

The historical dictionary and the digital age: Steps of a transformation process

Nathalie Mederake

1. Media change and the dictionary

Dictionaries consist of many things at once, but first and foremost they offer a rich documentation basis for languages by providing a survey of their state via structured access to meanings and definitions. Historical dictionaries in particular reflect how the use of words has evolved. Admittedly, throughout their creation and editing process, printed dictionaries are affected by inconsistencies much more than might appear at first glance or than we would like them to be: as books, they seem to be most stable objects after all. Historical dictionaries, in particular, are end results of long-term academic projects and therefore predestined to show such inconsistencies. The reasons are manifold: an ever-changing project staff, an ever-updated corpus, the modification of metalexigraphic standards during a decades-long working process, etc. After their publication in printed form, established historical dictionaries are often retro-digitized, which is when new discussions on the transformation of lexicographic data arise. Although this causes discomfort to traditional lexicographers, the changes brought about by the world of digital media have to be met in the context of historical lexicography too.

Media change has been experienced in various forms, for example, in the early days of cinematography artists began by filming theatre plays on stage before realizing that so much more was possible in cinema (and eventually TV, etc). Similarly, in the field of lexicography, there is no need for retro-digitized dictionaries to hide behind the traditional print product anymore: these dictionaries have a high potential to evolve into much wider-range information tools in the digital world.

This world is by all means a very complex one. As we can see, crowdsourced products or user-generated dictionaries are having quite an impact on the lexicographic landscape nowadays, which may also affect retro-digitized dictionaries. We expect the information we look up to be up-to-date, and do not bother to ask who made a change of what and when. Notwithstanding doubts and reservations, this development may be useful as future (historical) dictionaries won't exist as mere physical (and thus immutable) objects anymore. One could also argue that the storytelling of words is

already happening entirely virtually. We are experiencing a golden age in terms of studying and appreciating words, and dictionaries are more accessible than ever, as Slate journalist Stefan Fatsis recently noted about Merriam-Webster¹.

Nowadays, lexicographers can benefit from the development of modern tools that expand possibilities in dealing with and exploring knowledge about language. Therefore, they have started to equip dictionaries with additional information, e.g. maps to show regional distribution or timelines to emphasize the occurrence of a word, i.e. anything that might be helpful for potential users. Indeed, the concept of a dictionary is being reinvented. However, it depends on the type of dictionary to what extent certain aspects can be taken on board in the creation of a multi-layered user interface. As the options are numerous, it can be challenging to make and actually *have a choice* that is not dictated by space availability. The new interpretation of existing media puts the lexicographer in the position of showing the public *all* collected information that became available thanks to decades-long work. However – while conciseness is valuable – constraints other than limitation on the number of characters with the paper model do exist. How is it possible to depict useful and relevant information for the user? Which access mechanisms should be implemented and what about the ongoing transformation in the first place?

2. Transformations within *Deutsches Wörterbuch*

The story of the *Deutsches Wörterbuch* (DWB) is a telltale of the transformation process in terms of the past, present and future of historical dictionaries. It is one of nearly two hundred years, and some turbulent times at that. Jacob Grimm and Wilhelm Grimm, the brothers who started to work on DWB in 1838, each had his own specific workflow and handling of the entry structure. Thus, inconsistencies existed from the very beginning. Also, the entire concept of the dictionary was influenced by numerous editors, who continued the



Nathalie Mederake

received her MA in German philology, history of law and sociology and PhD in German linguistics from the University of Göttingen. She has been editor and research associate of the *Deutsches Wörterbuch*'s revised edition at the Göttingen Academy of Sciences and Humanities since 2008, and is responsible for the production of manuscripts and prepress, leading development and digitization projects including the list of references, digital access to users and a cross-lingual vocabulary portal. Her other research topics concern the dynamics of Wikipedia entries from linguistic perspective and heading a task group on the accessibility and interoperability of retro-digitized dictionaries as part of the European Network for e-Lexicography.
nmedera@gwdg.de

1 http://slate.com/articles/life/culture/box/2015/01/merriam_webster_dictionary_what_should_an_online_dictionary_look_like.html/

Deutsches Wörterbuch

The *Deutsches Wörterbuch* (DWB, the German Dictionary) is the largest and most comprehensive dictionary of the German language in existence. Encompassing modern High German vocabulary from 1450 to the present, it also includes loanwords adopted from other languages and covers etymology, historical development of meaning, attested forms, synonyms, usage peculiarities and regional distribution within the German-speaking world. The dictionary's historical linguistic approach is illuminated by examples from primary source documents. The 32 volumes of DWB, published between 1854 and 1961, list more than 330,000 headwords in 67,000 print columns. Final steps to fully revise and update the letters A-F according to modern academic standards are underway at the Göttingen Academy of Sciences and Humanities and are due to be completed in 2016, with the new volumes published by S. Hirzel Verlag.

<http://woerterbuchnetz.de/DWB/>
<http://www.uni-goettingen.de/de/118878.html/>



The brothers Jacob and Wilhelm Grimm

Grimms' work, each leaving their own trace. The subject matter and objectives of DWB recurrently led to extensive discussions and workflow reorganization. Therefore, the first edition of this dictionary – if not the concept of a historical dictionary altogether – lived to see small but numerous transformations within roughly 120 years until its completion. The answers to what should be said and shown have been constantly subject to change in this first phase of its history.

The fact that the first edition of DWB has eventually turned into a 32-volume leviathan with a web interface, which now exists for over ten years, marks a second step in the transformation process. The volumes were interlinked and enhanced with detailed search options for the complete dictionary: apart from full-text searches, more complex searches within individual entries were made possible using an annotated database. The data was encoded in XML, meeting the guidelines of the ISO Text Encoding Initiative (TEI). Furthermore, the dictionary's data has become part of a network including other historical and dialect dictionaries of German² providing an extensive array of word-related information.

It goes without saying that this step in the transformation process is media-induced. Dictionary makers and computer linguists have adapted to the digital opportunities of lexicography. A new technology is, indeed, liberating. From a current perspective, it is more than obvious, though, that already very basic and self-evident amendments like the implementation of search options and integration into a Web interface reinvent the *dictionary system* we used to know (in print media). Regarding the consequences, Granger (2012, 10) states: "It shows that all facets of the field are undergoing a transformation so profound that the resulting tools bear little resemblance to the good old paper dictionary".

Another aspect of transformation comes into play with DWB's revision. Thoughts about a second or revised edition has already been underway when the paper version of the first edition was almost finished in the late 1950's. Through the years the dictionary had become a fundamental work regarding German philology and historical sciences. Back in the 1950's, however, only the latest volumes met eligible contemporary scientific standards. The first six volumes written by Jacob and Wilhelm Grimm in the 19th century clearly needed to be updated to fit into the dictionary's concept

as a whole. The revised edition project, starting in the 1960's, was then justified by being a so-called "repair solution" to the first six volumes. Conceptually and lexicographically tied to the other dictionary volumes, the revision project was designed to make DWB a coherent oeuvre. However, the revised edition is *not* a supplement to the first edition, but presents an adjusted storytelling form for the history of words. It is based on a new and extensive corpus that adheres to high quality standards. Its entries are structured in a succinct way in terms of etymology, notes of explanatory matter and usage, precise definitions and quotes to mark noteworthy usages throughout the centuries. Entries from the revised edition differ substantially in concept from the very first ones written by the Grimm brothers.

For the time being, the revised edition only includes the letters A-F in the printed version. Nevertheless, the importance of the first and the revised editions as sources for the history of dictionary making in Germany is unrivalled. They form a fundamental work for all users with questions on the origins of German words. However, in the case of the revised edition, a possible digital concept that may go beyond the Web interface (which already exists for the first edition) has not been a topic of discussion yet. Considering such a concept might lead to new and important developments beyond "the good old paper dictionary", while at the same time focusing on potential trouble spots: accessibility and usability of words' stories in historical dictionaries.

3. In need of new paths

It is not only the advance of search engines or search optimization that is becoming a focal point. What we see today is that digital enhancements are not concerned any more only with displaying the content of a dictionary, but with emphasizing the role of the potential (or even actual) user. Despite the fact that users of a historical dictionary in general (and DWB in particular) have not yet been the subject of considerable studies, it seems to be a worthwhile objective of a digital interface to catch the user's interest for the dictionary's content and suggest means of handling the lexical information. Lexicographers (of DWB) must be aware of the fact that, as Lew (2011, 248) states in a survey on online dictionaries of English, "without proper guidance, users run the risk of getting lost in the riches". Therefore, the goal is to fathom new ways of access that were not possible for the printed book. Not least do different kinds of media-induced performances provide

2 <http://woerterbuchnetz.de/>

