STATE-OF-THE-ART ON MONOLINGUAL LEXICOGRAPHY FOR NORWAY (NORWEGIAN BOKMÅL AND NYNORSK)

Oddrun GRØNVIK
University of Oslo, Faculty of Humanities

Sturla BERG-OLSEN, Marit HOVDENAK, Knut E. KARLSEN
The Language Council of Norway


Monolingual lexicography for Norwegian started some decades after political independence from Denmark in 1814. Since 1885 two written standards have been recognized, one based on Danish as spoken in Norway (today Bokmål), and one based on the Norwegian vernacular (Nynorsk). Both are fully described in major scholarly dictionaries, now completed and freely available on the web. Both receive some public funding, with a view to further development. Because of frequent orthographic revisions, at first aimed at bringing the written standards closer to each other, spellers dominated the market through most of the 20th century. Today linguistic stability is aimed for, incorporating only such changes in the written standards as are supported by general usage. The first general monolingual defining dictionaries Bokmålsordboka and Nynorskordboka, covering the central vocabulary of each written standard, were first published as parallel volumes in 1986, and are now undergoing revision at the University of Bergen in cooperation with the Language Council of Norway. These dictionaries are now stored in databases, are available on the web and as a free smartphone app. Public funding of monolingual mother tongue lexicography is seen as an investment in essential linguistic infrastructure, as is bilingual lexicography between the Nordic languages and Norwegian, while other bilingual lexicography is dealt with by private publishers.

Keywords: Norwegian, Scandinavian languages, Linguistic infrastructure
1 INTRODUCTION

Monolingual lexicography has a short history in Norway, compared to many other European countries. This has to do with the history of standardizing modern Norwegian. After the political separation from Denmark in 1814, Danish remained the chief written standard language until about 1900, although the lexicon could be seen to diverge from Danish. However, although Norwegian and Danish are cognate languages, Danish was unsatisfactory as the standard language both for political reasons and because of its remoteness from the Norwegian vernacular. The Norwegian vernacular was documented during the 19th century, and a written standard proposed, which in 1885 was given equal status to Danish by the Norwegian Parliament. Since then two written standards have been in use for Norwegian, now named Bokmål and Nynorsk.

From after 1900 and until about 1970, the aim of official language policy in Norway was to bring the two standards together in one common language, and this was to be brought about through orthographic reform. Since 1970, official policy has shifted to maintaining the two standards on equal terms, and respecting the linguistic choices expressed in general usage (Vikør 1995, p. 204).

As a consequence of a language policy expressed in the orthography, the Norwegian focus for a long time was centered on spellers (simple orthographic dictionaries listing only headwords, part of speech and occasionally additional information in order to distinguish between homographs), and not on language description through monolingual lexicography, centered on definitions and documented usage. Norwegian is therefore well equipped with spelling dictionaries at all levels. The first monolingual dictionaries to give a full description of Bokmål and Nynorsk, with definitions, usage examples and the full apparatus of a modern dictionary, were Bokmålsordboka (Landrø and Wangensteen 1986) and Nynorskordboka (Hovdenak et al. 1986), which were produced in parallel and published as twin dictionaries in 1986.

Scholarly monolingual lexicography started in Europe about 1800 with the advent of comparative language studies, necessitating collections of instances of language use, where the collections are maintained for quality control, reference and new use. Scholarly monolingual lexicography came late to Norway, starting about 1920, with the collection of materials for Norsk riksmålsordbok
NRO documented Bokmål as expressed in literature written in Norway after 1814 and as used by the educated classes early in the 20th century. NRO had two additional volumes published in 1995. A revised and updated edition incorporating all former materials was published as a web dictionary at the end of 2017. This dictionary, renamed Det Norske Akademis ordbok (NAOB), aims to give wide coverage of the Bokmål vocabulary in a current orthography.

Scholarly lexicography for Nynorsk started with the work of Ivar Aasen (1813–1896) and is most fully expressed in his dictionary of 1873. But Aasen’s lexicographic work is bilingual – definitions are expressed in Danish. In 1929 efforts started to collect materials for a large monolingual dictionary covering the Nynorsk written language and the Norwegian vernacular from the whole country. The result is Norsk Ordbok (NO) I–XII (1966–2016) (cf. Grønvik, 2017).

2 GENERAL MONOLINGUAL DICTIONARIES AVAILABLE FOR NORWEGIAN

If we regard “a general monolingual dictionary” as a dictionary fully equipped with definitions and usage examples, then there are four general dictionaries for Bokmål (Norsk riksmålsordbok; Guttu 2005; Landfald 2006; Landrø and Wangensteen 2006) and two for Nynorsk (Norsk Ordbok; Hovdenak et al. 2006).

The dominant e-dictionaries are Bokmålsordboka (based on Landrø and Wangensteen 2006, updated on the web since 2015) and Nynorskordboka (based on Hovdenak et al. 2006, updated since 2012). Both are available on <ordbok.uib.no>. A revised and extended version of Guttu 2005 is available through the commercial dictionary service Ordnett. The web dictionary NAOB (based on Norsk riksmålsordbok and Guttu 2005) has been available online since 21 December 2017.

The dictionaries that come the closest to serving as national symbols are Aasen (1873) and Norsk riksmålsordbok, owing to their status as pioneer dictionaries of a language form, but these dictionaries are no longer used for current reference as the content is outdated.
3 DICTIONARY USAGE

Search numbers for Bokmålsordboka and Nynorskordboka have been high for years, rising sharply from 2004 to the present day. There are no long-term statistics for preferred reference dictionary, but the Norwegian data from the European survey of dictionary use and culture, received last year, confirm a tendency seen for several years in the logging system: the majority of dictionary users prefer to use the portal for Bokmålsordboka and Nynorskordboka, and they prefer to search both dictionaries together, rather than the one corresponding to their chosen form of Norwegian. It seems that many people like looking at the object – the Norwegian language – from more than one angle.

4 SPELLING DICTIONARIES – THEIR USE AND STATUS

Most people are acquainted with spellers from school and work. There are many different ones. Two dominant spellers are Tanums store rettskrivningssordbok (Bokmål) and Nynorsk ordliste. These have been market leaders since the 1950s, though too comprehensive for elementary school use. The Language Council publishes a web list of authorized school spellers which currently has 13 spellers for Bokmål and 5 for Nynorsk (The Language Council, 2017).

5 SPECIAL MONOLINGUAL DICTIONARIES FOR SCHOOL, FOREIGN LEARNERS ETC.

There are printed spelling dictionaries designed for school use, but no especially adapted defining dictionaries. There are several electronic spell checkers. The major commercial resources provided to schools are Ordnett and Clarify (formerly known as iFinger). Other resources used in schools are Clue, CD-ord and Lingit. The defining dictionary portal <ordbok.uib.no> is used in secondary education.

Dictionaries for Norwegian as a foreign language exist in electronic form. The bilingual LEXIN series is aimed at immigrants and covers 18 languages. They are designed as unidirectional teaching dictionaries with Norwegian as the source language and the other language as the target language, but could probably also be used by Norwegians trying to master the target language. Use is free of charge.
Inter-Nordic communication is encouraged through the dictionary website Islex, which covers the Nordic language pairs through small bilingual bidirectional dictionaries. Use is free of charge. A children’s dictionary for the Nordic languages, including Bokmål and Nynorsk, Nordisk miniordbok, was published some years ago, without full coverage of the orthography.

Bilingual dictionaries for the languages generally taught as second and third languages in school (English, German, French, and Spanish, etc.) are generally aimed at the Norwegian market, but also used by foreign learners of Norwegian. This segment of dictionary production lies under the domain of commercial publishing.

6 FUNDING OF NORWEGIAN MONOLINGUAL LEXICOGRAPHY

In the mid-sixties, it was recognized by the Norwegian government that the Norwegian language market was too small for large monolingual lexicographical projects to be privately funded. A research institute for lexicography was proposed (Gundersen, 1967, p. 120), and established at the University of Oslo in 1972. Since 2016, this activity, with the Norwegian Language collections, has been located at the University of Bergen. The multivolume scholarly dictionaries Norsk Ordbok and Norsk riksmålsordbok were both publicly funded, and the materials are now part of the Language Collections at the University of Bergen. Det Norske Akademis ordbok is owned by a private publishing company, Kunnskapsforlaget, but has received considerable public funding. The rationale for public funding is that the large scholarly dictionaries serve as reference points for (smaller) commercial ventures within lexicography. Public funding of mother tongue lexicography is in short seen as a political investment in linguistic infrastructure. The full form registers for Bokmål and Nynorsk – Ordbanken (The Word Bank) – can be downloaded for free.¹

Since the collections as well as the dictionaries were created with public funding, the public is offered access to the digitized parts of the underlying collections as well as the dictionaries themselves. The collections are heterogeneous, covering ca. 140 years of documentation, and include digitized slip archives, corpora, retro-digitized dictionaries of different kinds, dialect records and maps. These materials are organized through a common index, the Meta

Dictionary (cf. Grønvik and Ore, 2018), and can be searched item by item or through the Meta Dictionary at the website of the Language Collections.²

The scholarly dictionaries *Bokmålsordboka* and *Nynorskordboka* are publicly funded and owned by the Language Council and the University of Bergen. They are maintained by the Language Collections unit at the University of Bergen. A major five-year revision (2018–2023) project for *Bokmålsordboka* and *Nynorskordboka* has been granted funding by the Ministry of Culture, and an editorial staff is established at the University of Bergen. These very popular dictionaries have had no thorough overhaul of contents since the first publication in 1986, more than thirty years ago, so this project has been well received.

The Language Council has also published a pair of spellers (*Bokmålsordlista* and *Nynorskordlista*), but these have not been kept updated. Other spellers are published commercially.

An important issue is whether monolingual dictionaries are trusted to tell the truth about how a language is written and used in the linguistic community. The trust level for Norwegian monolingual dictionaries has not been looked into, but the portal ordbok.uib.no seems to enjoy a high level of usage and trust.

7 AVAILABLE DICTIONARY FORMATS

Both paper versions and web versions are available for *Bokmålsordboka* and *Nynorskordboka*, but the web versions are the updated ones. Minor corrections are made on a day to day basis, while revised alphabet sections will appear according to a publication plan for the total ongoing revision project. The DWS for the database system used by these two dictionaries has style sheets for paper production. If new paper editions are judged financially viable, it is possible to produce a paper edition.

*Bokmålsordboka* and *Nynorskordboka* are available through a browser³ and through an app (“Ordbokene”). The browser version has responsive design. Ordnett⁴ also provides an app for accessing their dictionaries. Some

---

² [https://www.uib.no/ub/101277/spr%C3%A5kinfrastruktur-norsk-spr%C3%A5k-og-norsk-e-namn](https://www.uib.no/ub/101277/spr%C3%A5kinfrastruktur-norsk-spr%C3%A5k-og-norsk-e-namn)
³ [https://ordbok.uib.no/](https://ordbok.uib.no/)
⁴ [https://www.ordnett.no/](https://www.ordnett.no/)
dictionaries are produced only for the digital medium. This goes for both; the Islex\(^5\) and the LEXIN\(^6\) series exist only as electronic resources. In addition, Det Norske Akademis ordbok\(^7\) is available as an electronic publication.

The dictionary writing system (DWS) for *Norsk Ordbok*, *Bokmålsordboka* and *Nynorskordboka* was in 2014 defined as a research result at the University of Oslo and registered with Inven2, the innovation company of the University of Oslo.

8 WEB DICTIONARIES AND THE INTERNET IN NORWAY

Internet is in general use throughout the country although coverage may be uneven due to Norwegian topography. Internet coverage as of today ranges from 2G to 4G+ (Telenor, 2017), and public policy is that there should be full coverage in or near permanent dwellings and places of work. At present 97% of the population aged 9–79 have internet access (Medienorge, 2017), while on average 87% access internet every day (SSB, 2017).

9 ACCESS TO ELECTRONIC DICTIONARIES FOR NORWEGIAN – FREE OR PAY SERVICE?

*Bokmålsordboka* and *Nynorskordboka* at the dictionary portal ordbok.uib.no are freely available, as are most of the other electronic resources published by the Language Collections at the University of Bergen. The scholarly dictionary *Norsk Ordbok*,\(^8\) which covers all Norwegian dialects as well as the Nynorsk standard language, is freely available on the web from the letter I and to the end of the alphabet, the last letter being Å. A revision and upgrading project for the web edition of Norsk Ordbok A–H is planned, and has received preparatory funding from 2019 from the Ministry of Culture.

There are also some pay services, the most important ones being ordnett.no and Clarify, which represents a number of Norwegian dictionary publishers (the latter company recently took over iFinger in Norway).

\(^5\) http://islex.is/nn
\(^6\) http://lexin.udir.no/
\(^7\) https://www.naob.no/
\(^8\) http://no2014.uib.no/perl/ordbok/no2014.cgi
Neither Ordnett nor Clarify publish cost details unless you approach them as a customer, and Clarify only accept institutions as customers. Both Ordnett and Clarify have a wide range of bilingual dictionaries in addition to the monolingual ones, and deliver dictionary services to many schools and local authorities.

10 DICTIONARY USE AS A SCHOOL SUBJECT

The national school curriculum for Norwegian states that students by year 7 should have acquired the skill of using different kinds of digital and paper-based dictionaries (Udir, 2013). Also see comments in the section for dictionary use below.

11 LANGUAGE CORPORA – PUBLIC FAMILIARITY AND USE

Most Norwegian language corpora require authentication before use, which effectively puts them beyond most people’s reach, though some are freely accessible. Anyone who takes an interest in language documentation will have heard the word “corpus”, but very few outside the academic community will have used one.

There are some exceptions. One is Nynorskkorpuset, a corpus of Nynorsk of 105 mill tokens covering the period 1866–2013. This is available to the general public through the website of the Language Collections at the University of Bergen. Nynorskkorpuset is also available through SketchEngine. Another exception is HaBiT Norwegian Web Corpus 2015 Bokmål (1.18 billion tokens) and Nynorsk (55 mill tokens) launched by the Text Laboratory, University of Oslo in cooperation with Masaryk University in Brno, the Czech Republic within the framework of the HaBiT project. These corpora are also available through SketchEngine.

12 LEXICOGRAPHY AND CROWDSOURCING IN NORWAY

If one considers all unsolicited contributions as a form of crowdsourcing, then there is a certain amount of support, especially in connection with dialect web sites, where people pass on information about words, expressions and usage. One instance is the website “Talemålet i Valle og Hylestad” (“The vernacular of
Valle and Hylestad), which has more than 13,000 entries in a simple but consistent dictionary format, is competently moderated and frequently updated.

Norway has strong language associations (because of our two competing standards and their different political and social affiliations). Many local branches have at some point engaged in dialect lexicography, as have branches of other national organisations.

No initiatives of importance concerning slang dictionaries or urban dictionaries are generally known or in use, though attempts have been made.

13 THE STATUS OF LEXICOGRAPHY IN GENERAL IN NORWAY

Lexicography is recognized as a specialized and necessary discipline both in relation to the mother tongue and in learning foreign languages. There is growing recognition of lexicography as a source of language and culture documentation in schools, business and public administration.

The awareness of lexicography use at a more advanced level is limited – judging by the letters the dictionaries get, people tend to look at dictionaries as a source of orthographic information rather than a source of precise concept description, though the number of comments on definitions and usage examples is growing. The assumption that a word has to be recorded in dictionaries in order to be safe to use motivates a great deal of the contact between dictionary users and dictionary staff.

14 RESEARCHING DICTIONARY USE IN NORWAY

Little has been done in Norway so far regarding gathering information about dictionary use. The Language Council did a survey of the use of web dictionaries in elementary and secondary schools in 2014 (Ims, 2015). This survey had responses from roughly 500 teachers and 5000 school pupils. The most important findings were that dictionary use increases with age, that secondary school pupils in programs preparing for tertiary education used dictionaries weekly or more often, and that Bokmålsordboka and Nynorskordboka were the best known resources. The report is published only in Norwegian (The Language Council, 2015). An earlier study (Karlsen and Rødningen, 2008) surveys the

---

9 http://www.vallemal.no/frames.jsp?menu=3
use of monolingual dictionaries in Norwegian mother tongue teaching in upper-secondary school, in the light of curricula for compulsory school and upper-secondary school. This survey shows that 86 percent of the teachers provide training in dictionary use, but that attention is mainly given to the formal aspects of language (orthography). An investigation of the textbooks used in mother tongue teaching shows that dictionary use is not a subject, not even where the subject matter strongly supports the use of reference works.

15 CONCLUSION

Mother tongue lexicography has come a long way in Norway in the last fifty years. The chief developments have been getting the first modern monolingual definition dictionaries, computerizing the major parts of the Language Collections, computerizing lexicography as a discipline and bringing out Norsk Ordbok (NO) and Det Norske Akademis Ordbok (NAOB). With these projects completed, together with the 2015 and 2012 updates of Bokmålsordboka and Nynorskordboka through the dictionary portal ordbok.uib.no, the Norwegian public has an adequate online selection of modern defining dictionaries, and Norway has quality information on the Norwegian language to offer for commercial and scholarly development. Access to language corpora will hopefully stimulate interest in language exploration and usage.

A major revision project is afoot from the beginning of 2018 for Bokmålsordboka and Nynorskordboka, which have had no thorough overhaul of contents since the first publication in 1986, more than thirty years ago. It would be natural to look at the needs and wishes of different user groups in connection with this project, and also to develop better systems for registering and specifying use, and developing usage statistics.

REFERENCES

Dictionaries:


Clue. Retrieved from https://clue.no/


Islex. Retrieved from http://islex.is/


Lingit. Retrieved from https://lingit.no/

NAOB. See Det Norske Akademis ordbok.

NO. See Norsk Ordbok.


NRO. See Norsk riksmålsordbok.


Other:


The Language Collections (2016). Available at https://www.uib.no/ub/101277/spr%C3%A5kinfrastruktur-norsk-spr%C3%A5k-og-norske-namn


STANJE ENOJEZIČNE LEKSIKOGRAFIJE: NORVEŠKA (BOKMÅL IN NYNORSK)

Enojezična leksikografija se je na Norveškem razvila nekaj desetletij po politični odcepitvi od Danske leta 1814. Od leta 1885 sta veljavna dva pisna standarda, eden temelječ na danščini (bokmål) in eden na norveškem dialektu (nynorsk). Oba sta opisana v večjih akademskih slovarjih, ki so zdaj prosto dostopni na spletu. Oba standarda sta deležna javnega financiranja, z namenom nadaljnega razvoja. Zaradi pogostih pravopisnih reform, ki so bile sprva namenjene predvsem priблиževanju obeh pisnih standardov, so večino 20. stoletja na trgu prevladovali predvsem pravopisni priročniki. Danes je cilj jezikovna stabilnost, tako da so v pisnem standardu sprejete zgolj spremembe, ki so utemeljene v splošni rabi. Prva splošna normativna slovarja, Bokmålsordboka in Nynorskordboka, ki sta pokrivala osrednje besedišče obeh standarov, sta izšla vzporedno leta 1986, zdaj pa sta v pregledu na Univerzi v Bergnu, v sodelovanju z Norveškim jezikovnim svetom. Slovarja sta zdaj zapisana kot podatkovni bazi in sta na voljo na spletu ter v obliki brezplačne aplikacije za mobilne telefone. Javno financiranje enojezične leksikografije materinščine je razumljeno kot naložba v kritično jezikovno infrastrukturo, podobno pa velja tudi za dvojezično leksikografijo v povezavi med norveščino in drugimi skandinavskimi jeziki, medtem ko za druge dvojezične slovarske projekte skrbijo zasebni založniki.

Ključne besede: norveščina, skandinavski jeziki, jezikovna infrastruktura

To delo je ponujeno pod licenco Creative Commons: Priznanje avtorstva-Deljenje pod enakimi pogoji 4.0 Mednarodna. / This work is licensed under the Creative Commons Attribution-Share-Alike 4.0 International.

https://creativecommons.org/licenses/by-sa/4.0/