

Language Technology in the Faroe Islands

by

Hjalmar P. Petersen

1 Introduction

At the moment there are not many electronic texts or text banks available in Faroese, and neither are there any morphological analyses in electronic formats. This presentation will therefore be short, but we do believe that we cover everything of relevance, which might be of use to Nordic and other linguistic researchers. We provide information about *where* the researcher or others with an interest in the subject can get access to the material, possible contact names, telephone and fax numbers and, if available, email addresses.

2 Dictionaries

From the early nineties, a certain amount of dictionaries have been published in the Faroe Islands. Some are available in electronic form, others only in print. The following dictionaries are available on CDs and/or floppy discs.

Donsk-føroysk orðabók (Danish-Faroese Dictionary) was published by *Føroya Fróðskaparfelag* in 1995, and this book is available in electronic form on four floppy discs. It was made in the Faroese editing program RiSt¹. This program has good search facilities. As the discs are installed on a PC, you have more search options. You can search on single words, parts of words/or morphemes and idioms, word sequences and information about word usage, i.e. whether the individual words are colloquial, local, belong to certain scientific terminologies etc. By pressing *Ctrl T* it is possible to search for Danish as well as Faroese words, inflection patterns such as neuter (n.), and parts of a word.

This program is developed for PC only, not for Mac.

Contact:

¹ Ritstjórnarforrit = Editing program.

Føroya Fróðskaparfelag
Handan Á 5
FO-100 Tórshavn
Tel. +322074
Fax+322074
email: maudh@setur.fo

Føroysk orðabók

The above description of the features of *Donsk-føroysk orðabók* (Danish-Faroese dictionary) also goes for *Føroysk orðabók* (Faroese Dictionary) from 1998, and *Føroysk Samheitaorðabók* (Faroese Thesaurus) from 2000. These two are available on CDs.

When installed on a PC, these three dictionaries are placed in a shared file, *Skjáttan* (The File), and the same search facilities are used for all three dictionaries. You can move freely between the dictionaries in *Skjáttan* by pressing *Ctrl S*.

Føroysk orðabók (Faroese Dictionary) is available from the same contact address as *Donsk-føroysk orðabók* (ie. Føroya Fróðskaparfelag etc.), whilst the thesaurus can be acquired from the publisher on the following address:

Stiðin
Hornavegi 16
FO-188 Tórshavn
Færøerne
Tel. +311589
e-mail: zakarias@post.olivant.fo

The publisher *Stiðin* also has an English-Faroese dictionary (*Ensk-føroysk orðabók*), which is based on a dictionary program from the large Danish publisher Gyldendal from 1992. The contact is *Stiðin*, Hornavegi 16 etc. , i.e. the same address as above.

There are other dictionaries available in Faroese. *Stiði* has published a Danish-Faroese dictionary, but this is not available in an electronic version.

Føroyamálsdeildin (Department of Faroese Language, Literature and Folklore) is at the moment working on a Faroese-Danish dictionary, a technical dictionary and an Italian-Faroese dictionary. All of these will probably be published in both printed and electronic versions. So will a Faroese-German dictionary edited by Ulf Timmermand, and a Russian-

Faroese dictionary edited by Johnny Thomsen. These people can be contacted on the following addresses:

Ulf Timmermand
Klokkaragøtu 1
FO-100 Tórshavn
Tel. 314776

Johnny Thomsen
undir Heygnum 3
FO-100 Tórshavn
Tel. 313509
e-mail: johnnythomsen@email.dk

3 Texts

Newspapers

There are two main Faroese newspapers, *Dimmalætting* and *Sosialurin*. These are published all days of the week except Sundays and Mondays.

Dimmalætting

Dimmalætting has an electronic archive dating back to 1998, and a majority of their material can be found here. This material is available to all people who subscribe to the paper.

The texts found in the archive of Dimmalætting can be searched as in normal text searches and cover the following fields:

- Editorial texts
- News from the Faroe Islands (Heimatíðindi)
- Foreign news (Útlendsk tíðindi)
- Greetings (Heilsanir)
- Readers' letters (Lesarabrøv)
- Classified advertisements (Gulu síðurnar)

Dimmalætting uses an electronic system developed by the company *Zenit*. The system works in the way that the texts are placed on the Intranet, and from here they are manually placed on the Internet and in the newspaper archive.

All texts are available on the Internet apart from news from the Faroe Islands and the classified advertisements.

You search through the texts as in a normal Word search, that is, you press *Ctrl B* and then you type in the word or sequence you wish to look for.

As mentioned, it is necessary to subscribe to *Dimmalætting* in order to gain access to these texts.

Contact:

Dimmalætting
Smyrilsvegi 13
FO-100 Tórshavn
Tel. +341200
Fax. +341201
Website: www.dimma.fo

Sosialurin

Sosialurin has electronic texts dating back to September 1997, and these texts are available if you subscribe to the paper.

In principle *Sosialurin* save all texts on the Internet under the title *Internet Sosialurin*. Texts can also be found in dvd format at the editing office of the paper. This format can be converted into a readable format without any major problems.

At *Internet Sosialurin* you can search as in normal text searches (*Ctrl B*), and further information can be obtained from:

Sosialurin
Argjavegur 26
Argir
Tel. +311820 / +314871
Fax: +314720/ +310817
email: post@sosialurin.fo
Website: www.sosialurin.fo

Printing houses

The printing houses have a variety of texts dating back to the late eighties, which can be searched as in normal text search. The texts are not generally accessible for the public, as the publishers, writers and the printing houses themselves own the copyrights for the material.

However, it is possible to get access to a selection of the texts for research purposes by contacting the printing houses mentioned below. You probably need to enter into a written agreement with the publishers, writers and printing houses, and they might want some form of financial compensation for the use of the texts.

None of these texts are available on the Internet.

Einars prent

Address:
Einars Offsett
v/ Eydnu Djurhuus
Sjúrðargøtu 19/Heygsbreyt 14
Postbox 173
FO-110 Tórshavn
Tel. +315193 / +311533
email: einarspr@post.olivant.fo

Einars Prent has got electronic texts dating back to 1984. These are difficult to access, but texts from about 1988 are more easily accessible. The search facility is similar to normal text searches. Please contact the above address for further information about the various texts and access to these.

Hestprent

Address:
Hestprent
Dvørgastígur 7
FO-100 Tórshavn
Tel. +314555
Fax +318026
email: hestoy@hestprent.fo

Hestprent has got electronic texts dating back to 1988. The texts are in various dissimilar formats and editions, but they can be converted to readable formats, even though this is a time consuming process, to make it possible to search for eg. words and sequences as in normal text searches.

The printing house has between 600 and 900 different examples of printed matters plus other material such as printed adverts etc. They mainly use Mac.

In principle there is no public access to the texts, as the publishers own the copyrights. It is, however, possible to gain access to the texts for research

purposes, possibly at a certain cost. For further information contact the above address.

Estra

Estra has various electronic texts saved on CDs in Mac format, and you can search these as in a normal Word program.

It is possible to gain access to the texts by contacting the owners of the copyrights for the texts, i.e. the publishers, writers and the printing house itself.

Address:

Estra Prent
FO 513 Syðrugøtu
Færøerne
Tel. +282205
Fax +441673
email: klestra@post.olivant.fo

The text bank

At Føroyamálsdeildin (Department of Faroese Language, Literature and Folklore) at Fróðskaparsetur Føroya (University of the Faroe Islands) there is a text bank based on a program called DT search. Here it is possible to search for words, sequences, parts of words (morphemes) and to see the words in a context.

Anyone who wishes to gain access to the text bank should contact the person mentioned below, however at the moment it is not possible for everybody to gain access to the bank, as some texts are restricted by certain clauses.

The selection of texts is broad. There is prose, non-fiction, news, websites, Internet chat and other texts. The text bank is continuously updated.

Contact:

Jógvan í Lon-Jacobsen
Fróðskaparsetur Føroya
Føroyamálsdeildin
V.U.Hammershaimbgsøtu 16
FO-100 Tórshavn
Tel. 315304
email: jogvanl@setur.fo

Texts from Lagtinget

The Faroese Lagting (local government) has various texts on the Internet, amongst these the code of statutes, on the following address: www.logting.fo.

It is possible to search these texts as in a normal Word program, and the website also has links to other sites.

4 Blindastovnur Føroya

Blindastovnur Føroya [The Faroese Institute for the Blind] is currently working to develop speech synthesis for Faroese. It is available in a beta version, which is being processed into a diphone version. This is based on recordings made by the Icelandic researcher P. Helgason at Department of Phonetics at the University of Stockholm, and S. Gullbein, who mainly works with the Faroese material.

A continuous morpheme analysis has been created; this is especially to establish morpheme borders, as one would in Faroese expect a short vowel in front of two consonants, but this is not always the case. For example it is *beuksteavur/* (bókstavur 'letter') rather than */bøksteavur/* (bókstavur 'letter').

This morpheme analysis is not available in any printed form or in any manual, as it has been continuously used in the practical work of developing the synthesis or rather the set of rules behind the synthesis.

The developers of the synthesis have also worked with the distribution of stress in loanwords, and they have so far formulated ten rules, which are not completely comprehensive, however. Again there is no printed material available, as these results have been implemented directly in the rules behind the synthesis.

In the development of the speech synthesis, the researchers have acquired an extensive text bank (app. 10 mill. words) with texts from the newspapers *Dimmalætting* and *Sosialurin* and texts from Faroese Internet sites and chat – in general they have used the Internet and the newspapers to build up a frequency lexicon.

They also have transcriptions of spoken Faroese, where deviations from normal Faroese pronunciation have been labelled.

This material is not available to the public, but for more information you can contact:

Blindastovnur Føroya
att. Sjúrdur Gullbein

Íslandsvegi 10A
FO-100 Tórshavn
Tel. +318222
email: sgull@post.olivant.fo

5 Føroyska Málstovan

The Faroese Language Council (Føroyska Málnevndin) publishes word lists, a leaflet about language usage (Orðafar), orthographic rules, the legislation for names and different links at the website www.fmn.fo, and the secretary of the Language Council can be contacted for further information on the address below. The links included on the website are mainly links to other Nordic language councils and information for the Faroese about where to find Nordic dictionaries on the Internet.

Address:

Føroyska málnevndin
V.U. Hammershaimbøgøtu 16
FO-100
Tórshavn
Tel. +312397
Fax. +318448
email: mariuss@setur.fo

6 Morphological analysers

As mentioned at the beginning, there are no morphological analysers available for Faroese. However, the researcher Tim Wentslau is currently working on one. This is mainly to be used for children with reading disabilities. The program works in the way that you type in for example (in Danish, as he is using a Danish model) *spi-* in Word. All uses of *spi-* are shown and by choosing eg. the sequence *spis-* (*eat-*), all connections with *spis-* are shown. The program has not yet been developed for Faroese, but a Faroese version is currently being developed.

Contact:

Tim Wentslau
FO-826 Trongisvágur
Føroyar
Tel. 371555
Website: www.auxilior.com