

In: Merja Kytö, Matti Rissanen and Susan Wright (eds), *Corpora across the centuries*. Amsterdam: Rodopi, pp. 23-31.

# A Corpus of Irish English

*Raymond Hickey*  
*Department of English*  
*University of Munich*

## 0. Introduction

Corpora vary greatly both in design and size. For any prospective compiler the main question is whether his intentions are liable to satisfy a need taken to exist among fellow linguists for a particular corpus. In the case of a major project covering something as global as present-day English one can safely assume that there will be demand for what one produces. Furthermore if one has large-scale financial and staff resources then the question of size can be treated quite generously. Considerably more attention must be devoted in advance to these questions when the projected corpus is to cover a rather specialised need. In the case of the corpus being compiled by the present author a number of decisions have had to be made in advance to render the project feasible and hopefully useful to linguists working in the field of corpus processing.

The first major decision to be made when compiling a corpus of a variety of English is whether one includes historical or just present-day material. Another decision, independent of the first, concerns the representativeness of the texts chosen. In the present instance this question posed itself at the very outset. Assuming that an historical element is to be included, should a corpus of Irish English offer a cross section of literature written in English by authors who were Irish by birth or affiliation? If so, are the usual criteria of literary merit to be the guideline for the quantity of material by a certain author to be entered. To quote a case in point, if this were so it would be necessary to give considerable weight to the works of W.B. Yeats and G.B. Shaw. But the matter can be seen in quite a different light. Going on the assumption that users of a corpus of a variety of English will be interested in determining what linguistic features are characteristic for just that variety, one could choose texts which are linguistically representative. There may be cases where literary merit and the interest of the variety linguist meet, as in the writings of J.M. Synge, but this is more the exception than the rule. Authors like Shaw and Yeats are not particularly interesting linguistically as both use very standard forms of English.

The decision of the present author has been made for a corpus which illustrates the language traits of Irish English at their most salient. This has meant that some authors are included who are not normally regarded as particularly meritorious from a literary point of view, e.g. Lady Augusta Gregory. Others are represented by a work or works which are not necessarily regarded as their masterpieces but which show

their portrayal of Irish English at its best. This is evident in the case of Shaw whose *John Bull's other island* is the only work incorporated and in that of Yeats who is represented by the play *Cathleen ni Hoolihan*.

As the projected corpus is to include historical material, it has seemed sensible to start from the earliest attestations of Irish English and divide the texts according to period and genre. The latter aspect is important when the goal of the corpus is borne in mind. Of the main literary genres, drama is the one which clearly is most typical of any given variety of a language as it consists of representations of the spoken word. However, in the older period of Irish English this genre is not to be found so that here only poetry is included which is perhaps dubious in its varietal authenticity but which for want of other material must be accepted.

## 1. Periodisation of corpus

**Medieval Period.** The older period of Irish English is the medieval one which begins historically with the arrival of mercenaries from Wales in 1169 and which lasts until the final defeat of the united Irish forces at the battle of Kinsale in 1601. In this period the available linguistic material is scanty. Indeed the bulk of it is contained in a set of poems to be found in the Harley 913 manuscript of the British Museum and which are available in an annotated edition by Wilhelm Heuser in 1904 who took some liberties in punctuation and expanding abbreviations (Kosok, 1990:22). Going on some onomastic evidence these poems are regarded as Irish in provenance and are referred to as the *Kildare Poems* after the mention of a monk Michael of Kildare as the author of one of the poems; they probably stem from the beginning of the 14th. century. To these should be added the poem 'The virtue of herbs' (Zettersten, 1967). Prose documents from the early period are scarce indeed. From the 14th. century there are the *Acts and Statutes of the City of Waterford*; from the 16th. century there is the motley *Book of Howth* (Kosok, 1990:28), neither of which are linguistically particularly interesting. For the present corpus only the poetry just mentioned has been incorporated.

The linguistic continuation of the medieval period is not to be found in the literature of the 17th. and 18th. century but in the attestations of an archaic dialect from the south east of the country (Co. Wexford). It is called after the two baronies where it was spoken, Forth and Bargy. This variety of English is more like a form of Middle English (Hickey, 1987) which has been strongly influenced by Irish, at least lexically. It survives only in the form of glossaries (see Vallencey, 1788 and Poole, 1867) which were compiled at the end of the 18th. century before the dialect died out. Both these glossaries are to be found in the corpus, along with one or two texts in the dialect (poems and songs). In the corpus the glossaries are available in database form which facilitates lexical examination. They may be processed as databases and/or converted into texts quite easily using the software by the present author (Hickey, 1992).

**Early Modern Period.** When considering the early modern period in Ireland a strict distinction in types of English, (i) and (ii) below, must be made which remains of importance up to the present day.

- (i) More or less genuine representations of Irish English by natively Irish writers.
- (ii) Stretches of texts by non-Irish writers where the non-native perception of Irish English is portrayed.

There is a remarkable amount of material, mostly in the form of drama, stretching in time from the very end of the 16th. up to the mid 18th. century (Bliss, 1979). This literature contains material of both of the above types and will form the bulk of this section of the corpus. Notable among the dramatists whose works are to be found here are William Congreve (1670-1729), R.B.Sheridan (1751-1816), George Farquhar (1678- 1707) among others (see list below). There are a few samples from this period which are not from plays, such as the poem *The Irish Hudibras* by James Farewell or Swift's *A Dialogue in the Hybernian Stile*.

Type (ii) literature should not be underestimated in its value for the variety linguist. It is interesting in that it reveals what features of Irish English were salient and thus registered by non-native speakers. These features have gone into forming the linguistic notion of the *Stage Irishman*, a stock figure in much drama from this period onwards (Duggan, 1969[1937]).

**The 19th. and 20th. centuries.** For this period only one genre is presently represented, i.e. drama, for the reason outlined above that it is likely to contain most examples of specifically Irish English features as it is written speech.

The two main authors here are John Millington Synge and Sean O'Casey. These literary figures are in a way complementary. Synge is to rural Ireland what O'Casey is to urban Ireland, above all Dublin. O'Casey was himself a native of Dublin, while Synge, although not a native of the west of Ireland, studied the life and language of its inhabitants and attempted to represent this faithfully, at least in his early plays. It is true of both authors that their later plays are stylistically more idiosyncratic and less typical of a general form of the rural or urban varieties of Irish English.

Other dramatists some of whose typical works are included in this section are Dion Boucicault (1820-1890), Lady Augusta Gregory (1852-1932) along with Shaw and Yeats (see remarks above).

A conceivable expansion of the current corpus would be to include some of the many prose works by Irish writers. Some of these evince a certain linguistic interest. For instance the author Maria Edgeworth (1767-1849) has stretches of typically Irish English speech in her novel *Castle Rackrent* as do the authors Edith Somerville (1858-1949) and Violet Martin (Ross) (1862-1915) in their *Experiences of an Irish RM* as well as the northern writer William Carleton (1794-1869) in his *Traits and Stories of the Irish Peasantry*. This direction could be continued into the 20th. century to include the major prose works of writers like James Joyce or Flann O'Brien. However, whether this will be realised depends very much on financing and staffing as the volume of such works is very much greater than that of the dramas, poetry and glossaries already envisaged so that entry and correction would be considerably more time-consuming.

## 2. Processing a Corpus of Irish English

To gain maximum benefit from examining a corpus one must use appropriate software. For the corpus at hand it is recommended to avail of the programme suite *Lexa* (Hickey, 1992) also by the present author. The lexical analysis of texts and a variety of information retrieval tasks can be carried out easily; furthermore the processing of lexical material is facilitated by the use of the supplied database manager in the package. Those interested in syntactically deviant features of Irish English can, for instance, examine the dramas with the appropriate software for locating syntactic contexts. Statistics can be gained from any corpus texts and integrated into statements on the linguistic nature of such texts.

## 3. Preliminary list of works in A Corpus of Irish English

### Medieval Period

- 1) Kildare Poems
- 2) The Virtue of Herbs
- 3) Vallencey's glossary of Forth and Bargy
- 4) Poole's glossary of Forth and Bargy

### Early Modern Period

William Congreve (1670-1729)

- 1) The way of the world

Thomas Sheridan (1719-1788) (father of following)

- 1) Captain O'Blunder or The Brave Irishman

Richard Brinsley Sheridan (1751-1816)

- 1) St.Patrick's Day
- 2) The School for Scandal

George Farquhar (1678-1707)

- 1) The Beaux' Stratagem
- 2) The Twin Rivals

Oliver Goldsmith (1728-1774)

- 1) She stoops to conquer

This section also contains a number of minor pieces and extracts which are useful in gaining a picture of what Irish English was like at the beginning of the modern period.

### The 19th and 20th centuries

Dion Boucicault (1820-1890)

- 1) The Colleen Bawn

John Millington Synge (1871-1909)

- 1) Riders to the Sea

- 2) The Tinker's Wedding
- 3) The Well of the Saints
- 4) In the Shadow of the Glen
- 5) The Playboy of the Western World
- 6) Deirdre of the Sorrows

Sean O'Casey (1884-1964)

- 1) The Shadow of a Gunman
- 2) Juno and the Paycock
- 3) The Plough and the Stars

Lady Augusta Gregory (1852-1932)

- 1) Spreading the News
- 2) On the Racecourse
- 3) Hanrahan's Oath
- 4) The Workhouse Ward

George Bernard Shaw (1856-1950)

- 1) John Bull's Other Island

William Butler Yeats (1865-1939)

- 1) Cathleen ni Hoolihan

**Lexical databases of Irish English** (see also 'Medieval Period' above)

- 1) J.Clark (1917)
- 2) P.W.Joyce (1910)

## References

- Bliss, Alan. 1979. *Spoken English in Ireland 1600-1740*. Dublin: Dolmen Press.
- Clark, J. 1917. *The vocabulary of Anglo-Irish*. St.Gallen: Zollikofen.
- Duggan, G.C. 1969. *The stage Irishman (a history of the Irish play and stage characters from the earliest time)*. First edition 1937. New York: Benjamin Blom.
- Heuser, Wilhelm. 1904. *Die Kildare-Gedichte. Die ältesten mittelenglischen Denkmäler in Anglo-Irischer Überlieferung*. Bonner Beiträge zur Anglistik Bd. 14. Bonn: Hanstein.
- Hickey, Raymond. 1987. 'A lost Middle English dialect: the case of Forth and Bargy', in: J.Fisiak (ed.), *Historical dialectology*. 235-272. Berlin: Mouton-de Gruyter.
- Hickey, Raymond. 1992. *Lexa. Corpus processing software, 3 Vols.* Bergen: Norwegian Computing Centre for the Humanities).
- Hickey, Raymond. 1993a. 'The beginnings of Irish English', *Folia Linguistica Historica*.
- Hickey, Raymond. 1993b. 'Corpus data processing with Lexa', *ICAME Journal*.
- Hogan, James Jeremiah. 1927 *The English language in Ireland*. Dublin: Educational Company of Ireland.
- Kosok, Heinz. 1990. *Geschichte der anglo-irischen Literatur*. Berlin: Erich Schmidt.

- Joyce, P.W. 1910. *English as we speak it in Ireland*. (Reprint 1979 + 1988, Dublin: Wolfhound Press). London: Longmans, Green & Co.
- Poole, Jacob. 1867. *A glossary, with some pieces of verse, of the old dialect of the English colony in the baronies of Forth and Bargy, County of Wexford, Ireland*. London: J.R.Smith.
- Vallancey, Charles. 1787-1788. 'Memoir of the language, manners, and customs of an Anglo-Saxon colony settled in the baronies of Forth and Bargie, in the County of Wesford, Ireland, in 1167, 1', *Royal Irish Academy. Transactions* 2: 19-44.

CELT, the Corpus of Electronic Texts, is Ireland's longest running Humanities Computing project. It brings the wealth of Irish literary and historical culture to you on the Internet, for the use and benefit of everyone worldwide. It has a searchable online textbase consisting of over 19 million words, in 1638 contemporary and historical documents from many areas, including literature, medicine, and the other arts. This experimental search interface uses Google's custom search. It will search resources on both CELT and our former sister project, Multitext. CELT in a nutshell. A Corpus of Irish English was published with Corpus Presenter by John Benjamins, Amsterdam in 2003. The descriptions below assume that you are viewing the corpus with Corpus Presenter as the two come together on the CD accompanying the book (see descriptions elsewhere on this website). Getting started. A Corpus of Irish English consists of over 70 texts with a time span of nearly 600 years. The material is arranged in a manner which reflects the main division in the history of Irish English into an earlier period, from the late 12th century to the end of the 16th century, and a later period, f 100+ million word corpus of British English, 1980s-1993. Freely-available online. Allows for an extremely wide range of searches.Â The British National Corpus (BNC) was originally created by Oxford University press in the 1980s - early 1990s, and it contains 100 million words of text texts from a wide range of genres (e.g. spoken, fiction, magazines, newspapers, and academic). The BNC is related to many other corpora of English that we have created. These corpora were formerly known as the "BYU Corpora" , and they offer unparalleled insight into variation in English . Click on any of the links in the search form to the left for context-sensitive help, and to see the range of queries that the corpus offers. Both Irish and English are the official languages of Ireland. Two languages for two civilisations, a challenging part of Irelandâ€™s history. A conversation with a native Irish English speaker will do wonders for your comprehension and your accent. (Source: Engin\_Akyurt). Currently, the Irish government attempts to make the two languages equal by giving the Irish language the status of the countryâ€™s first official language. However, Ireland is a predominantly English-speaking country in terms of its popular culture. Furthermore, the Celtic languages (which also include Scottish Gaelic and Manx)