

Foclóir na Nua-Ghaeilge

The bibliographical databases for the period 1600-2000 are published at the address www.fng.ie. They are presented at this point in two different segments by way of a drop-down menu since the corpus 1600-1882 is considered a closed one, whereas the 1882-2000 database lacks material from periodicals.

Bibliographical database 1600-1882

The database presents all the titles brought together by way of a concordance on the CD-ROM. It is an online version of the material recorded under the button **Text List** on the main screen.

All validated printed texts for the period are included. The material ranges across annals, other historical documents, biographies, poetry, prose, Finn tales, romances, religious, legal, meteorological and medical texts, lives of saints, and translations of medieval European literature. It is an amalgam of texts published in book form and of shorter ones published in academic journals or in installments in monthly, seasonal or annual periodicals.

Further information on the database and on the corpus itself can be found on the CD-ROM, **Corpas na Gaeilge, 1600-1882, the Irish Language Corpus**.

Bibliographical database 1882-2000

The database contains both “opportunistic” material, i.e. electronic texts donated to the project by publishers and individuals, and “principled” selections, i.e. chosen titles based on a diversity of text-types.

It brings together titles which may be appropriate in creating a 20th century corpus.

Since it is intended primarily for lexicographical purposes, it contains, in addition to fiction and non-fiction, as one might expect, school textbooks, documentation from the European Union, annual company reports, etc., - texts which are unlikely to be of interest to Irish academics but which are essential for lexicographers.

There are now 1904 texts listed in the database. It began with a list of texts which were captured between 1982 and 1994 as exemplars, illustrative of correct usage in the various dialects primarily, and of the best creative writers in those dialects. In addition, there were a number of works by non-Gaeltacht writers and an occasional text on technical matters, making a total of 160 texts.

However, with the change of focus on the project from drafting the dictionary itself to the creation of corpora – an approach which had arisen in lexicography throughout the world – it soon became apparent that it would be necessary to add to that number of texts in order to produce a corpus that would be representative of the language of the 20th century. An additional list was compiled c. 2000 proposing an additional 340 texts, making up a corpus of 500 texts which would contain approximately 20 million words. But in light of the continuous revision by experts of the number of recorded words necessary in order to be representative, yet a further review was required.

For example, the **British National Corpus** contains 100 million words drawn from 4000 texts; the **Bank of English** contains 323 million words and now, the **Collins Dictionary, 175 Years of Publishing**, 9th Edition (2007) draws on a corpus of 2.5 billion words. 20 million words would not therefore be adequate for Irish.

Representativeness

Particular attention was paid to making the database representative in various ways. Fiction and non-fiction were covered; texts were selected from throughout the country where available; lesser-known authors were included as well as the well known.

In the category, **Fiction**, there are novels, short stories, poetry, plays, folklore; there are original works and translations; under **Non-fiction** are found (auto)biographies, historical narrative, research works on the dialects, books on technical subjects from algebra to zoology, annual company reports, etc.

However, it must be remembered that the amount of material available is virtually limitless. It was necessary to include titles from which excerpts were included in the dictionary draft, **A-** to **Al-**, (unpublished), - which perhaps explains the significant number of Latin textbooks included. The compilation of a further database is already underway, which will contain all types of periodicals, including learned journals, magazines of general interest, newspapers and news bulletins. It is there one will find, for example, **Irisleabhar na Gaedhilge**, which played such a dominant role at the beginning of the Revival. The list of works by various scholars and authors, such as Dáithí Ó hUaithne, Brian Ó Cuív and Breandán Ó hEithir will be greatly amplified there. Although certain works of criticism have been included in this database, it is in the periodicals database that one will find reviews. However, one must not overlook the history of the language – the fact that there is a range of subjects barely represented in Irish, if represented at all. It also highlights the disparity of cover in the dialects. Compare the volume of titles under **M** and **LM** vis à vis **U** and **LU**. Where

would Ulster be without the works of “Máire”? It also illustrates the paucity of women authors.

Processing

The databases are presented in order of titles but the user may process them to suit by clicking on the name of the selected column. Search and Advanced Search facilities are provided.

Spelling and diverse versions of names

The 1882-2000 database encompasses almost one hundred and twenty years of Irish publishing, two-thirds of which period had elapsed before the introduction of **An Caighdeán Oifigiúil agus An Litriú Caighdeánach**. The absence of such a standard can be noted, especially in the spelling of authors’ names and of publishing companies. Spelling details have been taken directly from the text throughout, whether accurate, inaccurate or differing from other works by the same author. For instance, Peadar Ó Laoghaire is also listed under Ua Laoghaire and O’Leary; Niall Ó Dónaill under Ó Domhnaill; Seán Mac Meanman under both Mac Meanman and Mhac Meanman. Sometimes the name is in English or the author is identified by a pen name. Dubhghlas de hÍde can also be found under Hyde and An Craoibhín Aoibhinn. It should be noted that surnames beginning with Ó are all listed before those beginning with O’. Certain women authors should be sought both under their maiden name and their married name. E.g. Ní Fhlathartaigh and Uí Ógáin. Therefore, account must be taken of these complications in order to obtain a full listing of an author. This is equally true of the 1600-1882 database.

Notation system

C / M / U : traditional language of the Gaeltacht from 1882 onwards ...[and]
reliable research on Gaeltacht Irish.

LC / LM / LU : original literary work by a native of the Gaeltacht.

AC / AM / AU : a translation other than of modern technical subject matter
by a native of the Gaeltacht.

NL : original work by a non-native Irish speaker and modern technical subject
matter by a native of the Gaeltacht.

L : pre-1882 writings

T : Terminology

< **Foclóir na Nua-Ghaeilge – Forbairt agus Fadhbanna**, Tomás de Bhaldraithe.

Celtic Languages and Celtic Peoples. Proceedings of the 2nd North American
Congress of Celtic Studies.iii Halifax. August 16-19, 1989, 203.

[Note : **C** = Connachta; **M** = Mumhain; **U** = Ulaidh; **L** = Litríocht; **NL** = Nua-
Litríocht; **A** = Aistriúchán; **_nua**= a new edition published after the author's death]

C / M / U also applies to oral poetry recorded from a native of the Gaeltacht,
post-1882.

Beathaisnéis I – IX had a central role in notation selection. Information taken
therefrom determined whether an author was classified under **C / M / U**, etc.
rather than **NL**. The decision was occasionally a problematic one.

There are a number of entries under **L**-. These are texts which should have been in the
1600-1882 Corpus, but which were overlooked for one reason or another. [e.g. L105,
L106, L107, L374]; in other cases, the entry refers to the editor's preface to a text
in the 1600-1882 corpus.

NLP is a notation coinage referring to Nua-litríocht, Páistí. Only one entry is
recorded.

Source references

In addition to **Beathaisnéis I-IX** already mentioned, the following sources were used to select titles :

An tÚrscéal Gaeilge, Alan Titley, An Clóchomhar, 1991.

An Gearrscéal sa Ghaeilge, 1898-1940, Aisling Ní Dhonnchadha, An Clóchomhar, 1981.

Stair Dhrámaíocht na Gaeilge, 1900-1970, Pádraig Ó Siadhail, Cló Iar-Chonnachta, 1993.

An Litríocht Réigiúnach, Máirín Nic Eoin, An Clóchomhar, 1982.

Litríocht agus Pobal, Gearóid Denvir, Cló Iar-Chonnachta, 1997.

Clár Litridheacht na Nua-Ghaedhilge, 1850-1936, Risteárd de Hae, Leabharlann na hÉireann, 1938-1940.

1. Na Leabhra
2. Filidheacht i dtréimhseacháin
3. Prós i dtréimhseacháin

Bibliography of Irish Linguistics and Literature 1942-71, Rolf Baumgarten, Dublin Institute for Advanced Studies, 1986.

Publishers' catalogues, especially those of Coiscéim, An Gúm, Cló Iar-Chonnachta, An Clóchomhar and Sáirséal agus Dill.

Corpais na Gaeilge

Since historical dictionaries are, by nature, long-term, FNG is committed to providing interim products to the public. The CD-ROM, **Corpas na Gaeilge, 1600-1882, the**

Irish

Language Corpus, published in 2004, was the first of a series in hand.

Forthcoming corpora:

Máirtín Ó Cadhain, ed. Úna Uí Bheirn

Peadar Ó Laoghaire, ed. Eilís Ní Mhearraí

Irisleabhar na Gaedhilge, ed. Déirdre D'Auria

Seán Ó Ruadháin, ed. Eithne Ní Ghallchobhair

Feedback

This database is work in progress and there are obvious omissions as it stands.

Opinions and suggestions from the public will be welcome. We would greatly appreciate an electronic copy of a suggested text, should such exist.

Suggestions and / or corrections should be sent on the appropriate form, found at

www.fng.ie, to

fng @ ria.ie

General comment should be sent by email to

e.nighallchobhair@ria.ie

Further information :

Foclóir na Nua-Ghaeilge

Acadamh Ríoga na hÉireann

19 Sráid Dásain

Baile Átha Cliath 2

Phone: +353-1-6090639 / 6090624 fng@ria.ie www.fng.ie