Revision and Digitisation of the Early Volumes of *Norsk Ordbok*: Lexicographical Challenges

Sturla Berg-Olsen, Åse Wetås
Norsk Ordbok 2014, University of Oslo
sturla.berg-olsen@iln.uio.no, ase.wetas@iln.uio.no

Abstract

2014 will see the work on the 12th and final volume of the academic dictionary *Norsk Ordbok* (NO) finished. Still, the dictionary will remain heterogeneous due to variation in editorial practice throughout the project and incomplete in the sense that its early volumes are not digitally available. The online version of NO currently only covers the alphabet from the letter *i*. This paper describes the present state of the different parts of NO and argues that the early volumes of the dictionary must be revised and digitised to bring them up to the standards of the rest of the work. The revision and digitisation will not only give the dictionary a unitary profile but also make it possible to use it for a number of other purposes and facilitate the continuous process of keeping the dictionary up to date. The paper discusses some of the lexicographical challenges involved in the planned revision project and displays examples of the changes that must be made to the structure of the early material. It also touches upon questions concerning project organisation and funding.

Keywords: Academic dictionaries; Digitisation; Project planning

1   The history and present status of *Norsk Ordbok*

Norsk Ordbok (NO) is an academic dictionary covering Norwegian Nynorsk and all Norwegian dialects. The dictionary will provide a scholarly and exhaustive account of spoken Norwegian and of texts written in Nynorsk from 1860 up till today, and is to be completed during 2014, the year of the bicentenary of the Norwegian constitution. From 2002 the dictionary work has been organised in the time-limited project organisation Norsk Ordbok 2014 (NO 2014). The project owner is the Department of Linguistics and Scandinavian Studies at the University of Oslo. In 2014, the finished work will include more than 300,000 entries, published in 12 volumes.

When NO was conceived in the late 1920s, Nynorsk was still a written language in the making, and the standard was continuously fed by Norwegian dialect words. The proponents of Nynorsk wanted to make a comprehensive scholarly dictionary building on the works of the famous Norwegian 19th century linguist Ivar Aasen. The immediate goal behind the dictionary was to develop Nynorsk further, and to raise the prestige of the new written standard. The combination of dialects and written standard in one dictionary – somewhat unusual in a wider European context – was considered a natural

choice, given the crucial role dialect data had played in the codification of Nynorsk from the outset. Even today the editors of NO regularly write entries entirely based on dialect material. This process often includes codifying the spelling and inflection of these words according to the Nynorsk standard.

The collection of data for a new and comprehensive dictionary of Nynorsk started in 1930. A dictionary board of trained lexicographers instructed and supervised more than 550 volunteers, who during these early years collected dialect data from all over the country and made it possible for the dictionary board to build up a huge slip archive. The learned dictionary board also supervised the extraction of literary excerpts from Nynorsk literature, both fiction and non-fiction. In addition, they compiled a draft manuscript combining Ivar Aasen's dictionaries (Aasen 1850 and 1873) with a range of other canonical dictionaries dating from 1870 to 1910, also adding data from glossaries and local dictionaries dating from 1600 to 1850 (Skard 1932). This draft manuscript for the new, academic dictionary was finished by 1940.

The editing of the dictionary started in 1946, and the first volume of NO was published 20 years later, covering the alphabet from the letter *a* to the adjective *doktrinær*. The original plan was to make a 2–3 volume dictionary, but in 1966 the chief editor estimated that 8–9 volumes would be needed to cover the whole alphabet (Hellevik 1966). During the first 50 years the editing of the dictionary progressed slowly. At the same time the source material grew, and so did the dictionary entries in volumes 2, 3 and 4. All the work was done manually, the slips sorted on the lexicographer's desk and the manuscripts prepared in handwriting.

In 2002 the work was reorganised and moved to a digital platform, making the editing process a lot more efficient. Increased funding allowed the project to employ more editors, and the work gained speed. During the period 2005–2013 7 volumes were published, with the last volume to be finished in 2014. However, the volumes produced before 2002 (volumes 1–4 and roughly half of volume 5) remain only partly digitised and show a number of discrepancies compared to the latter volumes. This has to do with changes in editorial practices that were implemented along the way. The digitisation of the volumes produced before 2002 and the revision of the contents of these volumes to bring them up to date are essential tasks that must be undertaken after the completion of the last volume. This will ensure that NO is a homogeneous dictionary that meets the scholarly standards of the age of electronic corpora and can be updated continuously in the future. Only when the entry database covers the whole alphabet can it be used for other purposes (e.g. the extraction of semantic structures to form the basis of a Nynorsk word net, the extraction of subsets of entries for new, thematic dictionaries etc.). In addition, revisions of the entry database itself can then be organised thematically, and not necessarily alphabetically.

Since 2012 an online version of NO has been available, but this version only contains the material from the letter *i* onwards. A complete online version is dependent on the complete digitisation of the early material and adaptation of this material to the database system used.

The reorganisation into the time-limited project NO 2014 also led to a change in profile for the dictionary. During the whole history of the dictionary, there has been a strict emphasis on constructing a

scholarly work that meets scientific demands. However, the editorial profile of the earliest volumes was that of a scientific paradigm still concerned with nation-building. The dictionary was part of the work to document and elaborate on Nynorsk as a cultural object and to further standardise this language, which was still in its formative stage. In the modern project organisation, the emphasis is on editorial practice as descriptive research work. This results in the inclusion of entries that were earlier not considered part of Nynorsk proper, but which have entered Nynorsk during the last 50 years. To take one example, during the work on the entries starting with the Norwegian privative prefix *u-* a number of instances were discovered where the 'positive' counterparts of these words from the earliest parts of the alphabet were not covered (words like *bekvem* 'comfortable', *bekymra* 'worried', *bemanna* 'manned' etc.). These are loan words from Danish and German that were earlier only used in Bokmål, but they are now also part of modern Nynorsk and should thus be included.

2   The microstructure of *Norsk Ordbok*

The microstructure of NO entries is fairly similar to that found in other comprehensive scholarly monolingual dictionaries, such as the OED, the *Dictionary of the Danish Language* (ODS) and the *Swedish Academy Dictionary* (SAOB). Each headword is followed by a section containing information on early lexicographical sources listing the word and etymology, as well as pronunciation (mainly for borrowed words) and alternative written forms of the word. This section also provides attested dialect forms of the word with geographical indications. Only dialect forms that do not follow automatically from general and well-known rules of sound correspondences in Norwegian dialects are included. The introductory section is the part of the NO entries which has seen the most variation and change during the project period. In the early volumes there was a certain degree of experimenting both with the order of the information given here and the structuring of this information. The digital platform used from 2002 onwards ensures stringency, but the variation found in the introductory section in the early volumes presents big challenges when it comes to digitisation.

The part of the entry following the introductory section is fairly straightforward, with potentially three explicit levels of senses, each sense customarily followed by literary sources and/or geographical indications, as well as examples of usage. In the early volumes, multi-word expressions are treated largely on a par with ordinary examples. Starting from the letter *i*, such expressions have been edited as sublemmas, appearing in boldface.

3   Challenges involved in the digitisation and revision of the early volumes

The goal of the revision project is to bring volumes 1–5 up to the same standards and give the entries in these volumes the same structure as that found in volumes 6–12. The contents of volumes 1–5 must be evaluated in view of the present editorial policies and revised on all levels where necessary in order to reflect these policies. This involves restructuring, adding information and also (particularly in volumes 3–4) removing some information. The result will be a homogeneous product reflecting the Nynorsk of the 21st century as well as the history of this written standard and the diversity of the Norwegian dialects.

There are several possible ways of digitising the oldest volumes of the dictionary. One solution could be OCR-scanning. This process was chosen for the first online version of the *Swedish Academy Dictionary* (SAOB) in 1997, but the result was considered unsatisfying and also turned out to be very expensive (Mattisson 2012). SAOB is currently going through a second re-digitisation process. This time the printed text is punched and stored in digital files in China. When this part of the process is finished, the SAOB editorial staff themselves will process the files by hand into valid XML. A similar process was chosen by the Society for Danish Language and Literature when they digitised their 28-volume *Dictionary of the Danish Language* (ODS) in 2005 (cf. ODS FBTS). The solution chosen for the ODS and for the second digitisation of the SAOB seems to be a good choice for older dictionaries where all the text is produced as typed manuscripts to feed a print version. The situation for NO is not quite similar to these works. Firstly, the dictionary has been produced on a digital platform from the letter *i* onwards. When the work on the 1st edition finishes in 2014, approximately 2/3 of the dictionary entries will be digital entries feeding both the online dictionary and the printed version. Secondly, the punching part of the digitisation process is already done for the oldest volumes of NO. In order to make an online version which covers the whole dictionary, and in order to complete the dictionary database, the only fully satisfactory solution for our dictionary will therefore be to integrate the digitised text from the oldest volumes into the already existing entry structure of the digital dictionary.

The current state for volumes 1–5 of Norsk Ordbok is that the two first volumes were punched and proofread in 2001–02. The manuscripts for volumes 3–4 and the part of volume 5 that covers the letter *h* were produced in simple word processing programmes, and supplied with tags either during the editing process or afterwards. The original text for the oldest volumes of NO thus existed as digital manuscripts as early as in 2002. In 2005, the Norsk Ordbok 2014 project organisation made a pilot study on the integration of this digital text into the modern database system. The adaptation of the texts into the new and stringent database format proved too difficult and too time-consuming for the time-limited project organisation, and was therefore put on hold.

The entries from volumes 1 and 2 are integrated in the database system of NO 2014, but only in an incomplete version. The text is not in line with the current quality when it comes to consistency, and it does not give a complete coverage of older source material. Volumes 3 and 4 are partly integrated in

the database, but a lot of the text is not fitted into the correct fields, and the huge amount of dialect data and information on etymology is lacking altogether. The part of volume 5 covering entries starting with the letter *h* is not integrated in the database at all.

3.1 Why digitisation *and* revision?

Why is it so important to do the digitisation and the revision in one integrated operation? As mentioned above, the project organisation made a pilot study in 2005 to see if it would be possible to load the text of the oldest volumes into the modern editorial database. The pilot revealed that a lot of work has to be done to make the old text fit into the strict categories of the new editorial system, and that work inevitably also involves revision. One way of presenting the whole dictionary digitally without performing this integrated process of digitisation *and* revision would be to publish the oldest volumes as searchable PDFs on the Internet. This would be very unsatisfactory for several reasons: low user-friendliness, no possibility to perform searches across the base, lack of access to multi-word expressions in the earliest volumes, lack of possibility to do thematically based revisions and use the dictionary contents for other purposes etc.

Producing a digital dictionary which is identical to the printed version of NO is not the best solution in the view of the project organisation. Instead, we want to fit the entries from volumes 1–5 into the modern editorial database format. Preserving the contents of the oldest volumes in detail would force us to extend the existing database structures in order to adapt it to the structure and the idiosyncrasies of the old entries. Our goal is instead to modernise and standardise these entries and adapt them structurally to the modern online dictionary format.

3.2 Structural changes related to the digitisation

The first four and a half volumes of the dictionary were produced manually. The entries of these volumes are of a high quality for their time, but they often have a very tiered structure (Atkins & Rundell 2008:249) and from time to time include entry-specific structuring of data. This practice is possible and probably inevitable when the manuscripts are produced by hand, but it meets problems with the introduction of a digital production platform.

In 2002, the senior editing staff of the dictionary did a huge job extracting an ideal entry structure from the early volumes. This was used for setting up the electronic editing schema of the modern, digitised dictionary. The entry structure at the macro level (entry status, flat vs tiered structure, content selection etc., cf. Atkins (2008: 36ff)) was created on the basis of what was conceived to be the best practice of the old volumes, but this still leaves a lot of information that will not fit into the categories of the schema, and that will need to be given elsewhere in the entries or, if deemed superfluous, deleted. The planning of the entry structure at the macro level is much in line with the process of dictionary planning described by Atkins (2008), but for a dictionary project that has already published five

volumes the options when setting up the macro structure are not open in the way they are when planning new dictionary projects (see also Cantell & Sandström 2012: 166f).

Another task associated with the digitisation of the material is the electronic linking of words in definitions, etymologies and elsewhere, as well as the linking of the first part of compounds to the correct basic word. Such links are an integral part of the structure in the latter volumes, and must be added also in the early material. This linking also requires that the structure of the older volumes is possible to adapt into the new data base system.

3.3 Structural changes related to the revision

Several structural changes must be performed in the old material in order for it to meet the requirements of NO's present editorial practices; a few examples will be mentioned here. As stated above, multi-word expressions were in the early volumes treated more or less on a par with ordinary examples, while starting from the letter *i* they have been edited as sublemmas in boldface. In order to attain a unitary structure throughout the dictionary, multi-word expressions in the early volumes must be identified and changed into sublemmas. A case in point is the phrase *bita i graset* ≈ 'bite the dust' (literally 'bite in the grass'), seen in figure 1. The phrase appears as an example under sense 1a in the entry **bita**, but clearly deserves the status of sublemma in a revised version the entry.

> I **bita** (*í*) v **bit, beit, biti** [VAgd16-,L,L 280,A, R3; målf òg ft pl *bito* (Hall), *beto* (Va); fp òg *bite, bete*; gno *bita (bit, beit, bitum, bitit)*; grunntyd 'kløyva'; tyd 6 vel av eng. *beat* 'slå']. **1) a)** […]
>
> || *bita i graset*. 1) (eigl om hest som stoggar og tek seg ei grastugge når han møter folk; overf om folk) stogga og tala til ein som ein råkar (Rog): *han Ola for forbi meg utan å bita i graset*. 2) bita i bakken; tapa: *dei lyt bita i graset som hev minste makti* (RomsOrdt)

Figure 1: Part of the entry *bita* with the multi-word expression *bita i graset*.

The structure of senses will – particularly in longer entries – need to be made flatter, more transparent and thus easier to navigate. The fact that the editors have access to a much larger body of linguistic data today (including a ~100 mill. word corpus) than when the early volumes were produced has contributed to less tiered sense structures in the latter volumes, and this will necessarily also be the case for the early material after revision.

There are a lot of structural features where the early volumes differ from today's editorial practice, and where structural revisions along the lines of the present editorial guidelines are required. One example concerns the use of usage labels; certain labels are no longer in use, such as *lbr* (*lite brukande* ≈

'should be used with caution'), which is connected with a certain puristic inclination in the early years of the dictionary. Figure 2 shows two entries with this label from volume 1:

> **behandla** v **~a** [ty] lbr. **1**) stella med etter ein viss metodisk framgangsmåte for å nå eit visst resultat; handsama; ha føre: *behandla ein for feber* (Ra.F) / *behandla ei sak.* **2**) fara fram (slik el slik) mot, fara åt (slik el slik) med: *eg hugsa og korleis han behandla Husmanns-Knut* (HolmS 57). **behandling** f lbr, det å behandla; stell, handsaming; førehaving; medferd.

Figure 2: The entries *behandla* 'treat' and *behandling* 'treatment' are equipped with the label *lbr*, although they are widespread in modern Nynorsk.

Another example concerns the labels *zool* (zoology) and *bot.* (botany), which were earlier used for all definitions covering names of animals and plants respectively, but are today restricted to official terms, while e.g. local names for plants lack the label *bot.*, but are electronically linked to the official term.

3.4   Revision of the lemma list

Faced with the task of producing a definite number of volumes on the basis of a certain amount of data, the project NO 2014 has developed effective methods for determining which lemmas should be included and how much space each entry should occupy (Grønvik 2006). The existing lemma list in volumes 1–5 must be revised using the present criteria for inclusion in NO and taking into account the material we have at our disposal today, which is a lot larger than when the first volumes were edited and includes a corpus dominated by 21st century newspaper texts. Neologisms and words that were previously not represented or poorly represented in the material must be included, together with lemmas of German or Danish provenance that were left out for puristic reasons but are used in modern Nynorsk (cf. section 1). In other cases lemmas that were originally included must be excluded – especially in volumes 3–4, where the inclusion criteria were clearly more liberal than today. Thus one can fairly frequently find entries that are based on hapax legomena (figure 3) or exclusively on occurrences in bilingual dictionaries (figure 4). These entries do not qualify for inclusion in the dictionary according to the present editorial guidelines.

> **franske-flokk** m flokk av franskmenn (Maul.I I,64).
> **fred-stor** adj poet., sj, fredfull: **i fredstore ævelengdi* (Hovd.SS 109).

Figure 3: Entries in volume 3 based on hapax legomena.

fransk-etar m [ett ty *franzosenfresser*] person som ottast el hatar franskmennene (Ra.E u *gallophobe*; Ra.F u *gallophobe*; Vo.Ty u *franzosenfresser*).

fransking f gallisisme, franskbragd (S.D u *gallicisme*; Vo.Ty u *gallizismus*).

Figure 4: Entries in volume 3 based exclusively on occurrences in bilingual dictionaries.

The oldest entries of the dictionary are not more than some 70 years old. This means that the diachronic dimension of the work itself is less challenging than for dictionaries with a production period that stretches over more than one century. Still, the oldest parts of NO show that some entry revision is needed. New entries have to be added, some old entries should be removed altogether and a lot of existing entries need revision due to broader and sounder empirical evidence, language change or both.

3.5 Revision of the dialect data given in the entries

The dialect material at the editors' disposal is substantially larger today than 80 or even 30 years ago. The geographical indications regarding special dialect forms and dialectal uses of words and word senses can thus be supplemented, in many cases possibly justifying the use of larger areas instead of single counties (the county is the smallest unit used for geographical references in NO). At the same time, the geographical indications in some of the early volumes reflect a more liberal practice than the one followed today, and they must be checked to make sure that the dictionary reflects the actual dialect material at our disposal.

The method of presenting dialect forms has changed somewhat during the history of NO; in particular, volumes 3–4 present such forms in greater phonetic detail and with more parallel forms than both the earlier and the latter volumes (cf. figure 5). Here the revision must imply a certain degree of simplification, following methods established in 2002 and later.

fredag m [Fyresdal1698 *frædajen*, VTel1821 118 *frædajæn*,A,R; målf *fre′dag* (Vanl, sjeldnare *fre′da*), *fræ′dag* (mest sfjells, sjeldnare *fræ′da* Furnes), *fræ′dæ* (Vestf), *fræ′dæu* (ØvRendal), *fred′da(g)* el *fræd′da(g)* (Agd sumst,Sokndal, Stav), *frei(d)′dag* (JrR), *frad′da* o l (Trysil, Tynset), *fre`dag* (Elsfjord,Gimsøy,Borge i Lof, Bjarkøy); gno *frjádagr*, mno *fræigjadagr* svar. t gno **friggjardagr* 'Friggs dag' etter lat *dies Veneris* 'Venus' dag',T]

mån-dag m [Fyresdal1698 *måndajen*, Stav1698, VTel1821 118 og 121 *mådajæn*, C, A, R, R2; skr òg *mandag* (ØklandGA362,Furs.ENE44,Hovl.SV63,I.LindSD 131); målf òg *man-* (vanl), *mån-* (heller vanl), *mon(n)-* (Torsnes, Vågå R, Uvdal, ATel R2, SAukra, Rindal), *må-* (Vinje i Tel R, Mo i Tel R), *mårn-* (Alvdal, Grong); norr. *mánadagr* 'månedag', ett lat *dies lunae*, frå gr]

Figure 5: The introductory section of the entry *fredag* 'Friday' from volume 3 and that of *måndag* 'Monday' from volume 8. Note the differences in the notation of dialect forms (introduced by 'målf' and 'målf òg' respectively).

3.6   Revision of definitions

In the majority of the entries the actual wording of the definitions will hardly require a lot of revision. Still, since the publication of volume 1 in 1966 the language has certainly undergone quite a few changes on the level of semantics and pragmatics – changes that must necessarily lead to adjustments in a number of definitions. Obvious examples are words that were earlier used neutrally, but later developed derogatory connotations and often have become obsolete altogether. In figure 6, the definition of *australiar* 'Australian' is '(white) person belonging in or coming from Australia'; the first word in the definition should definitely be deleted. The second headword, *australneger*, is no longer in use due to its derogatory character. As NO is a descriptive dictionary which documents actual (historical) usage, the entry should be preserved, but the definition must be updated to reflect the stylistic properties of the words and the fact that it is obsolete. The modern neutral term *aboriginar* 'Aborigine, Native Australian', which is not found in volume 1, must also of course be added.

> **australiar** m [-*ra'*-; til *Australia*, avl. t lat
> *australis* 'sørleg' av *auster* 'sønnavind'] (kvit)
> person som eig heime i el er frå Australia.
> **australneger** m [-*ra'*-] ein av det opphavlege
> folket i Australia.

Figure 6: The entries *australiar* 'Australian' and *australneger* 'Australian negro' from volume 1.

On a more general level, there is a tendency in some of the early material towards focusing on the particular rather than the general and to posit separate word senses where the present practice would prefer lumping rather than splitting. Thus especially in longer entries there will be a need for revising or rewriting definitions. Entries that lack a definition altogether but meet the criteria for inclusion in the dictionary must of course be provided with a definition.

Integrating new source material and meeting the requirements of a modern scholarly dictionary

The new source material – including corpus data – must be integrated at all levels of the early volumes of the dictionary. This will be reflected in the addition of new entries (cf. 3.4), the creation of new senses in existing entries, the introduction of new examples, especially the addition of more recent examples (sometimes due to reasons of space and clarity replacing some of the existing examples) as well as in new, updated geographical indications (cf. 3.5).

It is essential to ensure that the early volumes meet the requirements of a modern scholarly dictionary. This implies, firstly, that every entry must be linked to its source material and, secondly, that all entries and all word senses must have a documented source material behind them and contain at least one source reference at the level of definition and/or example. In the electronic version of the dictionary the links between entries and source material will be made explicit, enabling users to verify the information given and potentially falsify it.

> **dyvels-** el **divels-klo** f [målf *divels-* (Mas-fjorden); ett ty *teufelsklaue*, ndl *duivelsklauw*, eng. *devils claw* eigl 'djevelsklo'] sjøm., fag., (mest i pl) eit par jarnkrokar i sams stropp til å huka i bjelkar o l som skal heisast; slike krokar i enden av eit seglskaut til å festa i storseglet på ein større båt med sprisegl.

Figure 7: The definition in the entry *dyvelsklo* 'devil's claw (a kind of split hook)'from volume 2 lacks source references. One or more references must be added, or the entry must be excluded from the dictionary.

4   Planning and implementing the revision project

As part of the planning it must be decided to what degree the dictionary entries should be rewritten. A plausible strategy is to assume that smaller entries – which constitute the majority – require only revision, while at least a part of the larger entries (especially large verbs and function words) will benefit from being re-edited. This re-editing must be performed with the editor at all times keeping a keen eye on the existing entry and making sure that all essential information that is given there and can be verified is transferred to the new version.

In the modern NO 2014 organisation, all the relevant source material is digitised and stored in a structured relational database system. This makes it possible to quantify relative space for each entry and to estimate the work load for the staff as a group and for each single editor. The experience from the last 12 years of project work shows that this way of working gives a high degree of prediction when it comes to how much time and money are needed to perform the whole operation of revising and digitising the oldest parts of the dictionary.

The whole of the source material behind the earliest volumes is included in the dictionary database system, and this provides a very sound way of estimating the work load for doing the integrated digitisation and revision work. For the whole bulk of 112,500 lemmas it is possible to make fairly accurate estimates that also take into account that some entries will be revised, while others will gain from a full rewriting. Based on experience with producing the last seven volumes of the dictionary over a period of 12 years, feeding both a printed publication (each volume includes 800 pages of entries) and an Internet version, the NO 2014 organisation estimates that the digitisation and revision of the first volumes will be possible with a staff of 16 editors working full time over a period of five years. This is approximately 45 % of the amount of work that was put into volumes 6–12.

5   Funding

The revision project will have a total cost of some 70 million NOK (approx. 8.5 million EUR). The production of NO has so far been funded by the University of Oslo and the Norwegian Ministry of Culture in a joint agreement, but this funding ends in 2014. Language infrastructure, including dictionaries, is cost-intensive and involves huge amounts of manual work. Norway is a relatively small language community, and the commercial potential of the basic language infrastructure resources for Norwegian is quite low. This means that in order to reach the central goals on the field of Norwegian language policy, the building up of basic language resources needs public funding.

A NO dictionary database covering the whole alphabet span will not only offer the public a comprehensive description of spoken Norwegian and written Nynorsk. The full dictionary database will also be an important component in future Norwegian language infrastructure and language technology. In this perspective, public funding of the digital integration of volumes 1–5 of Norsk Ordbok in the dictionary database would hopefully be within reach. In 2013 the Language Council of Norway set up a policy document for dictionaries and other basic lexical resources for the Norwegian languages, including Sami and the official minority languages of Norway. This policy document states the importance of a complete and updated online version of NO, and also states that this needs public funding (LCN 2013-08 and LCN 2014-03).

6   Conclusions

A lot of work is needed to bring the oldest volumes of NO up to the same digital standard as the rest of the dictionary. During the 80 years that have passed since the dictionary work started, the language itself, linguistic theory and preferred publishing platform have all changed. These changes have in turn led to changes in lexicographical practice. For a scholarly dictionary to be scientifically sound and relevant to the dictionary users, it is necessary to revise and upgrade its contents. For the dictionary database to become complete, it is not an option to choose only digitisation, or doing the process in two separate steps.

7   References

Aasen, I. (1850). *Ordbog over det norske Folkesprog*. Christiania: Carl C. Werner & Comp.
Aasen, I. (1873). *Norsk Ordbog med dansk Forklaring*. Christiania: P.T. Mallings Boghandel.
Atkins, B.T.S. (2008). Theoretical Lexicography and its Relation to Dictionary-Making. In: T. Fontenelle (ed.) *Practical Lexicography. A Reader*. Oxford: Oxford University Press, pp. 31-50.
Atkins, B.T.S., Rundell, M. (2008). *The Oxford Guide to Practical Lexicography*. Oxford: Oxford University Press.

Cantell, I. & Sandström, C. (2012), Hur blir en traditionell, tryckt ordbok en webbordbok? In: B. Eaker, L. Larsson, A. Mattisson (eds.) *Nordiska studier i lexikografi 11*. Rapport från Konferensen om lexikografi i Norden Lund 24–27 maj 2011, pp. 157-168.

Grønvik, O. (2006) Verknader av digitalisering på materialvurdering, redaksjonell metode og opplæring. In: *Nordiske Studier i Leksikografi*, 8, pp. 129-142.

Hellevik, A. (1966) Til fyrste bandet. In: NO volume 1, pp. XV-XVI.

LCN 2013-08 = En samlet ordbokpolitikk etter 2014 (letter from the Language Council of Norway to the Ministry of Culture). Accessed at: http://bit.ly/1i06QPN [08/04/2014].

LCN 2014-03 = Norsk ordbokpolitikk (memorandum from the Language Council of Norway to the Ministry of Culture). Accessed at: http://bit.ly/OBoRYU [08/04/2014].

NO = (1966-) *Norsk Ordbok. Ordbok over det norske folkemålet og det nynorske skriftmålet*. Oslo: Det Norske Samlaget. Web edition accessed at: http://no2014.uio.no [02/04/2014].

ODS = (1918-2005) *Ordbog over det danske Sprog*. København: Gyldendal. Web edition accessed at: http://ordnet.dk/ods [02/04/2014].

ODS FBTS = Fra bog til skærm. Accessed at: http://ordnet.dk/ods/fakta-om-ods/fra-bog-til-skerm [08/04/2014].

OED = *Oxford English Dictionary*. Accessed at: http://www.oed.com [02/04/2014].

SAOB = (1898-) *Ordbok över svenska språket*. Lund: Svenska Akademien.

Skard, S. (1932) Norsk Ordbok. Historie – plan – arbeidsskipnad. Oslo: Det Norske Samlaget.