

# **User studies, user behaviour and user involvement - evidence and experience from The Danish Dictionary**

**Henrik Lorentzen, Lars Trap-Jensen**

Society for Danish Language and Literature, Copenhagen, Denmark  
E-mail: hl@dsl.dk, ltj@dsl.dk

## **Abstract**

Launched in November 2009, The Danish Dictionary (Den Danske Ordbog, DDO) has an online history of just over six years. During that period we have monitored and analysed the way the dictionary is used and also made various adjustments as a consequence of both our own efforts and the feedback we have received. In this presentation, we want to share with you the experience we have gained over the years. Our focus will be on four main aspects: (1) user study, (2) search engine optimisation, (3) logfile analysis and (4) user involvement.

**Keywords:** user studies; user behaviour; search engine optimisation; logfile analysis; user involvement

## **1. User study**

In 2011, we carried out a user study (Johannsen, 2011; summarised in Lorentzen & Theilgaard, 2012) consisting of two parts: an online survey supplemented by qualitative interviews with selected user profiles. In the survey, which was answered by just over 1,000 respondents, we asked who the users were, in what situations they primarily used the dictionary, and how they evaluated the dictionary on various parameters. Our main findings from this part were:

- The typical user is a well-educated man or woman in their thirties. They use the dictionary primarily for reception purposes and usually in a professional capacity. The average user is satisfied with the dictionary and is a returning customer.
- The users generally find the dictionary appealing and easy to use. They appreciate the simple layout free from advertisement, chat function, social media sharing and other functionalities that distract their attention. First-time users are less generous in their appraisal than returning users.
- The most popular look-up items among frequent users were: meaning, use and spelling. Etymology is popular with both frequent and new users. Fixed expressions, the existence of a word and

inflection appeal more to frequent users.

- Among the critical remarks, users report that they want more headwords, more synonyms, etymology and LSP vocabulary. Furthermore, they request more usage examples and would like these to be less complex.

The qualitative part consisted of interviews with 10 respondents selected from different backgrounds in terms of age (17–61 years), gender (6 male, 4 female), education and dictionary experience, each interview falling in two parts. In the first part, the users were asked to perform six different look-up tasks. While performing the tasks, they were invited to think aloud while the examiner recorded what they said and did. In the second part, the respondents were interviewed about the tasks, their strategies and their use of DDO in general.

The interviews confirmed certain parts of the online survey, for example that Google was used by 50% as their first choice when looking up words, and by 30% as their second choice. The main problems with the tasks were to do with (1) selecting the right homograph and (2) locating fixed expressions.

Following the user study, we decided to make various changes to the website, the most important being:

- In the event that a search had no match in the base, the function 'Did you mean' was copied from the right-hand column to the central field as some users were unable to find it in the original position.
- In the event that a search had no match DDO, but did match an entry in the historical dictionary *Dictionary of the Danish Language* (Ordbog over det danske Sprog), also accessible from our website [ordnet.dk](http://ordnet.dk), we decided to notify the user and create a link. Since then, we have added links to other dictionaries developed by our institution, all digitised versions of historical dictionaries such as *Meyers Fremmedordbog*, a comprehensive dictionary of foreign words from 1924, *Holbergordbog*, a digital version of the printed five-volume dictionary of the complete works of the author Ludvig Holberg (1684-1754), and *Moth's dictionary*, the first comprehensive dictionary of Danish from c. 1700 (for further details, cf. Troelsgård & Sørensen in press). In Sørensen, Lorentzen & Trap-Jensen (2015) further plans are demonstrated to extend linking to external resources.
- In the event of multiple search results due to homographs, we

decided to add very short glosses to assist the user in finding the right entry. Most of this work has been completed but is not yet implemented on the public website.

## **2. Search engine optimisation**

Given the substantial proportion of users entering the site via Google, we made a decision in 2011 to allow crawlers from search engines to index the entire content of our two dictionaries in order to improve the page rank in Google results. This affected 100,000 pages for DDO, and 350,000 pages for ordnet.dk in total (including the historical dictionary *Dictionary of the Danish Language*). The effect was remarkable: from 3,000 daily visits in September 2010, the number rose to 19,600 in January 2012, and the figure has continued to grow steadily ever since. Today, we have more than 120,000 visits on a peak day, with an average of about 100,000, placing DDO among the top 150 websites in Denmark, among reference sites only surpassed by Wikipedia.

In addition to substantially more traffic on the website, we have also observed an increasing proportion of traffic coming from Google since the indexation. In September 2010, 49% of the visits came from Google, in January 2012 the proportion had risen to 84%, and now, in January 2016, the percentage has stabilized at 90%.

## **3. Logfile analysis**

If one wants to improve a site's page rank in a search engine result, allowing indexation is just one of a series of actions that may be initiated. One also needs to look at the contents of the dictionary. In our case the most helpful tool was logfile analysis. In 2011 and 2012 we were particularly concerned with looking at no-matches. In 2010 about 20% of queries in DDO yielded no result (Lorentzen & Theilgaard, 2012). If users look for words that are not in the dictionary, there is no quick fix, but it is a different matter if they search for strings that are actually in the dictionary, only not handled by us so as to secure a match in the database. When analysing the logfiles we discovered quite a few searches for inflectional forms that we had not previously included, and we therefore decided to include first the present participle of verbs and the genitive of nouns, and later in 2012 also imperative forms of verbs. These initiatives, among others, have reduced the no-match ratio from about 20% in 2010 to 9% in 2015 (Sørensen, Lorentzen & Trap-Jensen, 2015).

We also used logfile analysis to compare user behaviour with a corpus-based approach to lemma selection (cf. Trap-Jensen, Lorentzen & Sørensen, 2014). We were interested in investigating whether corpus frequency corresponds to the way users actually search. In other words, if we write entries that are never looked up, and if on the other hand, users search for words that are not in our dictionary, this is a good argument for changing our lemma selection procedure. The study revealed that corpus frequency is in fact a useful tool in lemma selection, but it is not equally suited for all word frequencies. It is especially suitable for the 60,000 most frequent entries, whereas for less frequent words (ranked 100,000 and beyond) the advantage is negligible as the absolute frequency is so low that it does not really matter what entry is edited next.

Importantly, the study also revealed that very few entries are never consulted. In absolute numbers, only 202 entries were never looked up, corresponding to 0.2% of all headwords, and 5,000 entries had a look-up frequency of 3 or below (Trap-Jensen, Lorentzen, Sørensen 2014).

#### **4. User involvement**

Interaction with the users is a valuable source of input for changes and improvement of the dictionary. We primarily interact with our users in three different ways: (1) mail contact, (2) reports on new words and (3) social media (Twitter).

Since DDO went online in 2009 we have received a large number of emails from users. Many of the users ask specific or general questions about words and language, but a substantial number of users make comments on existing dictionary entries and suggest modifications to definitions and examples and propose inclusion of new meanings, e.g. *drone* used by laymen as a toy. However, we are forced to be quite selective about answering emails as we receive much more mail than we have resources to answer. Priority is given to emails which are of sufficiently general interest contentwise, in which case we publish them in a separate section of “Language questions”, or which lead to technical or lexicographical improvement of the dictionary.

We also invite the users to contribute to the dictionary content by suggesting new words. The site offers a form (‘new words wanted’) to be filled in with information about the new word (or meaning) in question. The data ends up in a database shared with The Danish Language Council, and to both institutions the information is helpful when considering new entries for DDO and for the official spelling dictionary.

Finally, we have started using Twitter as a way of communicating with our users. The Word of the Day is posted every morning, and quite often users add comments or likes, or retweet the post. A couple of times we have taken the opportunity to ask our Twitter followers to weigh in in specific cases where the editors are in doubt or do not agree. This is a line of user involvement that we would like to explore in the future.

## 5. References

- DDO = *Den Danske Ordbog* (The Danish Dictionary). Accessed at: [ordnet.dk/ddo](http://ordnet.dk/ddo). (12 January 2016).
- Johannsen, S. (2011). *User survey of DDO for the Society for Danish Language and Literature* (Brugerundersøgelse af Den Danske Ordbog for DSL), unpublished internal student report.
- Lorentzen, H. & Theilgaard, L. (2012). Online dictionaries – how do users find them and what do they do once they have? In: Fjeld, R.V. & Torjusen, J.M. (eds.): *Proceedings of the 15th EURALEX International Congress*, Department of Linguistics and Scandinavian Studies, University of Oslo, 2012, pp. 654-660.
- Sørensen, N.H., Lorentzen, H. & Trap-Jensen, L. (2015). Dealing with “unwanted words” in an online dictionary – a non-invasive strategy. Presentation at *eLex 2015: electronic lexicography in the 21st century: Linking lexical data in the digital age*.
- Trap-Jensen, L., Lorentzen, H., Sørensen, N.H. (2014). An odd couple – Corpus frequency and look-up frequency: what relationship? In: Kosem, I. & Rundell, M. (eds.), *Slovenščina 2.0*, Volume 2 (2014), Issue 2, Trojina, Institute for Applied Slovene, Slovenia, pp. 94-113.
- Troelsgård, T. & Sørensen, N.H. (in press). Nyt liv til støvede ordbøger [New life to dusty dictionaries]. In: Gudiksen, A., Hovmark, H. & Theilgaard, L.D. (eds.), *Nordiske Studier i Leksikografi 13* [Nordic Studies in Lexicography 13].

This work is licensed under the Creative Commons Attribution ShareAlike 4.0 International License.

<http://creativecommons.org/licenses/by-sa/4.0/>

