

## Large Dictionaries in the Nordic Countries – Report from a Symposium in Oslo 16–18 January 2014

The 21st annual symposium of the Nordic Association for Lexicography was attended by 26 participants from Denmark, Finland, Iceland, Norway and Sweden. The selected topic, 'Large Dictionaries in the Nordic Countries', was deliberately ambiguous, as the adjective *stor* 'large' in the Scandinavian languages can be interpreted both as 'voluminous' and as 'great; influential'. This ambiguity was reflected in the 9 papers presented, their topics ranging from history of lexicography to the digitalisation of existing lexicographic resources and the creation of new digital interfaces.

Arnfinn Muruvik Vonen, Director General of the Language Council of Norway, talked about the need for large dictionaries in an overall dictionary policy in Norway. All kinds of dictionaries make up an important part of the total written-language infrastructure of a language community. Dictionaries may be bilingual or monolingual, specialised or general. Monolingual general dictionaries include orthographic dictionaries, defining dictionaries, and large scientific dictionaries. Large scientific dictionaries represent fundamental documentation of a language and often function as works of source and reference for smaller dictionaries. One current and one planned large scientific dictionary project of Norwegian were presented and discussed: *Norsk Ordbok 2014* ('Norwegian Dictionary 2014'), covering Norwegian spoken dialects and the Nynorsk written tradition, and *Bokmålets og riksmålets ordbokverk* ('The Dictionary of Bokmål and Riksmål'), covering that segment of written Norwegian that is not Nynorsk.

Lennart Larsson from the Swedish Academy Dictionary (SAOB) delved into the long history of SAOB, the first fascicle of which was published in 1898. When finished in a few years, SAOB will comprise 38 volumes. During this span of over 120 years, its level of ambition (and accordingly, its scope) has varied considerably. In his paper, Larsson showed how four periods can be distinguished in the history of the dictionary: two periods (until c. 1920 and c. 1960–2000) when the ambition was to account for virtually all words and uses of words that were documented, and two periods (c. 1920–1960 and c. 2000 onwards) when quite strict rules of restriction steered the work. Larsson presented examples of how the varying levels of ambition are reflected both in the selection of lemmas and in the way they are addressed.

The topic of the paper by Christian-Emil Ore and Åse Wetås, University of Oslo, was 'The large dictionary as digital resource: *Norsk Ordbok* in the digital age'. During the last 12 years, 7 volumes of *Norsk Ordbok* (NO) have been published, and work is under way on the 12th and last volume. The first four volumes, covering a–h, were published in the period 1966–2001 and are only available as printed text. From volume 5 onwards, NO has been edited on a digital platform. The electronic editing system, with its pivot indexing component (the Meta Dictionary) as the hub of the combined source database, text corpus and editing system, enables the linking of each dictionary entry to the relevant source material. NO is thus an electronic scientific publication with reproducible results. The editing system was designed to facilitate web publication and interlinking with other lexical resources. As soon as the publishing rights were cleared, the first version of the web dictionary was designed and implemented with little effort. The internet platform offers wide possibilities for adding new functionality and linking data, making user-friendly interfaces and advanced search function, illustrations, maps,

etc. The focus thus far has been on increasing the user-friendliness, e.g. offering digital maps displaying the geographical distribution of an entry word's senses and its dialect forms. Also the possibility for the users to search for multi-word expressions has been implemented. According to Ore and Wetås, the solution chosen for editing NO combines the best of two worlds, meeting both the demands of the print version (short, stringent texts) and those of the electronic version (production of entries within a strict and limited editing scheme and with links between entries and source material).

Sven-Göran Malmgren, University of Gothenburg, gave a historical survey of the *Glossary of the Swedish Academy* (SAOL) with particular focus on the forthcoming 14th edition of this glossary, which is regarded as an institution within Swedish lexicography. First published in 1874 with a small word-list of about 35,000 lemmas, SAOL is now a more dictionary-like product with about 125,000 entries (13th edition, 2006). Apart from the growth in the amount of lemmas, the improvement of the semantic descriptions is worth noting. The main task of SAOL is still to function as an orthographic and inflectional norm, but since the 10th edition, the need for more definitions – and somewhat more elaborate definitions – has grown stronger. One reason for this is that SAOL is by far the most well-known Swedish dictionary; in fact, many Swedes do not know of any other dictionary at all. In the 14th edition, this trend will continue. Apart from a set of very basic (and monosemous) words, every entry word should have some kind of explanation, explicitly or implicitly. In order to achieve this within the limited space, some default conventions are necessary. E.g., a simple word such as *val* 'choice'; 'whale' may have two senses. In this case, it turns out that by far most of the compounds with *val-* belong to the sense 'choice'; this information is consequently given in the entry **val**, and references will have to be made only in case of the sense 'whale'.

Øystein Eek and Boye Wangensteen from the Norwegian publishing house Kunnskapsforlaget presented the dictionary *Tanums store rettskrivningsordbok* ('Tanum's Large Orthographic Dictionary'), which has served as the authoritative spelling dictionary of Norwegian Bokmål for 75 years. The manuscript of the dictionary has recently been converted into a modern and simple, but well-tagged database. Redesigned digital and book versions are planned for 2014. The flexible spelling system of Norwegian Bokmål will be very visible in the new user interface, with full documentation of the variant form and flexion systems, and with numerous cross-references. Eek and Wangensteen voiced the thought that the publication might inspire the Language Council of Norway to re-consider some peculiarities of the Bokmål norm. They also invited the participants to discuss the criteria for lemma selection in a modern and comprehensive spelling dictionary of a language where new compounds pop up every day.

Marita Akhøj Nielsen from the Society for Danish Language and Literature gave a paper on the layout and impact of Matthias Moth's Dictionary. In 1697, Matthias Moth (1649–1719), the highest official in the Kingdom of Denmark-Norway, started his ambitious work on a Danish dictionary. Words and sayings from all over the Danish king's realm were collected, predominantly samples of the daily language of ordinary people. Moth edited the huge material all alone, changing the layout of both macro- and microstructure. The final edition consists of two parts: a dictionary of ordinary words and an encyclopaedia. It contains 109,096 headwords. Danish forms the core of the dictionary, but is accompanied by Latin equivalents. The dictionary was not published

until 2013, but the manuscript has served as an important source for later Danish dictionaries, such as *Dansk Ordbog udgiven under Videnskabernes Selskabs Bestyrelse* ('Danish Dictionary Published by the Board of the Society of Sciences', I–VIII, 1793–1905) and *Ordbog over det danske sprog* ('Dictionary of the Danish Language', I–XXVIII, 1918–1956). In October 2013 the Society for Danish Language and Literature published a test version online, [mothsordbog.dk](http://mothsordbog.dk), sponsored by the Carlsberg Foundation.

Lene Schøsler, University of Copenhagen, talked about one of the most important bilingual dictionary resources in Denmark, the Danish-French/French-Danish dictionaries edited by Andreas Blinkenberg and Margrethe Thiele and subsequently by Blinkenberg and Poul Høybye. The first edition (Danish-French) was published in 1937, based on the huge collection of handwritten translation material compiled by Thiele. The first French-Danish dictionary was published in 1961, and a number of revised and augmented editions followed up to 1997, when it was decided to publish both dictionaries on a searchable CD. At that moment the Danish-French part comprised approximately 172,000 Danish lemmas and more than 1 million translations of words and expressions into French. The French-Danish part comprised some 83,000 French lemmas and more than 1 million translations into Danish. The dictionaries are not only large in volume, but also important because of the lexicological principles adopted and their influence on Danish lexicography. Blinkenberg introduced a principled organisation of the entries. These display a pedagogical layout with clear definitions and adequate translations in the first part of the entry, labelled the head. Exemplifications, collocations and current expressions are found in the second part, the tail. There are cross-references between the translations in the head and the exemplifications etc. in the tail. In 2012 the dictionaries were transferred to a commercial platform ([www.ordbogen.com](http://www.ordbogen.com)). In her paper Schøsler discussed a number of predictable and unpredictable problems which arise when transferring printed dictionaries to an electronic platform.

Jón Hilmar Jónsson, Árni Magnússon Institute for Icelandic Studies, presented a lexicographic project he has been working on for a number of years. The project entails a substantial analysis of lexicographic data which serves as an important source for monolingual and bilingual dictionaries. The data is also published on the web as an independent dictionary ([www.ordanet.is](http://www.ordanet.is)). The sources of the data presented in this work are, to a large extent, phraseological dictionaries published by Jón Hilmar Jónsson in the period 1994–2005, with additional data from a large corpus of texts from Icelandic newspapers and periodicals. The syntactic relations of phrases and the morphological relations in compounds are analysed for the semantic relations they exhibit, e.g. synonyms, antonyms, and different semantic fields. The analysis presupposes unambiguous entries, and in this manner the list of entries can contain both headwords and phrasal (or multiword) entries. Multiword entries, which are especially prominent in verbs, are also analysed grammatically, facilitating a flexible order in the semantic classification in the online dictionary. The database contains over 300,000 entries, almost two-thirds of which are entries in the online dictionary. The distinctive feature of the dictionary is the prominence of phrases and multiword expressions, unusual in this type of work. The onomasiological aspect is in the foreground, strengthening the prerequisites for the definitions of both words and phrases.

Klaas Ruppel and Caroline Sandström from the Institute for the Languages of Finland gave a paper on large Finnish dictionaries in a historical perspective, discussing a number of works that are large in size and/or have had a large impact. The first dictionary of the Finnish language of greater importance was Ganader's *Nytt Finskt Lexicon* ('New Finnish Lexicon') from 1787. This dictionary remained as a manuscript, and was published in facsimile only in 1937–1940 and in printed version 1997. The second great Finnish dictionary, *Suomalais-Ruotsalainen Sanakirja* ('Finnish-Swedish Dictionary'), was compiled and published by Lönnrot 1866–1880; it had some 240,000 articles. The process of making and publishing dictionaries was an important part of the national movement in Finland during the 19th and 20th centuries. Work on a dictionary of the Finnish dialects started already in the 19th century; a large archive of materials with Finnish dialectal words was compiled both by professionals and through crowdsourcing. This archive now consists of some 8 million items. In 1985 the first part of the Finnish dialect dictionary was published, and since 2012 it has been published by the Institute for the Languages of Finland as an electronic dictionary. For the contemporary Finnish language three large descriptive and normative dictionaries can be identified: *Nykysuomen sanakirja* ('Dictionary of Modern Finnish', 200,000 entries) was published 1951–1961. It was followed by *Suomen kielen perussanakirja* ('Basic Dictionary of Contemporary Finnish') in the 1990s and by *Kielitoimiston sanakirja* ('Dictionary of Contemporary Finnish') in the beginning of the 21st century. Each of the two latter dictionaries consists of some 100,000 entries.

The symposium ended with a lively discussion of several of the points made during the individual papers. The topic of the next symposium, due to take place in the beginning of 2015, will be 'Dictionaries of Language Varieties in the Nordic Countries'. More information on this will be posted on the website of the Association ([www.nordisksprogkoordination.org/nfl](http://www.nordisksprogkoordination.org/nfl)).

Sturla Berg-Olsen  
Norsk Ordbok 2014, University of Oslo  
Postboks 1021 Blindern  
NO-0315 Oslo  
Norway