Tracking dialect history: *A Corpus of Irish English*

*Raymond Hickey*

*Essen University*

**1 Introduction**

The present chapter attempts to describe the text corpus called *A Corpus of Irish English* (Hickey 2003a) which contains some 90 texts which attest the history of Irish English from its beginnings (in written form) in the early 14th century to the early 20th century. The considerations in this chapter are diachronic in nature and hence complement the synchronic examination of Irish English to be found in the chapter by Jeffrey Kallen and John Kirk in volume one of the current work. Because of its historic nature, the corpus consists solely of texts, but again this kind of corpus could provide background information for corpora of present-day varieties of English such as that described by Anderson and Kay (volume one). The corpus also links up with atlas-type projects such as that reported on for Dutch by Barbiers, Cornips and Kunst (volume one) or that already published as Hickey (2004a).

The purpose of *A Corpus of Irish English* is to put at the disposal of interested students and scholars a set of texts, the linguistic examination of which can help to answer questions concerning the genesis of Irish English and to throw light on issues of dialect development, especially with reference to overseas forms of English in what has come to be known as “new dialect formation” (Hickey 2003b, Gordon and Trudgill 2004).

The electronic files of *A Corpus of Irish English* consist of various types, depending on the source of the data and the literary genre they represent. The geographical source of some of the text types can be given, for instance the glossaries for the dialect of Forth and Bargy all stem from the south-east corner of Ireland. Literary works by identified authors can also be located, at least going on their place of birth. In most cases this also provides an indication of the dialect of Irish English which is being represented as in the case of Sean O’Casey (1884-1964). However, in other cases, such as that of the playwright John Millington Synge (1871-1909), – see section 3.3 below – the question of dialect portrayed in plays is difficult to answer clearly as Synge maintained that he was representing the speech of uneducated peasants in the west of Ireland. The question of what dialect is represented is naturally of concern where works are anonymous, as with the many text fragments from the beginning of the early modern period, i.e. the late 16th and early 17th centuries. Another feature of many texts throughout the history of Irish English is that they are satirical in nature, that is, they use what is regarded as Irish English to parody the speech of the Irish. This matter is especially critical as many of the authors of such satirical works were in fact not Irish.
It is fair to assume in such cases, for instance with Ben Jonson’s *Irish Masque at Court* (1605) (see edition in Jonson 1969), that the speech used is a caricature based on a small selection of features which the non-Irish noted in the speech of the Irish. While such texts are obviously not necessarily reliable descriptions of varieties of Irish English in the early modern period, they nonetheless reveal what features were salient to non-Irish observers.

### 2 Background to *A Corpus of Irish English*

The present corpus was compiled over a period of several years during which texts were collected, digitalised (largely through scanning) and prepared for easy manipulation by corpus analysis software (Hickey 2003a). At the initial stage various decisions had to be made which determined the later form and content of the corpus. Among these early considerations was one about whether to include historical or just present-day material. The decision came down in favour of historical material, largely because this material had not been available in electronic form up to then. Another decision, independent of the first, concerned the representativeness of the texts chosen. Assuming that an historical element was to be included, it was decided to offer a cross section of literature written in English by authors who were Irish by birth or affiliation. A further consideration was whether literary merit was to be a guideline for the quantity of material by a certain author to be entered. If this were the case then it would have been necessary, for instance, to give considerable weight to the works of W. B. Yeats and G. B. Shaw. But this matter was seen in a different light. Going on the assumption that users of a corpus of Irish English would be interested in determining what linguistic features are characteristic of just that variety, those texts which are linguistically representative of Irish English were given preference. There were cases where literary merit and the interest of the variety linguist met, as in the writings of Sean O’Casey and perhaps John Millington Synge, but this was 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 purpose of *A Corpus of Irish English* is to illustrate the language traits of Irish English at their most salient, especially from a historical perspective. This has meant that some authors are included who are not normally regarded as particularly valuable from a literary point of view, e.g. Lady Augusta Gregory and Dion Boucicault. Others are represented by works which are not necessarily regarded as their masterpieces but which clearly show their portrayal of Irish English. This is evident in the case of Shaw whose *John Bull’s Other Island* is the only work included by this author for just that reason. The same applies to Yeats who is represented by the plays *Cathleen ni Hoolihan* and *The Countess Cathleen*. As the corpus includes historical material, it seemed sensible to start from the earliest attestations of Irish English and divide the texts according to period and genre. In the older period of Irish English only poetry is to be found and it was included although some of it is perhaps dubious in its dialectal authenticity but which for want of other material must be accepted.

The texts of the current corpus are largely literary in nature. This has to do with availability and relative certainty of source. There are, however, some other possible avenues which could be explored in the construction of historical corpora. One of these is presented by court verbatim transcripts, from the Old Bailey for instance, which exist...
for Irish defendants over the centuries and which could be assumed to be a relatively accurately rendering of at least the syntax of Irish English during previous periods.

2.1 Structure of corpus

The literary remains of Irish English can be traced back to the early 14th century. Despite the relatively long period for which there are documents their actual number is small, certainly when compared to that for England in the same period. The remnants of medieval Irish English can be counted on the fingers of one hand. The situation improves in the early modern period which can be seen as beginning around 1600. In the early 17th century the English put a policy of determined settlement (plantation) into effect (Palmer 2000), first of all in the north of the country with Scots settlers (and some English) who were encouraged by James I of England to move to Ulster and fill the vacuum left behind after the defeat of the Gaelic forces in this region.

These plantations reached their culmination in the mid-17th century with the re-allocation of land in the east to English mercenaries and servants loyal to Oliver Cromwell. There was a parallel expulsion of the native Irish from these lands to the poorer west of the country which increased in population as a result and which was the area where Irish survived longest.

For the early modern period a number of text types are available. The most frequent are satirical pieces in which an Irish character makes an appearance and is usually ridiculed. Prose sketches in this vein are also common and one has direct portrayals, if somewhat brief, of Irish English by Irish writers as well. Swift’s *Irish Eloquence* and *A Dialogue in Hybernia Stile* fits into the latter category.

Various prose fragments are contained in the present corpus as these document the transition period from a mainly Irish-speaking country to one in which the former native language is recessive and no longer of importance for the development of English. The formative early modern period is one in which much transfer from Irish into English can be observed, no doubt due to the considerable bilingualism which was typical for broad sections of the population during the changeover from Irish to English.

For the nineteenth century drama is the most important genre. It accounts for the large number of plays in the corpus for this period. With the literary revival at the end of the 19th century the emphasis is no longer on comic portrayals of more or less ridiculous Irish figures (see Duggan 1969 [1937], Leerssen 1996, Romani 1997) but with a more authentic representation of Irish characters as illustrated by the works of Synge. Dramatic realism is to be found in the plays of O’Casey which contain dialogue representing the language of the Dublin working classes.

Alongside the many plays of linguistic interest in the nineteenth century there are also novelists and short-story writers who used Irish English in their work in conscious contrast with more standard forms of English for deliberate effect. This is true of another writer represented *A Corpus of Irish English*, the northern Irish writer William Carleton (1794-1869) in whose *Traits and Stories of the Irish Peasantry* (1830-33) Irish English is found in the speech of the socially low-standing peasants. This situation is different from that in Synge where there is no discrimination on the basis of language (Todd 1989: 73), especially as there is no standard usage in the plays with which the more idiosyncratic forms of the language could contrast.
2.2 Questions of genre

A Corpus of Irish English consists largely of drama. Because the primary aim of the corpus is to offer written representations of Irish English it appeared sensible to concentrate on the genre in which the spoken word is central. The decision to concentrate on drama for the modern period meant that prose writers were largely excluded. Nonetheless some prose was integrated into the corpus. For instance, the 19th century Northern Irish writer William Carleton (see previous paragraph) is represented by a selection of these tales. This is also true of the Banim brothers, John (1798-1842) and Michael (1796-1874), from whose Tales of the O’Hara Family (6 vol., 1825-26) a set of extracts were taken.

3 Periods represented in the corpus

3.1 Medieval period

The first period of Irish English began with the arrival of mercenaries from Wales in 1169 and lasted until the final defeat of the Irish by the English in 1601 (during the Tudor conquest of Ireland, Moody and Martin (eds) 1967: 174ff.). In this period the available linguistic material is scanty. Most 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 certain liberties in punctuation and expanding abbreviations (Kosok 1990: 22). There is also a more recent edition by Angela Lucas (see Lucas 1995) 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 poems ‘The virtue of herbs’ and ‘On blood letting’ (Zettersten 1967). 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), comments on which are to be found in Henry (1958).

For the present corpus only the poetry just mentioned has been incorporated. An analysis of the linguistic features of these texts is to be found in Hickey (1997). For more general remarks on the initial stages of Irish English, see Hickey (1993).

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), 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 1988) which has been influenced by Irish, at least lexically. It survives only in the form of glossaries (Vallancey 1788 and Poole, published by Barnes in 1867) which were compiled at the end of the 18th and the beginning of the 19th centuries respectively before the dialect died out. Both these glossaries are to be found in A Corpus of Irish English, 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 supplied database management software (see descriptions in Hickey 2003a).
3.2 Early modern period

When considering the early modern period in Ireland (from 1600 onwards) a strict distinction in types of English, see (1) and (2) below, must be made. This distinction has continued to be important up to the present day.

1) More or less genuine representations of Irish English by native Irish writers.
2) Stretches of texts by non-Irish writers where a non-native perception of Irish English is found.

There is a remarkable amount of material available, 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 forms the bulk of this section of the corpus. Notable among the dramatists whose works are to be found here are William Congreve (1670-1729), Richard Brinsley Sheridan (1751-1816), George Farquhar (1678-1707), see below. There are a few prose samples from this period, such as The Irish Hudibras, probably by James Farewell, and Swift’s A Dialogue in the Hybernian Stile.

The value of type (2) literature should not be underestimated. 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]; Kosok 1990: 61ff.; Leerssen 1996; Romani 1997, Sand 2000).

3.3 The 19th and 20th centuries

The previous two centuries are represented in the corpus mostly by drama because, as outlined above, this genre is likely to contain most examples of specifically Irish English features given that 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 (see his The Aran Islands 1899) 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).

4 Structure of the corpus
The sections of *A Corpus of Irish English* have been arranged in such a way that when the corpus has been installed using the software supplied in Hickey (2003a) — known as *Corpus Presenter* — the following hierarchical structure will be shown. The necessary control files for the tree to be shown within the software are supplied with the CD-ROM which accompanies Hickey (2003a).

**Introduction**
Overview of Corpus

**Middle Ages**
- Kildare Poems (1)
- Kildare Poems (2)
- A Treatise on Gardening
- *The Virtue of Herbs*
- The Pride of Life

**Forth and Bargy**
- Vallancey’s glossary (1788)
- Poole’s glossary (early 19c)

**Fingal**
- The Fingallian Dance (1650-60)
- The Irish Hudibras (1689)
- Purgatorium Hibernicum (1670-5)

**Drama**

16th century
- Anon.: Sir John Oldcastle (1599/1600)
- Shakespeare: Henry V (1599/1623)
- Anon.: Captain Thomas Stukeley (1596/1605)

17th century
- Cuffe: The Siege of Ballyally Castle (1642)
- Dekker: The Honest Whore Part II (1605/1630)
- Dekker: Old Fortunatus (1599/1600)
- Head: Hic et Ubique (1663)
- Jonson: The Irish Masque (1613 /1616)
- Randolph: Hey for Honesty (c. 1630/1651)
- Shadwell: The Lancashire Witches (1681/1682)
- Anon.: The Welsh Embassador (1623)

18th century
- Breval: The Play is the Plot (1718)
- Centlivre: A Wife Well Managed (1715)
- Congreve: The Way of the World (1700)
- Farquhar: The Beaux’ Stratagem (1707)
- Farquhar: The Twin Rivals (1702/1703)
- Goldsmith: She stoops to conquer (1773)
- Michelbourne: Ireland Preserved (1705)
- R. B. Sheridan: The School for Scandal (1777)
- R. B. Sheridan: St. Patrick’s Day or The Scheming Lieutenant
- Th. Sheridan: The Brave Irishman (1740/1754)

19th century
- Boucicault: The Colleen Bawn (1860)
- Boucicault: Arragh na Pogue (1864)
- Boucicault: The Shaughraun (1875)
- Gregory: Hanrahan’s Oath
- Gregory: On the racecourse
Gregory: Spreading the news
Gregory: The workhouse ward
Wilde: The Importance of Being Earnest (1895)
Yeats: The Countess Cathleen (1899)
Yeats: Cathleen Ni Houlihan (1902)

20th century
Shaw: John Bull’s Other Island (1904)
Synge: In the Shadow of the Glen (1903)
Synge: Riders to the Sea (1904)
Synge: The Well of the Saints (1905)
Synge: The Playboy of the Western World (1907)
Synge: The Tinker’s Wedding (1909)
Synge: Deirdre of the Sorrows (1910)
O’Casey: The Shadow of a Gunman (1923)
O’Casey: Juno and the Paycock (1924)
O’Casey: The Plough and the Stars (1926)
O’Casey: The Silver Tassie (1928)
Behan: The Quare Fellow (1954)
Behan: The Hostage (1959)

Novels
19th century
Edgeworth: Castle Rackrent (1801)

Prose
19th century
Banim: O’Hara Tales (1825-26)
Carleton: Traits and Stories (1830-33)

Varia
17th century
Anon.: Bog-Witticisms (c. 1687)
Anon.: Páirleamh Chloinne Thomáis (1645-50)
Anon.: The Irishmen’s Prayers (1689)
Anon.: John Dunton, Report of a Sermon (1698)
18th century
Anon.: Peadar Ó Doirnín, Muiris Ó Gormáin (1730-40)
Anon.: The Pretender’s Exercise (?1727)
Anon.: A Dialogue between Teigue and Dermot (1713)
Swift: A Dialogue in Hybernian Stile (c.1735)

It is obvious from the above list that there are different text types contained in *A Corpus of Irish English*. These texts can be examined using the options contained in the *Corpus Presenter* software suite. For instance, there are modules for examining databases (which contain the glossaries for Forth and Bargy) and for interrogating the text files which contain prose and drama. Furthermore, when examining texts, one can specify if the entire collection (the whole tree as displayed on the computer screen) is consulted or just the texts in a single branch. One can also select individual files so that only these are subject to examination.

5 Interrogating the corpus
When specifying the parameters for a search with the supplied software various options are available to the user. One can search for simple strings, one can carry out searches using spelling variants or grammatical variants of words or sets of words which are determined by the user. There are many additional means of organising searches, for instance, it is possible to specify a syntactic frame and use this when examining files of the corpus. This option is especially useful in the context of Irish English. This variety and its sub-varieties are characterised more by non-standard syntactic features than by irregular morphology, a fact which is characteristic of varieties which owe their origin to language contact rather than to a long period of historical continuity, as is the case with Scots for instance. In the latter one finds more non-standard grammatical forms than in Irish English, compare the morphological variants as described in McClure (1994) with the few non-standard morphological forms as discussed in Hickey (2003c). This fact is in keeping with observations on shift-induced varieties of languages as shown in the case studies contained in Thomason and Kaufman (1988).

Searches using syntactic frames are useful when trying to determine if syntactic structures occur in the texts of a corpus. For example, it is possible to search for attestations of the aspectual types known to occur in Irish English to see what the distribution is like diachronically. Consider the immediate perfective which is expressed in Irish English by using the temporal adverb *after* followed by a continuous form of the verb as in *He’s after breaking the glass* ‘He has just broken the glass’ (Kallen and Kirk, volume one, also treat this structure). This structure is a calque on a source construction in Irish which uses the adverb *tar éis* ‘after’ with a non-finite verb form for the same purpose (for further details concerning the not entirely undisputed origin and present-day distribution of this structure in dialects of Gaelic, see Ó Sé 2004).

To locate instances of the immediate perfective in *A Corpus of Irish English* with the *Corpus Presenter* software all one needs to do is enter a syntactic frame as indicated in the following. When specifying the information for a search one can furthermore say whether a string represents an entire word or just a part of one. If the latter is the case, one may specify if this occurs at the beginning or end of a word (see rightmost column in the following table). There are a number of other parameters which can be set for a search and which increase flexibility and the likelihood of accurate returns, for instance, one can determine the number of possible intervening items between String 1 and String 2 in a search and use an input list for each of the strings, if necessary (see below). This option allows one to search for strings using spelling variants, an important consideration when looking at historical texts (see Barbiers, Cornips and Kunst (volume one) for similar discussions of the use of search engines used to interrogate databases).

Table 1  **Immediate perfective in Irish English**

<table>
<thead>
<tr>
<th>Syntactic frame</th>
<th>Entry</th>
<th>Scope</th>
</tr>
</thead>
<tbody>
<tr>
<td>String 1</td>
<td><em>after</em></td>
<td>whole word</td>
</tr>
<tr>
<td>String 2</td>
<td><em>ing</em></td>
<td>end of word</td>
</tr>
<tr>
<td>Intervening items: 1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Input list: —</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Type of return**

I’m [after] walk[ing] hundreds and long scores of miles...
The returns given in the box above show typical instances of the immediate perfective construction of Irish English (Hickey 1995). As can be seen from the returns it would be possible to specify that no intervening items occur between String 1 and 2 as in this structure the continuous form follows immediately on the temporal adverb after. If more intervening items were allowed then spurious returns like the following might occur: ‘... and there were men coming [after] them, and they hold[ing] a thing in the half of a red sail ...’ (Synge, Riders to the Sea), illustrating a different type of nonstandard structure (Filppula 1991), but not the immediate perfective of Irish English.

As mentioned above it is often sensible to use not a single string but an input list, any member of which can occupy the position of the string it is associated with. If one were examining the plays of Sean O’Casey, for instance, one would need to use an input list as there are two forms in his dramas, after and afther (the latter indicating the dentalisation of /t/ before /-r/ to [t] which is found historically in many forms of English, see the treatment of this issue in Hickey 1987).

Input lists are also useful if one wishes to check on inflected forms in a corpus. To take an example from O’Casey again: the word mot ‘girlfriend’ is typical of local Dublin English (the meaning evolved from ‘something very small’ via ‘something precious’ to its present meaning, Dolan 1998: 180). One might suspect that mot would occur in one of the plays. Indeed it does, but only in an inflected form, cf. ‘Never held a mot’s hand’ from The Plough and the Stars. Now if one searched for mot as an entire word then there would be no return. However, one could construct a small input list with mot, mot’s so that both forms would be searched for and the second would be returned from the play just mentioned. Yet another use of an input list would be to collect words which thematically belong together. For instance, one might be interested in words which indicate poor health and then devise an input list with items like ‘weak, poorly, feeble, dawny’, etc. The last item is a regionalism from English which is first attested in the 17th century in Ireland, retaining its meaning of ‘weak, of poor health’ (Dolan 1998: 93), and is attested in the plays of Sean O’Casey: ‘I’m terrible dawny, Mrs. Burgess’ (The Plough and the Stars).

A semantic issue in connection with the immediate perfective which has been the subject of discussion among scholars (see McCafferty 2003) concerns the time reference it is embedded in. In present-day Irish English it has an exclusively past time reference but originally it seems to have referred to the future, perhaps because it was influenced, at the beginning, by the use of English after as in He’s after his dinner ‘He is looking for his dinner’. Again an examination of A Corpus of Irish English helps in this respect. The returns with Corpus Presenter (see following graph) show that the instances with future time reference declined rapidly in the mid-1700s while the past time reference increased dramatically and continues as the only interpretation of this aspect in present-day Irish English (Hickey 2003a: 18f.).
Figure 1 *Shift in time reference for the after perfective as attested in A Corpus of Irish English*

![Graph showing time reference shift for the after perfective over centuries.](image)

A further example of a syntactic search in *A Corpus of Irish English* is offered below. It concerns the possible occurrence of the habitual aspect which is frequently marked by *do* and *be* in the present in Irish English. For the search which is illustrated in the following table, the forms *do, does, don’t* and *doesn’t* were used in an input list. As *be* is the only form found in habitual constructions, an input list was not necessary for the second string.

**Table 2  Habitual aspect in Irish English**

<table>
<thead>
<tr>
<th>Syntactic frame</th>
<th>Entry</th>
<th>Scope</th>
</tr>
</thead>
<tbody>
<tr>
<td>String 1</td>
<td><em>do, does, don’t, doesn’t</em></td>
<td>whole word</td>
</tr>
<tr>
<td>String 2</td>
<td><em>be</em></td>
<td>whole word</td>
</tr>
<tr>
<td>Intervening items:</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>Input list:</td>
<td>yes (for 1)</td>
<td></td>
</tr>
</tbody>
</table>

**Type of return**

Is it often the polis [do] [be] coming into this place, master of the house?  
the clumsy young fellows [do] [be] ploughing all times  
the like of them [do] [be] walking abroad with the peelers  
...the way the needy fallen spirits [do] [be] looking on the Lord?  
(Synge, *The Playboy of the Western World*)
The returns show that this structure is attested quite frequently in the dramas of Synge (just one of which is quoted here for reasons of space). The matter of the habitual in Irish English does not rest here as there are a number of alternative expressions of this aspectual type. In the north of Ireland it is common to find the form *bees* for the habitual, e.g. *John bees working at the week-ends* so any consideration of this category would involve a search for this as well. In addition a verbal *-s* on persons outside the third person singular – especially the first person – may also indicate the habitual, e.g. *I meets my sister on a Friday afternoon*. Searches for such forms would also be necessary.

### 5.1 Checking on standard wisdoms

#### 5.1.1 The attestations of the habitual

The example of the habitual just discussed might appear fairly simple but it has far-reaching consequences. These result chiefly from the time-scale of the attestations of the habitual in diachronic Irish English. The texts of *A Corpus of Irish English* show that the habitual is not attested before the middle of the 19th century (see Hickey forthcoming for a detailed discussion of this). There are a number of problems with this finding. It is fair to say that the majority of scholars working on Irish English see the rise of the habitual in Irish English as a consequence of contact-induced change (Hickey 1995, Filppula 1999) during the language shift which set in earnest in the 17th century and which was largely concluded by the late 19th century. If this interpretation is correct then one would expect to find attestations of the habitual in Irish English texts from the 17th century onwards. However, these are noticeably absent, something which is not true of the immediate perfective with *after* (although this initially had future reference). There are two possible conclusions from this fact. The first is that habitual did indeed just arise at the beginning of the 19th century and hence is rightly only to be found from the 1840s onwards. The second conclusion is that the textual attestations of Irish English are inaccurate, i.e. that the habitual arose in the 17th century but did not find its way into texts until well into the 19th century. The value of text corpora like *A Corpus of Irish English* is in highlighting attestations and hence sharpening the awareness among researchers for the historical interpretation of well-attested features like the habitual aspect.

#### 5.1.2 Second person plural distinctions

Nonstandard varieties of English normally have some means to distinguish between singular and plural second person pronouns. One need only think of *y’all, y’uns, youse, ye, yez*, the Caribbean *unu* or Tok Pisin *yupela*. The form of interest in the current context is *youse* which is reputedly of Irish English origin. The rise of the habitual Irish English aspect in the 19th century is a likely reason for this. With the demise of *thou* in Ireland, *you* came to be understood as the pronoun with singular reference and the gap to be filled therefore was that in the
plural. The Irish second person plural pronoun is *sibh* [ʃ̂v], phonologically unlike anything available in English then or since. The Irish solution was to find a form which was different from *you* and which could function as a plural pronoun. Basically, there were two pathways open at that time. The first was to use the inherited *ye* as a marker of second person plural. The second, and apparently later option, was to create a synthetic plural by appending the regular plural suffix -*s* to the already present *you*, yielding *youse* [ju(ː)z]. Later a combined form arose, *yez* [jiz] or with reduction to [jez] or [jəz]. There would seem to be a chronological sequence involved here as shown in the following table (see Hickey 2003c for more details).

**Table 3  Plural second person pronouns in Irish English**

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>ye</td>
<td>from 12th century onwards</td>
</tr>
<tr>
<td>youse</td>
<td>not before early to mid 19th century</td>
</tr>
<tr>
<td>yez</td>
<td>not before mid 19th century</td>
</tr>
</tbody>
</table>

*Youse* is a regular plural formation by simple attachment of suffixal -*s* to the existing pronoun *you*. If one is making a case for *youse* being a specifically Irish development then one must exclude any English source. With the help of available text corpora this issue can be resolved with reasonable certainty. For instance, the sampler of the *Early English Correspondence Corpus* (Nevalainen 1997; Raumolin-Brunberg 1997; Nevalainen and Raumolin-Brunberg 2003) does not reveal a single instance of *youse*, although *ye* and *thou* abound (*thou* is by far the most common second person pronoun, 372 instances, with *ye* occurring 19 times). This holds for the 23 texts in the public domain version of this corpus, covering letters from the end of the 16th to the end of the 17th century).

Equally in the 138 texts of the Early Modern English section of the *Helsinki Corpus of English Texts* there is not a single instance of *yous(e)* or *ye(e)z*. (a further Irish form from *ye + {S}*). The situation for Ireland can be seen by examining *A Corpus of Irish English*. Here the form *yous(e)* occurs abundantly in the plays of John Millington Synge (1871-1909) and with later writers like Sean O’Casey (1884-1964) and Brendan Behan (1923-1964).

A further fact can be cited here to underline the Irish origin of *youse*. The form is found in England in only a few areas, Liverpool (Trudgill 1986: 139-141), Newcastle (Beal 1993) and in Scotland in Glasgow and spreading out from there in central Scotland (Macafee 1983: 51). It is hardly a coincidence that these are the areas of Britain with greatest Irish influence.

If the above considerations justify the assumption that *youse* is not English but Irish, the next question would be: when did it arise? When one looks at earlier writers a different picture emerges from that later on: Maria Edgeworth (1767-1849) in her novel *Castle Rackrent* (1800), which attempts to display the speech of the native Irish realistically, has many instances of *ye* (11) but not a single one of *yous(e)*. The *Oxford English Dictionary* gives the earliest attestations from Samuel Lover’s novel *Handy Andy. A tale of Irish life* (1842). This in general concords with *A Corpus of Irish English* where no attestations are to be found before the mid 19th century, a similar picture to that for the habitual aspect (see above). The conclusion to be drawn from these observations is that while *youse* can be safely regarded as Irish in origin, it does not date from the beginning of the major language shift (17th century), probably
because the original *ye* was used initially by Irish speakers and only later did the analogical form *youse* arise (much as did negative epistemic *must* as in *She musn’t be Scottish*, Corrigan 2000). Support for this contention comes from the greater occurrence and acceptance of the *ye* form along the western seaboard in areas where Irish maintained itself longest (see the discussion of *A Survey of Irish English Usage* in Hickey 2004a).

6 Further results yielded from corpora

6.1 False assumptions

One obvious use of text corpora is to confirm or refute opinions which have perhaps been held without any serious degree of questioning by scholars in the field. The rise of such opinions may be justified in some original observation, but the continuation may not be supported by attestations in a dialect. A good example of this situation can be seen when looking at continuous forms of the type *a*-V-ing, e.g. *They were out a-playing on the strand*. When investigating such structures some authors have pointed out the structural parallel with Irish and Scottish Gaelic (Majewicz 1984). Consider the Irish rendering of the English sentence just given: *Bhí siad amuigh ag imirt ar an trá*. [was they out at playing on the strand]. However, this obvious parallel would appear to be coincidental. The structure *a*-V-ing is well attested in British English during the colonial period, deriving historically from *on V-ing* with phonetic reduction of the preposition *on* much as in *asleep* from an earlier *on sleep*. This may well be the source for those varieties of American English which show this structure as Montgomery (2000), who is sceptical of the Celtic origin, rightly points out. To consolidate this view one can examine the texts in *A Corpus of Irish English* which has only a very few attestations in the many historical texts where one might expect the type *a*-V-ing to occur. In an examination of all the plays from the 19th century in *A Corpus of Irish English* only 2 of the 128 returns for the syntactic frame *a*-V-ing were instances of the structure being discussed here (the instances were *a-milking* and *a-waiting* in plays by Dion Boucicault where the author used a hyphen after the *a* to indicate the structure in question). Significantly, none of the plays of Lady Gregory, which are replete with putative Irish English structures, show instances of *a*-V-ing.

6.2 Slight attestations

The example just given is one where a small number of attestation suggest that Irish English is not the source of a feature found in varieties of English today, for instance in forms of American English (Montgomery 2000). Such slight attestations may also be simply the remnants of a feature which died out completely. This is the case with archaic features of English morphology which may have been present with early settlers but which were later abandoned. For instance, there is a proclitic form for the first person singular, *ch* ‘I’, (first noted by Alexander Gil in his *Logonomia Anglica* from 1621, Ihalainen 1994: 200) which is attested in the archaic dialect of Forth and Bargy (Hickey 1988) but is confined to this one case in Ireland.
6.3 The neglect of distinctions

A small number of attestation may, however, have a further source in a dialect corpus. This is the relative neglect of distinctions found in other varieties, specifically in more widespread varieties of British English. A clear example of what is intended here is provided by the use of the so-called ‘extended present’ of Irish English (Filppula 1997). By this is meant the use of a present form of a verb to encompass an action which stretches back into the past. In such cases, for instance, in sentences with the temporal adverbial since, e.g. *He has been here since we moved to Dublin*, English requires the present perfect. However, Irish English only uses the present and so neglects the tense distinction found in standard English, e.g. *We’re living here for ten years now*. A significant source for this usage in Ireland (it is also found in Scotland) may well be Irish where an equivalent to the present perfect of English does not exist.

Quantifying the neglect of distinctions is notoriously difficult and it must be admitted that corpora cannot always be of assistance here. Essentially, what is required is a list of contexts in which the distinction which is being examined might have occurred. Then one could count those in which it did actually occur. There are, however, ways of doing this with corpus processing software. For example, one could search for all contexts with the temporal adverb since and then manually assess whether they would contain the present perfect in more standard forms of English and whether they do so in the corpus texts of Irish English.

6.4 Deciding oneself

The returns of a retrieval run frequently require interpretation by the user of the corpus. More often than not, it is necessary to manually assess whether the returns are genuine examples of the structure one is looking for. This fact should not, however, be regarded as a reason for not using a corpus in one’s investigations. Sophisticated software with a user-friendly interface will allow users to decide quickly and easily what returns are spurious and what or not and will provide the option of saving the genuine cases.

The following example is an instance of information retrieval with manual assessment. The concern here was with the non-standard past tense forms seen and done. In the history of English the forms for the preterite and the past participle have varied considerably (Lass 1994, Cheshire 1994) and non-standard varieties have shown a tendency to use fewer different forms in the past, often just one for both finite and non-finite forms of a verb. This situation also applies to Irish English. Of the forms under consideration here, seen is by far the more common of the two.

Table 4  **Occurrences of preterite seen and done in Irish English plays**

<table>
<thead>
<tr>
<th>Authors</th>
<th>seen</th>
<th>done</th>
</tr>
</thead>
<tbody>
<tr>
<td>O’Casey (4 plays)</td>
<td>44</td>
<td>12</td>
</tr>
<tr>
<td>Synge (6 plays)</td>
<td>46</td>
<td>0</td>
</tr>
</tbody>
</table>

The above figures were gleaned by manually separating out the instances of the preterite from those of the past participle. The manual assessment of the returns did, however, reveal an interesting fact about the use of preterite seen. This is its almost exclusive
occurrence in the first person. It would appear that it is favoured as a form for narration, i.e. when a character in a play is recounting an incident or offering information. This type of finding is typical of manual evaluation of retrieval returns from a corpus: it was not what was being looked for and only came to light incidentally. As it turned out, 17 of the 44 instances in O’Casey’s plays were from *Juno and the Paycock* which has a large amount of first person narration. The same is true for *The Playboy of the Western World* by Synge which accounts for 18 of the returns. The play *Juno and the Paycock* also had the highest incidence of preterite *done*, but here (as in the other plays) it occurred in the third rather than in the first person (the ratio was 5:1 in *Juno and the Paycock*). Furthermore, one can note that Synge has not a single incidence of preterite *done*. Neither Lady Gregory nor Dion Boucicault do either, going on the plays in *A Corpus of Irish English*. Boucicault does, however, have an incidence of first person preterite *seen*.

### 6.5 Shared features

Finally it should be stressed that the attestation of a usage or structure in a dialect corpus like *A Corpus of Irish English* must be relativised and cross-checked in similar corpora for other varieties if such collections are available. For instance, the adverb *never* is used in Irish English to mark a punctual event in the past as in *She never called us yesterday*. Attestations of this usage are to be found in the Irish English corpus (and in present-day corpora of Irish English, see discussion by Kallen and Kirk in volume one), but the usage is common in northern England and in Scotland as well so that no exclusive claim can be made for it.

### 7 Conclusion

The present chapter presents an overview of a dialect corpus, in this case *A Corpus of Irish English*, which has been put at the disposal of interested scholars (Hickey 2003a). The collection contains the majority of texts which are available for the history of this dialect, or more accurately, this set of dialects. The different types of texts are discussed and the nature of the language which they represent is referred to. Some of the pitfalls and advantages associated with exploiting these texts for linguistic research are pointed out. In all, the latter seem to predominate and a number of sample analyses show the uses to which a careful and objective analysis of corpus data can be put and the benefits to be accrued from this.

### References


Vallancey, Charles (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 Wexford, Ireland, in 1167, 1168, 1169’, *Transactions of the Royal Irish Academy* 2, 19-41.