

Lars S. Vikør, Section for Norwegian Lexicography,
University of Oslo

Norsk Ordbok – the Crown of Nynorsk Lexicography?

Abstract

Norsk Ordbok 'Norwegian Dictionary' is a multi-volume dictionary of the Norwegian standard variety *Nynorsk* and the Norwegian dialects. It is one of the very few dictionaries which cover both a written standard language and the oral dialects on which this standard is based. It was initiated around 1930, based on dialect material collected by volunteers and stored in a vast card archive, and on a variety of written sources. At present, three of twelve planned volumes have appeared, reaching into *g*. The paper gives a historical outline of the project, followed by a brief description of its structure and the types of information it gives. This is exemplified by the treatment of one particular word, *bunad*. Finally, some fundamental problems are briefly discussed: 1) the selection of lemmas, 2) the character of the sources, 3) the treatment of dialect forms, 4) the sequence of definitions.

The full title of *Norsk Ordbok* is *Norsk Ordbok. Ordbok over det norske folkemålet og det nynorske skriftmålet* 'Norwegian Dictionary. A dictionary of the Norwegian popular language [i.e. the Norwegian dialects], and the Nynorsk written language'. This title at once indicates the dual aspect of the dictionary: It gives integrated coverage of both oral dialects and a written standard language.

This dual aspect is the most special distinguishing feature of *Norsk Ordbok* as a lexicographic work. Normally, dictionaries cover written standard languages or some aspect of them (or, in the case of pronouncing dictionaries, oral standard language). If a non-standard variety is the object, the dictionary is labelled dialect dictionary – and these may cover a single dialect or a wider dialect area. *Norsk Ordbok*, however, in principle covers all Norwegian dialects together with written Nynorsk within an integrated framework. The reason for this must be sought in the origin and history of the language and the ideology on which it was built: it was codified on the basis of a set of related dialects in the nineteenth century, and has since then kept a closer contact with its dialectal base than more traditional written languages.

Ivar Aasen and the origin of Nynorsk

Nynorsk is one of the two written standards of Norwegian, based on the rural dialects in opposition to the Danish standard language, which was the official one even in Norway in the nineteenth century. Nynorsk was codified by Ivar Aasen in the middle of that century, and its lexicographic foundations were laid by Ivar Aasen's own *Ordbog over det norske Folkesprog* 'Dictionary of the Norwegian popular language' from 1850, later greatly expanded and reissued in 1873 under the title *Norsk Ordbog med dansk Forklaring* 'Norwegian dictionary with Danish explanations'. Aasen's dictionary was a very original enterprise, in fact a dialect dictionary that did not describe any single dialect, but the totality of Norwegian dialects in the framework of a single orthographical and morphological code meant to unite all the dialects that descended from Old Norse.

Aasen's selection of words was normative, and of course limited by the fact that he had collected most of the material singlehandedly. After his time, Nynorsk expanded as it was taken into use as a language of education, literature, press, administration etc., and a new dictionary reflecting this development was soon felt to be necessary. Besides, the dialects entered an age of transition, and it was felt to be imperative to safeguard and preserve as much as possible of their traditional vocabulary. So, these two seemingly opposing trends united to create a demand for a new dictionary, covering the traditional vocabulary of the dialects and the modernized vocabulary of Nynorsk.

The history of Norsk Ordbok

The project *Norsk Ordbok* was started in 1930. It was initiated by the Nynorsk publishing house Det Norske Samlaget with subsidies from the government. In 1972, the project was taken over by the University of Oslo, and at present it is part of the Section for Norwegian Lexicography at the Department of Scandinavian Studies and Comparative Literature of this university.

The initial plans implied a dictionary comprising two to three volumes. In the 1930s, a preliminary draft was worked out, largely by molding together the contents of the dictionaries already existing, including Aasen's, and translating the definitions into Nynorsk. At that time, all the existing Nynorsk dictionaries had explanations in Danish or *Bokmål*, the societally dominant standard of modern Norwegian which evolved from Norwegianized Danish. One of the desires concerning the

new Nynorsk dictionary was that it should be monolingual, or rather monolectal.

After WW2 new people came into the editorial staff, and they started afresh to draft a larger dictionary where all the available material was to be used. The dictionary was to be issued in fascicles, which were to be bound in volumes – a volume contains five fascicles. The first fascicle was published already in 1950 to commemorate the centenary of Aasen's first edition of his dictionary. After that, the work progressed steadily. The first volume was published in 1966, *a – doktrinær*, and the second volume, *dokument – flusken* 'uncombed [hair]' followed in 1978. After that, another project for a time absorbed most of the energies and resources of the staff, namely the one-volume *Nynorskordboka* 'The Nynorsk Dictionary', which was meant as a practical dictionary for ordinary users. Only when this was completed, in 1986, did the production of *Norsk Ordbok* continue. The third volume was issued at the turn of 1994/95, reaching to *gigla* 'stand loosely, waver'. The present plan is that the dictionary should consist of twelve volumes.

The material and time span of the dictionary

The material used by the dictionary is of two kinds: on the one hand, excerpts from the Nynorsk literature, press etc; on the other hand, information on usage in the dialects. In the 1930s, *Norsk Ordbok* established a network of contributors in the countryside, mostly primary school teachers, but also farmers, clergymen and others with an interest in the local dialect and culture. They wrote down information on words, expressions, and word forms in their dialects on small cards which they sent to the editors of the dictionary. This continued into the 1950s and 60s, when the number of contributors and hence contributions decreased from natural reasons. In total more than 1000 persons have been involved in these activities. In total a quarter of a million cards with dialect information were sent in, and with the cards containing citations from printed texts the whole archive today contains around three million cards. All Norwegian dialects are represented in the archive, but urban dialects are less well represented than rural ones, since Nynorsk is a rural-based language at the outset, and rural dialects and rural culture has a special place in the ideology in the Nynorsk movement.

The time covered by the dictionary is in principle the period from 1525 till the present age – 1525 because we in the historiography of our language customarily regard this year as the dividing line between "Middle Norwegian" and "Modern Norwegian". But from this time until

Ivar Aasen appeared in the middle of the nineteenth century, no Norwegian written language existed, so the only material we have, is some dialect texts and glossaries of dialects made by vicars – particularly from the eighteenth century – a very few from the seventeenth, and actually nothing from the sixteenth. Only after the appearance of Aasen's work, i.e. from the latter half of the nineteenth century onwards, do we have large amounts of texts, recorded linguistic information etc. A time limit downwards has not been set, so that new texts and new items of information from the dialects are continuously being included in the files and used in the dictionary. This means, of course, that the last parts of the dictionary will reflect a younger stage of the development of the language than the first volumes. But it has to be admitted that the classical Nynorsk writers, like Ivar Aasen, Aasmund O. Vinje, Arne Garborg, Olav Duun and some others, are more intensively drawn on for citations than more modern ones.

The microstructure of the dictionary

The articles are divided into three parts: head – “sharp brackets” – body. The head of every article consists of the lemma followed by a grammatical classification. The form of the lemmas presents some problems for two reasons: Nynorsk – like Bokmål – has undergone several spelling reforms during this century, and the publication of the first fascicles of *Norsk Ordbok* took place between two such reforms, implemented in 1938 and 1959 respectively. In *Norsk Ordbok*, the spelling of 1938 is maintained, but as headwords all optional forms after 1938 and 1959 are given (seldom more than two forms of a word). The other problem is the fact that dialectal forms often diverge considerably from standard Nynorsk, and that many dialectal words have not been formally codified in a standard form. In such cases, the editors must coin a standard form per analogy, and there will often be a conflict between the spoken form in the dialects and a more etymologically motivated form. In cases when the dialect forms are very deviant, they are listed as headwords with reference to the standard form.

Grammatical classification means giving word class, gender for nouns and conjugation for verbs. Irregular inflection forms (in the standard language) are given separately.

Then follows a section within sharp brackets, containing formal information. First, the age of the word is indicated by citing the older (pre-Aasen) glossaries where it is noted, if any, and a reference to Aasen's and Ross's classical Nynorsk dictionaries is given if the word is

treated there. Deviant forms in the older writing tradition and the dialects are given; especially the forms in the dialects are very explicitly and thoroughly listed with geographical references. Variations in gender and inflection in the dialects is also given. In base words this part of the “sharp brackets” is the longest one. The last part of the “sharp brackets” gives the etymology very summarily.

The body of the article consists of definitions and examples of usage drawn from the excerpted texts or from the collected dialect material. This is of course the main part of the article, and one might use the treatment of one single word as an example of how thoroughly this is often done.

An example: *bunad*

The word *bunad* is well-known to all who have studied Norwegian culture. The modern definitions of this word is: “national costume, traditional costume associated with rural culture”. But *Nynorskordboka* (1986) has two additional meanings: ‘equipment in the household’ and ‘tools’. Ivar Aasen’s dictionary from 1873, however, puts ‘household’ first, followed by ‘equipment’ and ‘tools’, and only as meaning no. 3 we get ‘costume’, while the current meaning ‘national costume’ does not appear at all.

In *Norsk Ordbok*, the word is given a lengthy treatment. Its meaning is divided into eight senses, which can be summarized like this:

1. Household, dwelling place (as understood in the pre-modern, agrarian society, with houses and farm lands)
2. The period of time a certain farmer runs a farm
3. Preparations to produce something, for instance baking (rare)
4. Things needed for running a household, furniture, everything a house contains (still seen from the viewpoint of the agrarian society)
5. Tools, utensils, necessities
6. Cloth, clothing – both for everyday life and for celebrations etc. – from which sense derives the modern meaning ‘national costume, traditionally based local folk costumes’.
7. A kind of women’s headgear, “a red band around the head, set with pearls or decorated with small silver plates” (rare)
8. Body-shape (rare, probably not derived from the senses enlisted above)

The first sense, ‘household, dwelling place’, must be the oldest one, actually a kind of verbal noun derived from the verb *bu* (Old Norse *búa*),

which meant 'to live somewhere, to have one's home somewhere' (as it does in modern Norwegian), but at the same time it meant 'to run a farm or a household' – since in the pre-Viking age when these words were coined, to have a home implied running it, cultivating land etc. The second sense has a very clear connection with the first, and is certainly a derivation of it. No. 3 is only documented by a single proverb: *bunaden er stundom større enn baksteren* 'the preparations are sometimes bigger than the baking'. This sense must be quite special and rare. Sense no. 4 must also be derived from no. 1, and no. 5 is a specification of no. 4; it includes for instance tools for fishing, driving a cart, sailing a boat etc. Sense no. 6 is 'cloth, clothing, particularly national costume', and with the term 'national costume' we in fact understand 'traditionally based local folk costumes' – since these costumes differ from locality to locality, but together they function as a potent symbol of "Norwegian-ness". This is the familiar modern sense of *bunad*, exemplified by the expression: *gå i bunad 17. mai* 'to wear *bunad* on the 17 of May', the national independence day of Norway, when this kind of clothing is very extensively used, especially by women. The last two senses are more special, but no. 7 is clearly derived from no. 6. No. 8 stands isolated from the rest, and should perhaps have been treated as a different word, a homonym.

The difference between *Norsk Ordbok* and the other, smaller dictionaries, is that in *Norsk Ordbok* one gets a total picture not only of what the word can mean, but also of the connections between the different senses, and the development of the meaning and use of the word. *Bunad* had passed through a long development where its meaning has been gradually expanded through unconscious use of the word in senses a little different from, but related to the established ones. All the meanings given in *Norsk Ordbok* were current in the dialects in the period between 1850 and 1950 and probably longer, especially in the speech of older people. But in modern times the sense 'national costume' or 'local folk costume' has become the dominant meaning of the word to all Norwegians.

Computerization

Norsk Ordbok was of course produced manually from the start. In the 1980s, the computer was taken into use, and in the 1990s we are computerizing our card files. All our three million cards have recently been scanned and are thus available as electronic facsimiles. These facsimiles are now being registered in a database. Besides, we are scanning many of the written sources, particularly the classical writers in Nynorsk (Aasen,

Vinje, Garborg), but also a representative selection of later Nynorsk poetry and prose (factional and fictional). All this takes place within the framework of the so-called "Documentation Project", a national project which aims at the computerization of all collections belonging the Faculty of Arts of the University of Oslo and gradually also the other Norwegian universities. The resulting database is expected to be ready before the end of the present decade.

Norsk Ordbok has not progressed so far as one should have wished, having reached *g* after half a century. But we hope that the database will prove a powerful tool which will increase the speed of production. We should, I think, be able to complete the work before the middle of the next century, but a more precise assessment is impossible at the moment.

A particular problem is the tendency to make too long and elaborate articles. We schedule the whole work to comprise twelve volumes, which is a far cry from the two or three ones planned in the 1930s. But the articles we make at present are so long that if we continue this way, we will need sixteen volumes before we are finished. Therefore, a thorough discussion and a revision of our editorial style will be needed.

Problems

The dual character of *Norsk Ordbok* is not without its problems, partly of a kind experienced by all large dictionaries, partly some rather peculiar ones. I shall briefly discuss some of them here. They relate to 1) the selection of lemmas, 2) the character of the sources, 3) the treatment of dialect forms, 4) the sequence of definitions.

1) As to the selection of lemmas, *Norsk Ordbok* is very inclusive. It is a basic principle of Nynorsk that it is lexically open to inherited words in all the dialects of the country. In principle, therefore, every word that is found in the archive is eligible for inclusion in the dictionary, regardless of geographical distribution or frequency.

The number of words represented in the archive is, however, too large to allow for a mechanical application of this principle. Some kind of selection procedures have to be applied. Words of the following types are examined critically before being in- or excluded: 1) Accidental compounds. 2) Foreign words which are restricted to a specific domain of language, particularly science. 3) New words that have not showed themselves to be stable elements in the language. A particularly problematic category are the many loanwords from Danish and Low German in the dialects which are normally proscribed in standard Nynorsk, because of the purist tradition which still exerts a (by now rather weakened)

influence. Such words are usually included in *Norsk Ordbok* if they are documented from the dialects or the excerpted written texts, albeit with some sort of marking ("little used", "less acceptable", or something similar).

2) There is an inherent self-contradiction in our dictionary: although covering the oral dialects very extensively, its sources are not oral at all. As I have shown above, the dialectal material has been taken partly from published sources (dialect word lists, written texts in dialects, folkloristic material and fiction where dialect words are used etc.), partly from written cards sent in by our collaborators and stored in our archives. Records of genuine speech are not used at all. Our material is still vast, but there are certain limitations: Much of the information concerns dialect usage of the older generations, partly probably extinct by now. This is in itself valuable, since in this way, linguistic and cultural information which would otherwise be lost, can be made accessible to later generations (e.g. of the kind exemplified in connection with *bunad* above). But it gives a one-sided view of the language. Modern usage are almost exclusively retrieved from modern written sources – literature and newspapers. It is also a problem that all the material – both from oral and written sources – has been selected *in order to* be included in the files. This means that the words and usage included tend to be the unusual and special. Day-to-day, trivial usage is underrepresented. In the future, this will be corrected as far as written language is concerned, when we have proceeded with computerization and made an electronic corpus of literary and non-literary texts. It will be a great challenge for us in the future to combine the use of our valuable file material with a new corpus-based approach.

3) The dialect forms of the headwords are given in the so-called "sharp brackets". They are taken from our card files and from a so-called "dialect synopsis", a set of hand-written books containing dialect forms of a large number of common words from every municipality of the country, made and kept at the Section for dialectology of the Department of Scandinavian Studies and Comparative Literature. The problem can only be sketched preliminarily here, but it concerns explicitness and detailedness. Norwegian dialects can be vastly different, they cover a large geographical area, and the structures and the relations between them are very complex. They are also in flux, like dialects everywhere. How far should one go in reporting subtle phonetic and geographical details, or, on the other hand: how far can one go in synthesizing and simplifying a complex reality, choosing the most representative forms of larger areas?

4) As to the sequence of the definitions, the problem is to which degree one should try follow a historical or chronological principle, and to which degree other principles should be used: e.g. logic or frequency in use in the modern language. This is a universal problem for lexicographers, of course, but it is particularly pressing for a dictionary which intends to give a historically based picture of oral language, which means that sources illustrating semantic development must be largely lacking. The sequence of the senses of *bunad* above may illustrate the methods we try to apply: it intends on logical grounds to give a picture of how the senses have developed historically in relation to each other, although no dating of the different uses and senses is possible before we reach modern times, where the development of the 'national costume'-sense of *bunad* can be quite adequately accounted for.

Conclusion

Whether *Norsk Ordbok* is going to be the crown of Nynorsk lexicography is impossible to say at present. It is a visionary project, and if successfully completed, it will become the largest dictionary of Norway and of Norwegian, and in Scandinavia only surpassed by the large historical dictionaries *Ordbog over det danske Sprog* (28 vol.) and *Ordbok över svenska språket* by the Swedish Academy (at present 31 vol.). In that case, it will be a cornerstone for the knowledge of Nynorsk and Norwegian dialects. But the task of producing such a work is a difficult one, and we have to take some very important decisions to speed up the progress. One of the problems awaiting us is of a biological nature: all the present eight editors of the dictionary will reach the age of seventy during the period 2005–2017. Inevitably, then, the future of *Norsk Ordbok* will be a topic of discussion, and it will be decided during the coming two decades whether (and how) it will be carried through to its conclusion. The question mark in the title of this paper must, therefore, remain for the time being.