to ensure that the overall experience of the participants was one tempered by prior knowledge and therefore, not influenced by perceptions of experiencing something 'new'.

Further, for participants who were not native Norwegian language speakers, the participants had to have at least completed an A2 level Norwegian language course. This was stipulated so that nonnative Norwegian language speakers

would be able to read and understand the basic Norwegian language.

Lastly, since the evaluation was about the user interface (not the content of the dictionary) all the participants had to have basic IT skills.

Meeting the above criteria, a total of 20 participants were recruited for this study. Five were female and 15 were male. The age distribution of the participants was as follows: 30% - 18-24 years old, 60% - 25-34 years old and 10% - 35-44 years old.

There were 10 Norwegian participants and 10 nonNorwegian participants. When participants were asked about where they had initially heard about the Norwegian Online Dictionary, the response was that 50% used this dictionary in school, 40% were recommended to it by a Norwegian language teacher and 10% of the participants found out from a Google search.

Similarly, 85% of the participants used the dictionary in a web browser, 5% in app and 10% used both the platform app and browser. Lastly, when participants were asked about, whether it was easy to use the Norwegian Online Dictionary in the beginning, 60% of the participants found it difficult to use the dictionary in the beginning and the remaining 40% of the participants found it easy to use the dictionary in the beginning.

Ethical requirements for human participants were adhered to and met Norwegian standards. In brief, participants gave informed consent for their participation and no aspect of what they did or what was collected as data could identify a participant in any way.

Overall, from the details elicited above, it was felt that the participant sample recruited was a good representative sample of participants that did and would use this specific tool in real life.

C. Experimental Design

The experiment was designed with a within-users approach. This clearly meant that the participants would experience both user interfaces. This was felt to be the most suitable approach as it was important for the authors to have the participants able to make comparisons between the two user interfaces.

D. Variables

The independent variables were the two user interfaces, consisting of the prototype developed and the original dictionary user interface and the tasks used in the experiment.

The dependent variables were performance and user satisfaction.

The dependent measures were that for performance, the errors committed by participants in undertaking the tasks with the two user interfaces were recorded. An error was counted if a participant took more than one attempt to do a task, e.g. if a participant took three attempts to achieve a specific task, then two errors were counted.

User satisfaction was measured by means of a post-experiment questionnaire. The questionnaire covered several aspects to do with ease of use, ease of reading and layout.

E. Apparatus, Materials and Tasks

To design the prototype, a Huawei MateBook D laptop was used. The laptop was installed with windows 10 (64-bit operating system), 8GB RAM and an Intel core i5 processor.

The specifically designed tasks were as follows:

1) Search for the word ‘tur’ in either Bokmål or Nynorsk in both the dictionaries.
   a) Find in which dictionary it was easy to read the meaning list.
   b) Find the 3rd example of the 2nd meaning.
   c) Find the grammatical forms of the word (bøying).
2) Search for the word ‘prøve’ in both the dictionaries.
   a) Find the noun (substantive) and verb meaning list.
   b) Find the sub-meaning of the 1st meaning.
3) Search for the word ‘språk’ in both Bokmål and Nynorsk (begge) in both the dictionaries.
   a) Find in which dictionary it is easy to read the meaning list.

The above tasks were designed to be as realistic as possible and to mirror the potential use of a user using the dictionary.

F. Procedure

The recruitment process was started with an initial contact through email explaining the purpose of the study and the tasks to be performed. If the participant agreed to participate, then the recruitment criteria were checked. A suitable location was selected by the participant.

The experiment process was started by first obtaining informed consent from the participants both verbally and in written form.

Before beginning the tasks, participants were asked to complete a pre-experiment questionnaire. This covered issues about prior knowledge or experience of using the Norwegian online dictionary, age group, gender and level of Norwegian language for nonnative participants.

The next stage involved the participants carrying out the tasks, as described in the previous section. The tasks were carried out using both the prototype user interface and the Norwegian online dictionary. The ordering of which user interface was used first was rotated.

Once the tasks were completed a post-experiment questionnaire was completed by the participants. The questionnaire covered aspects of user experience of the two user interfaces. The questionnaire covered several aspects to do with ease of use, ease of reading and layout.

G. Results

Errors were recorded as participants did the tasks, as described above. The tasks involved finding the examples from the searched word, finding the grammatical forms (bøying) of the searched word, finding the different meaning list of the searched word, and finding the sub-meaning list of the searched word. For each of these aspects, errors were recorded for each user interface and then these were subjected to paired samples t-testing in order to determine if there were any statistically significant differences between the two user interfaces.

For the aspects of finding examples from the searched word, the t-test result for errors was as shown in Table I and Fig. 3.

TABLE I. ERRORS IN FINDING EXAMPLES FROM THE SEARCHED WORD

	Errors in Finding Examples From the Searched Word	
	Prototype New User Interface	*Norwegian Online Dictionary*
Mean	0	1.45
Standard Deviation	0	0.686
t(19) = -9.448, p = 0.0001, d = 2.253		

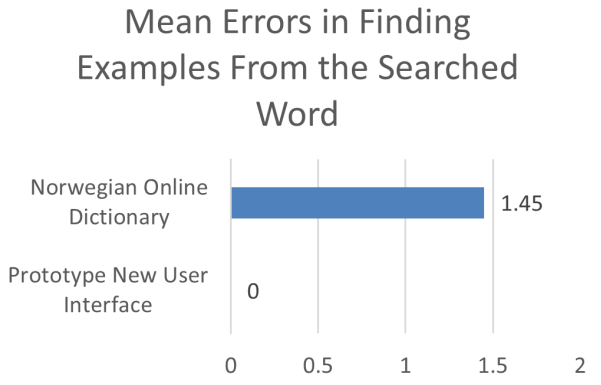

Figure 3. Mean Errors in Finding Examples From the Searched Word

This is a statistically significant result suggesting the Prototype new user interface incurred significantly fewer errors than the Norwegian Online Dictionary.

For the aspects of finding the grammatical forms (bøying) of a searched word, the t-test result for errors was as shown in Table II and Fig. 4.

TABLE II. ERRORS IN FINDING GRAMMATICAL FORMS

	Errors in Finding Grammatical Forms	
	Prototype New User Interface	*Norwegian Online Dictionary*
Mean	0	1.45
Standard Deviation	0	0.686
t(19) = -5.616, p = 0.0001, d =2.08		

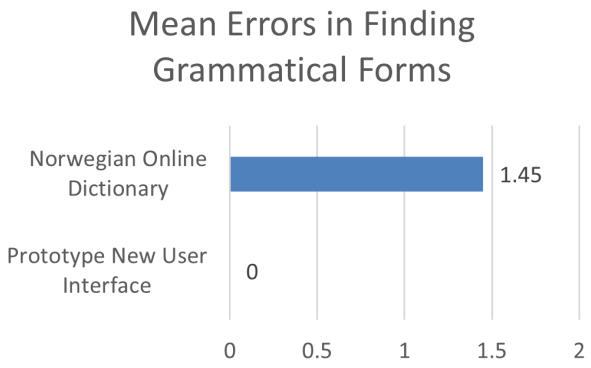

Figure 4. Mean Errors in Finding Grammatical Forms

This is a statistically significant result suggesting the Prototype new user interface incurred significantly fewer errors than the Norwegian Online Dictionary.

For the aspects of finding a different meaning list of a searched word, the t-test result for errors was as shown in Table III and Fig. 5.

TABLE III. ERRORS IN FINDING A DIFFERENT MEANING LIST OF A SEARCHED WORD

	Errors in Finding a Different Meaning List of a Searched Word	
	Prototype New User Interface	*Norwegian Online Dictionary*
Mean	0	0.42
Standard Deviation	0	0.507
t(19) = -3.618, p = 0.002, d = 0.704		

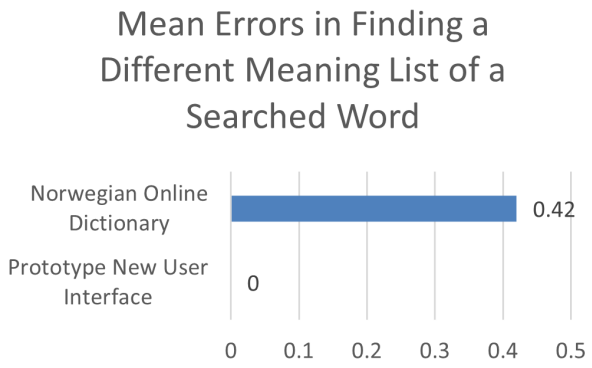

Figure 5. Mean Errors in Finding a Different Meaning List of a Searched Word

This is a statistically significant result suggesting the Prototype new user interface incurred significantly fewer errors than the Norwegian Online Dictionary.

For the aspects of finding a sub-meaning list of a searched word, the t-test result for errors was as shown in Table IV and Fig. 6.

TABLE IV. ERRORS IN FINDING A SUB-MEANING LIST OF A SEARCHED WORD

	Errors in Finding a Sub-Meaning List of a Searched Word	
	Prototype New User Interface	***Norwegian Online Dictionary***
Mean	0.25	0.5
Standard Deviation	0.444	0.607
t(19) = -2.517, p = 0.021, d = 0.294		

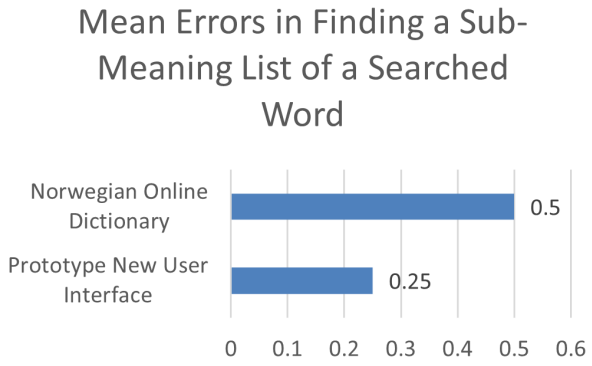

Figure 6. Mean Errors in Finding a Sub-Meaning List of a Searched Word

This is a statistically significant result suggesting the Prototype new user interface incurred significantly fewer errors than the Norwegian Online Dictionary.

Subjective satisfaction and experience were evaluated by using an eight-question questionnaire, where participants ranked their opinions for both user interfaces using a Likert type [19] scale, ranging from one to five. Each of the five numbers in the scale had the following meanings: 1 = Strongly disagree, 2 = Disagree, 3 = Neutral, 4 = Agree and 5 = Strongly Agree.

For the question concerning the meaning list being displayed with proper line spacing and being easy to read, the t-test result was as shown in Table V and Fig. 7.

TABLE V. MEANING LIST DISPLAYED WITH PROPER LINE SPACING AND EASY TO READ

	Meaning List Displayed With Proper Line Spacing and Easy to Read	
	Prototype New User Interface	***Norwegian Online Dictionary***
Mean	4.45	2.1
Standard Deviation	0.51	0.447
t(19) = 14.104, p = 0.0001, d = 2.826		

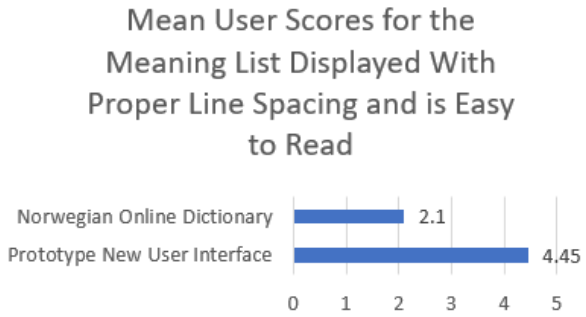

Figure 7. Mean User Scores for the Meaning List Displayed With Proper Line Spacing and is Easy to Read

This is a statistically significant result suggesting the Prototype new user interface is significantly easier to read when a meaning list is displayed along with proper line spacing.

For the question concerning the content from the word search having proper headings and labels that help to scan and find the required information, the t-test result was as shown in Table VI and Fig. 8.

TABLE VI. THE CONTENT FROM THE WORD SEARCH HAS PROPER HEADINGS AND LABELS THAT HELP TO SCAN AND FIND THE REQUIRED INFORMATION

	The Content From the Word Search Has Proper Headings and Labels That Help to Scan and Find the Required Information	
	Prototype New User Interface	*Norwegian Online Dictionary*
Mean	4.5	2.2
Standard Deviation	0.513	0.894
t(19) = 9.516, p = 0.0001, d = 2.523		

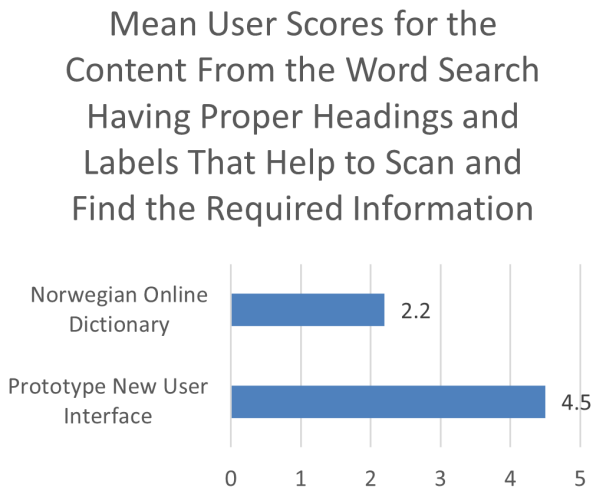

Figure 8. Mean User Scores for the Content From the Word Search Having Proper Headings and Labels That Help to Scan and Find the Required Information

This is a statistically significant result suggesting the Prototype new user interface is significantly better for finding information than the Norwegian Online Dictionary.

For the question concerning the content being cluttered and difficult to read, the t-test result was as shown in Table VII and Fig. 9.

TABLE VII. THE CONTENT IS CLUTTERED AND DIFFICULT TO READ

	The Content is Cluttered and Difficult to Read	
	Prototype New User Interface	*Norwegian Online Dictionary*
Mean	1.5	4.5
Standard Deviation	0.761	0.513
t(19) = -11.469, p = 0.0001, d = 3.366		

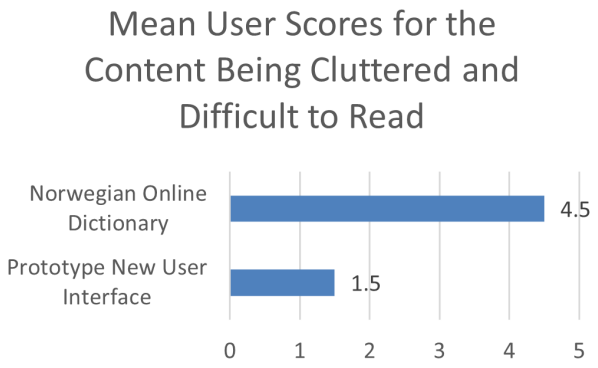

Figure 9. Mean User Scores for the Content Being Cluttered and Difficult to Read

This is a statistically significant result suggesting the Prototype new user interface is significantly less cluttered than the Norwegian Online Dictionary.

For the question concerning the sub-meanings being numbered and easy to find and navigate, the t-test result was as shown in Table VIII and Fig. 10.

TABLE VIII. THE SUB-MEANINGS ARE NUMBERED AND EASY TO FIND AND NAVIGATE

	The Sub-Meanings are Numbered and Easy to Find and Navigate	
	Prototype New User Interface	*Norwegian Online Dictionary*
Mean	4.25	1.75
Standard Deviation	0.716	0.851
t(19) = 9.05, p = 0.0001, d = 2.658		

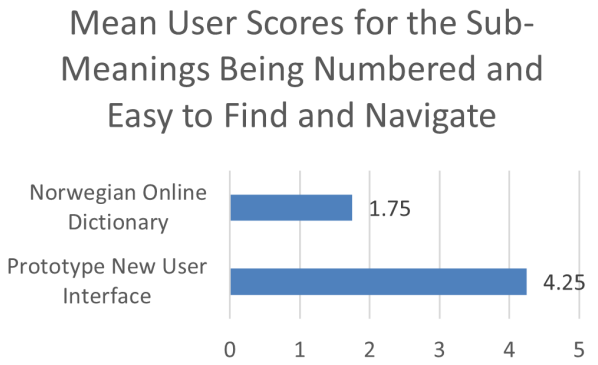

Figure 10. Mean User Scores for the Sub-Meanings Being Numbered and Easy to Find and Navigate

This is a statistically significant result suggesting the Prototype new user interface is significantly better for finding sub-meanings and navigation than the Norwegian Online Dictionary.

For the question concerning the design layout of the content being easy to understand, the t-test result was as shown in Table IX and Fig. 11.

TABLE IX. THE LAYOUT DESIGN OF THE CONTENT IS EASY TO UNDERSTAND

	The Layout Design of the Content is Easy to Understand	
	Prototype New User Interface	***Norwegian Online Dictionary***
Mean	4.45	1.65
Standard Deviation	0.605	0.587
t(19) = 11.332, p = 0.0001, d = 3.187		

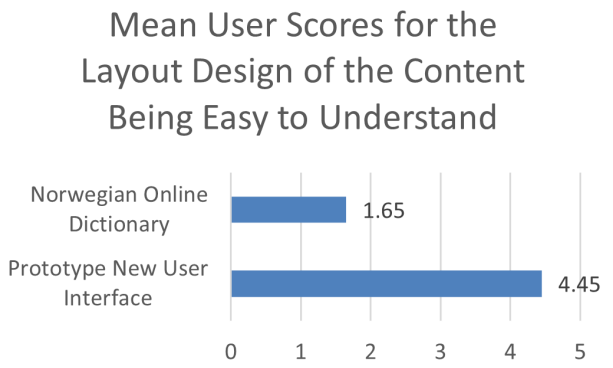

Figure 11. Mean User Scores for the Layout Design of the Content Being Easy to Understand

This is a statistically significant result suggesting the Prototype new user interface design layout is significantly easier to understand than the Norwegian Online Dictionary.

For the question concerning the grammatical forms of a word (Bøying in Norwegian) being easy to find, the t-test result was as shown in Table X and Fig. 12.

TABLE X. THE GRAMMATICAL FORMS OF A WORD ARE EASY TO FIND

	The Grammatical Forms of a Word are Easy to Find	
	Prototype New User Interface	***Norwegian Online Dictionary***
Mean	4.95	1.45
Standard Deviation	0.224	0.51
t(19) = 30.512, p = 0.0001, d = 4.542		

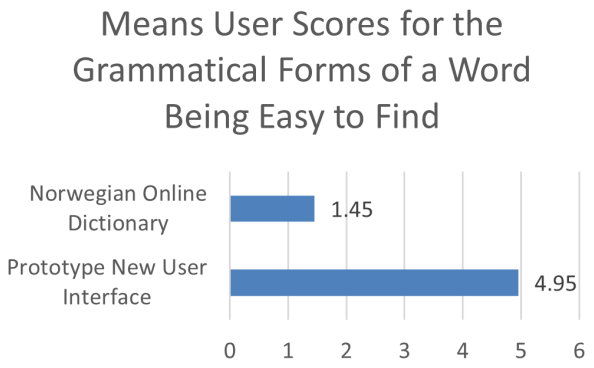

Figure 12. Means User Scores for the Grammatical Forms of a Word Being Easy to Find

This is a statistically significant result suggesting the Prototype new user interface is significantly easier for finding the grammatical forms of a word than the Norwegian Online Dictionary.

For the question concerning the examples being easy to read and find, the t-test result was as shown in Table XI and Fig. 13.

TABLE XI. THE EXAMPLES ARE EASY TO READ AND FIND

	The Examples are Easy to Read and Find	
	Prototype New User Interface	***Norwegian Online Dictionary***
Mean	4.75	1.95
Standard Deviation	0.444	0.887
t(19) = 13.161, p = 0.0001, d = 3.123		

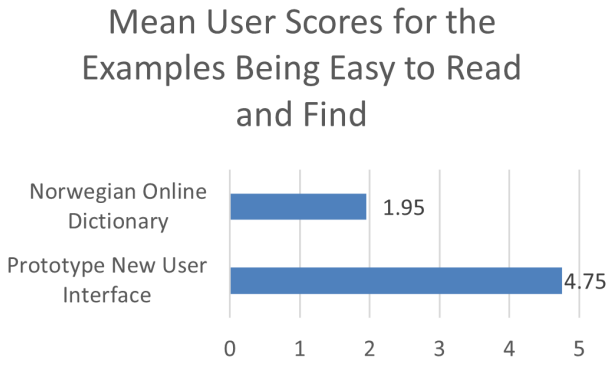

Figure 13. Mean User Scores for the Examples Being Easy to Read and Find

This is a statistically significant result suggesting the Prototype new user interface is significantly easier for reading and finding examples than the Norwegian Online Dictionary.

For the question concerning the option of a word meaning being displayed in both Bokmål and Nynorsk (Begge in Norwegian) and being easy to read and understand, the t-test result was as shown in Table XII and Fig. 14.

TABLE XII. DISPLAYING A WORD MEANING IN BOKMÅL AND NYNORSK IS EASY TO READ AND UNDERSTAND

	Displaying a Word Meaning in Bokmål and Nynorsk is Easy to Read and Understand	
	Prototype New User Interface	*Norwegian Online Dictionary*
Mean	4.45	2.6
Standard Deviation	0.51	0.681
t(19) = 16.907, p = 0.0001, d = 2.109		

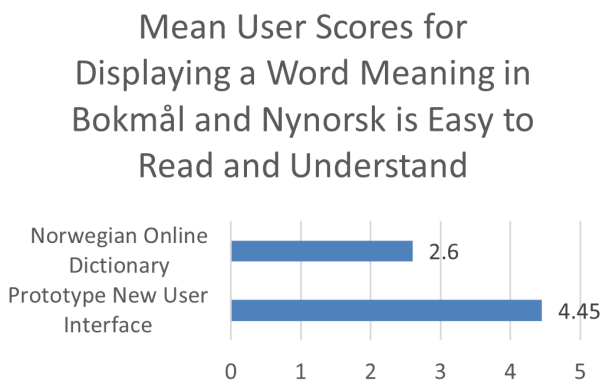

Figure 14. Mean User Scores for Displaying a Word Meaning in Bokmål and Nynorsk is Easy to Read and Understand

This is a statistically significant result suggesting the Prototype new user interface is significantly easier to read and understand for the word meanings in the two forms of the Norwegian language than the Norwegian Online Dictionary.

In the next section we will present some discussion and conclusions concerning the observed results and the overall design of the user interface.

5. DISCUSSION AND CONCLUSIONS

As can be seen from the t-testing done above all performance and user experience indicators were statistically significant in favour of the Prototype new user interface. Categorically, the participants made very few errors in their use of the prototype, while in the Norwegian Online Dictionary more errors were done by the users. The differences being statistically significant. Also, the eight user experience questions were all categorically showing a better experience with the new Prototype user interface.

As discussed in the 'New User Interface' section above, we made strong use of several existing user interface guidelines. The Universal Design principles [18] were key in the activity of redesigning the user interface (See Figure 2). Principle 1 – 'Equitable Use' and Principle 2 - 'Flexibility in Use' [18] concern a design being useful and marketable to different people with different skills. It also includes people's personal inclinations. These two aspects were specifically met by ensuring that the new user interface accommodated keyboard compatibility, screen readers, specified headings, labels, sections, contrast customization, page zoom customization, browser and device compatibility. Principle 4 – 'Perceptible Information' [18] is also linked to the implementation as Principle 4 specifically mentions screen reader technology in a software context.

The experimental evaluation results specifically showed several of the other principles to have been met within the dictionary context. Principle 3 - 'Simple and Intuitive Use' was clearly met as all the relevant user experience questions (See Results section) in the questionnaire ranked the prototype user interface as being easy to use and the fact that users made no errors during the tasks shows clearly that the user interface is 'intuitive'.

Principle 5 – 'Tolerance for Error' [18] can have wide application in software and web-based contexts. However, in our context of online dictionaries, the categorical results of users not doing any errors during the tasks, indicates that the new user interface is more tolerant to errors in the sense that the design is able to foster an interaction that is much less error prone.

In the task where participants were asked to find the specific example of the searched word, all participants were confused and had difficulties in finding examples in the Norwegian Online Dictionary, whereas in the new prototype participants found the specific example easily. This was helped by the user interface design and the bullet listing and indentation used in displaying the relevant content.

Similarly, in finding the grammatical forms (bøying) of a searched word, all the Norwegian participants could not find it in Norwegian Online Dictionary, even though they had been using this dictionary from their school days. Observation showed that they randomly clicked on different links until it was found. In the prototype, the users found it easily, as the grammatical forms were located at the end of the page (See Figure 2).

In finding a different meaning list of a searched word, participants were asked to find the noun meaning list and verb meaning list of the word 'prøve'. Errors were made by most of the nonNorwegian participants in the Norwegian Online Dictionary. All participants could easily find it in the prototype because of the improved user interface design including the heading and meaning list being specified.

In trying to find the sub-meaning of a searched word some of the participants were confused and made an error(s) in the prototype. However, the number of errors were less than in the Norwegian Online Dictionary. Some of the

participants also found difficulty in finding the meaning list in the prototype because they got confused concerning whether it was another example or a sub-meaning.

Overall, this paper makes a novel and significant contribution to knowledge in a user interface context that has not benefited by being investigated thoroughly and in an empirical manner. This is despite the fact that online dictionaries have a usage of millions of users. Our work did not deal with evaluating the content of the Norwegian Online Dictionary, but purely with the user interface. To our knowledge, this is a first in the world where a new dictionary user interface has been designed, evaluated and empirical data statistically analysed. The design presented in Figure 2 could be applied as a global template and thus improve usability and universal access to the Norwegian Online Dictionary and to other dictionaries with similar features.

Furthermore, the user interface that is designed and presented in Figure 2 is novel in several ways. The first is that the content of a search is presented along with the grammatical forms of a word using the available screen real estate in an efficient manner. The original version presented all the results in a narrow column, thus wasting much available space and potentially requiring more scrolling. Further, the grammatical forms relied on the user being able to detect that these were available via an on-screen click which in turn would display a pop-up window with the grammatical forms. This approach lacked usability and reduced universal design. Also, the new version of the user interface was more universally designed. The prototype has keyboard compatibility, can accommodate screen readers, specified headings, labels, sections, contrast customization, page zoom customization, browser and device compatibility, which were lacking in the original version. The statistical analysis presented and discussed in this paper gives clear evidence to suggest that the developed prototype is superior in terms of usability and universal design. In future work, we would recommend a wider user-based study charting the dictionary usage and user experience in daily life. We would also suggest aiming for a larger sample size in future work. In our own experience and in talking to other users of the dictionary, further study could be carried out in how words and terms are explained within the dictionary. Some nonnative Norwegian speakers also commented that the dictionary could be extended to include English/Norwegian. This could help nonNorwegians learn Norwegian better and also Norwegians learning or improving their English.

References

[1] Dictionary.com, "About," 2021, accessed March 2021. [Online]. Available: https://www.dictionary.com/e/about/

[2] Merriam-Webster.com, "About us," 2021, accessed March 2021. [Online]. Available: https://www.merriam-webster.com/about-us

[3] Language Council of Norway – Språkrådet, "Homepage," 2021, accessed March 2021. [Online]. Available: https://www.sprakradet.no/

[4] University of Bergen, "Revisjonsprosjektet," 2021, accessed March 2021. [Online]. Available: https://www.uib.no/lle/121561/revisjonsprosjektet-bokm%C3%A5lsordboka-og-nynorskordboka#brukstatistikk-for-ordb-kene-nbsp-Instituttforlingvistiske,litter\OT1\aereogestetiskestudier

[5] Wave Evaluation Tool WebAIM, "Wave evaluation tool," 2021, accessed March 2021. [Online]. Available: https://chrome.google.com/webstore/detail/wave-evaluation-tool/jbbplnpkjmmeebjpijfedlgcdilocofh/related?utm_source=chrome-ntp-icon

[6] MacMillan Dictionary, "Macmillan dictionary," 2021, accessed March 2021. [Online]. Available: https://www.macmillandictionary.com/

[7] Cambridge Dictionary, "Cambridge dictionary," 2021, accessed March 2021. [Online]. Available: https://dictionary.cambridge.org/

[8] Bokmålsordboka og Nynorskordboka, "Bokmålsordboka og nynorskordboka," 2021, accessed March 2021. [Online]. Available: https://ordbok.uib.no/

[9] J. Ling and P. van Schaik, "The influence of font type and line length on visual search and information retrieval in web pages," *International Journal of Human-Computer Studies*, vol. 64, no. 5, p. 10, 2006.

[10] R. Duchnicky and P. Kolers, "Readability of text scrolled on visual display terminals as a function of window size," *Human Factors*, vol. 25, no. 6, p. 10, 1983.

[11] M. Dyson and M. Hazelgrove, "The influence of reading speed and line length on the effectiveness of reading from screen," *The International Journal of Human-Computer Studies*, vol. 54, p. 28, 2001.

[12] J. Ling and P. van Schaik, "The effect of text and background colour on visual search of web pages," *Displays*, vol. 23, no. 5, p. 8, 2002.

[13] P. Fuertes-Olivera, "On the usability of free internet dictionaries for teaching and learning business english," in *Electronic Lexicography*, S. Granger and M. Paquot, Eds. Oxford University Press, 2012.

[14] U. Heid and J. Zimmermann, "Usability testing as a tool for e-dictionary design: Collocations as a case in point," in *Proceedings of the 15th EURALEX International Congress*, R. V. Fjeld and J. M. Torjusen, Eds. Oslo, Norway: Department of Linguistics and Scandinavian Studies, University of Oslo, Aug. 2012.

[15] SortSite, "One click web site testing," 2021, accessed March 2021. [Online]. Available: https://www.powermapper.com/products/sortsite/

[16] w3.org, "Web content accessibility guidelines (wcag) 2.1," 2021, accessed March 2021. [Online]. Available: https://www.w3.org/TR/WCAG21/

[17] Leavitt, M. O., Shneiderman, B., Bailey, R. W., Barnum, C., Bosley, J., Chaparro, B., Dumas, J., Ivory, M. Y., John, B., Miller-Jacobs, H., Koyani, S. J., Lewis, J. R., Page, S., Ramey, J., Redish, J., Scholtz, J., Wigginton, S., Wolfson, C. A., Wood, L. E., and Zimmerman, D. , "Research-based web design and usability guidelines 2006 edition," 2006, accessed March 2021. [Online]. Available: https://commons.erau.edu/publication/1028

[18] M. F. Story, "Maximizing usability: The principles of universal

design," *Assistive Technology: The Official Journal of RESNA*, vol. 10, no. 1, p. 9, 1998.

[19] R. Likert, *Technique for the Measurement of Attitudes*. Columbia University Press, NY, Jun. 1932.

Sayush Lal Shrestha holds a Master degree in Universal Design of ICT from Oslo Metropolitan University, Oslo, Norway and a Bachelor in Information Management from Tribhuvan University, Kathmandu, Nepal. Sayush Lal Shrestha currently lives in Nepal and works as a software engineer.

Dr Pietro Murano is a Computer Scientist with research specializing in all aspects of the user interface, interaction, user experience and software design (HCI and Usability Engineering). Dr Murano also teaches at the university level in Computer Science subjects. Dr Murano is currently an Associate Professor of Computer Science at Oslo Metropolitan University, Norway. www.pietromurano.org.