



Anderson, W. (2007) *The SCOTS Corpus: a resource for language contact study*. In: *Language contact and minority languages in Europe*. Studies in Eurolinguistics, 5. Logos Verlag. ISBN 9783832516444

<http://eprints.gla.ac.uk/4682/>

Deposited on: 17 March 2009

Wendy J. Anderson

The SCOTS Corpus: a resource for language contact study

1. Introduction

The Scottish Corpus of Texts and Speech (SCOTS)¹ is an ongoing project in the Department of English Language, University of Glasgow.² The ultimate aim of the project is to create a large electronic corpus of both written and spoken texts for the languages of Scotland, which will reflect the linguistic situation in current-day Scotland. The corpus will be freely accessible and searchable on the Web.

A corpus such as this of course can have any number of applications: in addition to studies of language contact, corpus methodology makes possible various types of self-contained investigation into the lexis, grammar, phraseology, pragmatics etc., of the languages of Scotland. The aim of this paper is to introduce the corpus, and highlight some of its possible applications in the study of language contact. Contact has been historically important for the languages of Scotland, and this is no less true today: I hope therefore to suggest ways in which the resource will be valuable for such research in the very near future.³

2. Background to the corpus

As the first corpus project focusing on modern texts in the languages of Scotland, SCOTS fills a gap in the corpus resources available for research. As far as the Scots language is concerned, older texts are well catered for by the Helsinki Corpus of Older Scots, which covers Scots material from the period 1450-1700⁴, and the Edinburgh Corpus of Older Scots, for the period 1380-1500.⁵ Other projects, such as the British National Corpus (BNC), contain some material in current-day Scottish English (including some spoken texts with Scottish speakers),⁶ but were not designed to reflect Scottish usage. By providing a

¹ The Scottish Corpus of Texts and Speech (SCOTS) is a venture by the Department of English Language and STELLA project at the University of Glasgow. The first stage of the project was grant-funded by the Engineering and Physical Sciences Research Council (EPSRC Grant no. GR/R32772/01): the current three-year stage is funded by the Arts and Humanities Research Board (AHRB Grant no. B/RE/AN9984/APN17387). The project website can be found at www.scottishcorpus.ac.uk.

² The author is postdoctoral Research Assistant on the SCOTS project (since May 2004). The project's Principal Investigator is Dr. John Corbett. The other members are Professor Christian Kay, Dave Beavan (Computing Manager), Jean Anderson, Dr Jane Stuart-Smith, Flora Edmonds, Louise Edmonds and Cerwyss O'Hare.

³ I would like to thank Dr. John Corbett and Professor Christian Kay for the invaluable suggestions they made to a draft of this paper.

⁴ See for example Meurman-Solin 1995.

⁵ For information on both of these corpora, see Meurman-Solin 2003. Meurman-Solin is also preparing a Corpus of Early Scottish Women's Writings, 1500-1800 (CESWW).

⁶ See the BNC bibliography at <http://sara.natcorp.ox.ac.uk/BNCbib/>

medium-sized corpus of (in the first instance) present-day Scots and Scottish English, SCOTS opens up a number of avenues for research, both as a single corpus, and in conjunction with other corpora for comparative purposes.

The early stages of the project involved setting up the infrastructure, and beginning to identify texts for inclusion. This resulted in a pilot corpus, which will form part of the corpus which we intend to launch publicly in November 2004. After its launch, the corpus will continue to grow as more texts are identified, processed and added. Our target over the next three years is 4 million running words of text (see below for further details).

Surprisingly little is known with much certainty about the current-day language situation in Scotland, such as the number of speakers of Broad Scots today, the range of functions for which Scots is used, and distinguishing features of Scottish English compared with the English of England. Any corpus which aims to reflect or represent the linguistic situation of Scotland of course also has to consider a number of languages other than Scots and Scottish English. While the focus of the current stage of the project is on these two varieties, subsequently we also aim to extend the corpus to texts in Gaelic, and non-indigenous community languages, such as Punjabi, Urdu and Chinese. This may take the form of compatible subcorpora which can be slotted in to the existing SCOTS framework.

The timing of the project fits very neatly with a heightening of interest in Scotland, its languages and its literature, which has of course come hand in hand with devolution from the United Kingdom parliament, and the re-established Scottish Parliament, which was formed in 1999, and which settled into its new accommodation at Holyrood in Autumn 2004.

3. Corpus methodology and the nature of the corpus

Corpus methodology is ideal for investigating a nation's linguistic situation, and a sensitively designed corpus can open the way for a wide range of different types of research – linguistic and sociolinguistic, as well as cultural or literary. Above all, a carefully constructed corpus enables a researcher both to test intuition, and derive qualitative and quantitative information from the data.

As far as corpus methodology is concerned, experience over the last thirty years has repeatedly shown that theoretical aims and practical constraints do not always correspond. Priorities have to be carefully weighed against each other. Size is of course an issue, but one that is in many cases less important than balance and representativeness. With regard to size, as noted above, SCOTS aims to collect 4 million running words of text over the next three years. At present we have close to half a million words in a state which is ready to be made available – i.e. with all permissions finalised and the texts suitably processed and formatted. If the pilot corpus can be taken to be predictive of the end result, then there are likely to be around 3500 text units. Our policy is to include whole texts where possible, as this is more valuable for investigations into larger linguistic units, such as discourse analysis: where whole-text inclusion is unacceptable for authors or publishers, but where

the text is desirable for other reasons, such as genre balance, however, then only extracts are used.

The corpus will include both written (80%) and spoken material (20%), the spoken material being held both as audio or video files and accompanying orthographic transcriptions. This enables study of the phonetics and phonology of the languages of Scotland, as well as research on the lexical, grammatical, phraseological, and discourse levels, from linguistic and sociolinguistic perspectives. The corpus is roughly synchronic: the texts currently being sought date from 1940 onwards, although this is an arbitrary date and would not preclude the incorporation of earlier texts if these were to become available. It will be possible in a subsequent stage to extend further back in time to include older written material, and of continuing to add to the corpus with more up-to-date, written and spoken, material in the future.

A perennial issue in corpus linguistics is representativeness,⁷ and this is a tricky issue for SCOTS too: indeed it is impossible *a priori* to achieve true representativeness.⁸ Especially where a corpus involves more than one language, and where one or more of these is not a standardised variety, somewhat arbitrary initial decisions have to be taken with regard to the relative proportions to be included. Similar decisions have to be taken in relation to genre coverage: a corpus of Scots which contained equal quantities of a range of genres might be useful for linguistic comparisons of these varieties, but cannot give faithful statistical information about the typical production of the language.

As regards genre and text type, the corpus in its current state already contains quite a wide range. Broad Scots is of course fairly widely used in literature, especially poetry, and has a long and proud literary history.⁹ It is used much less in business, administrative or educational contexts where Scottish Standard English is the usual choice, even of people who speak Scots on a daily basis.¹⁰ SCOTS already contains some fairly unusual genres of Scots: emails, journalistic writing, letters, as well as large quantities of prose fiction, poetry (including some in translation from other languages), drama texts, lectures, student essays, and conversation.

The genre balance of the corpus is something which will require continuous monitoring over the course of the project, while recognising with Leech that “We should always bear in

⁷ For a discussion of this issue, see Atkins, Clear and Ostler 1992, who explain that “the compiler of a general language corpus will have to evaluate text samples on the basis of *both* reception and production.”

⁸ See Douglas 2003 for a discussion of corpus design for a corpus of non-standardised language varieties, and the impracticality of stratificational sampling in such a case.

⁹ To mention briefly just a few of the most well-known: John Barbour’s *Bruce* from the late 14th century, through David Lyndsay, Robert Burns, Sir Walter Scott, Robert Louis Stevenson, Lewis Grassie Gibbon, right up to the present day and Irvine Welsh’s internationally famous *Trainspotting*.

¹⁰ It is interesting, however, to read on the Scottish Parliament’s website a version of its Parliamentary Guide in Scots. The website also contains versions in English, Punjabi, Chinese, Urdu, Bengali, Arabic, Gaelic, British Sign Language, and as of September 2004, Catalan, Spanish, Italian, German, French and Russian. See www.scottish.parliament.uk.

mind that the assumption of representativeness must be regarded largely as an act of faith” (Leech 1991, p. 27). The corpus already contains a fair variety of genres, and the next stage of work is to widen as well as deepen this variety, in order to try to achieve a balance. The texts that we have acquired so far have come from a number of sources: in particular from the general public responding to requests for texts on our website and early publicity material. We still, however, need a lot more texts, and would be delighted to follow up any leads or offers of material.¹¹

4. Issues and metadata

As for any corpus project which seeks to make recent texts publicly available, one of the most evident issues arising is that of copyright. We need to obtain explicit written permission from all authors of texts included in the corpus, and participants involved in audio material or transcriptions. This is done through a series of specially-designed forms completed by individuals submitting texts, and all of these permissions are recorded in the SCOTS administrative database, associated with the corpus. This database also enables us to log communications with contacts (such as authors, participants and submitters).

Another very important feature of the database is the sociolinguistic metadata which it contains, which makes the corpus texts more valuable for linguistic research. Contributors are asked to provide answers to a number of questions, giving demographic and linguistic information.¹² With the contributor’s consent, this data is then available to corpus researchers.

The demographic information includes: the author’s place of birth, place of residence, parents’ places of birth, educational attainment, occupation, parents’ occupations and religion. We also ask about the languages spoken, written, understood and read, and the normal context of use of these languages. This information, or as much as it is possible to collect in each case, is entered for all authors of written texts which are included in the corpus, and the participants in all of the spoken texts we include. We also collect considerable detail about the texts submitted, such as publication, or translation information as appropriate, text type and medium, intended audience, participant roles if the text is audio or video, the issue of awareness of recording, and the level of spontaneity of spoken language production.

These two main types of metadata combined will make generalisations possible, and allow correlations to be drawn, for example between a speaker’s place of birth and features of pronunciation, authors’ social background and lexical items, or grammatical constructions and textual genre. For researchers with an interest either in language contact in Scotland specifically, or in the general mechanisms of language contact, this is valuable information. A borrowing which is found to be spreading through Scots may be able to be tracked, if not to its point of entry, then at least to likely types of speaker involved in innovation. Since it

¹¹ Readers who might have material to contribute to SCOTS are invited to contact the project by email on scots@arts.gla.ac.uk.

¹² To reduce the burden on contributors, and with a view to providing comparable data for each text, for the most part this involves tick-box responses.

is one of the purposes of the corpus to enable patterns to be highlighted, the metadata included with texts must be as full as possible, in order not to pre-empt correlations, and to enable the material to speak for itself.

The metadata recorded about texts and authors feeds into the web-based search facility offered by SCOTS. The main search page at present allows for a number of types of search. A researcher can browse the titles of the texts contained in the most up-to-date version of the corpus, currently ordering the information returned by either author or title. Alternatively, a search can be carried out for texts of either written or spoken mode (in each case including or excluding poetry as required), or for texts of a particular time period. The author's surname, gender or region of birth or residence can also be used to restrict the texts returned. It is intended, in a later stage of the corpus, to enable searches on all of the fields used to record metadata.

Just as useful from a linguistic perspective is the facility for searching for a particular word or phrase in the corpus, either across the whole corpus, or across a defined subsection. We intend to offer a concordance facility in due course. It is also possible to download the texts, either as plain text or xml, and of course to view the metadata recorded about each text.

5. Language variety and contact in Scotland

The language contact situation today in Scotland is fairly complex, perhaps more so than a passing consideration suggests. Scottish English, (Broad) Scots, Gaelic, and a number of non-indigenous community languages are all spoken (with Indian, Pakistani and Chinese as the largest ethnic groups).¹³

To state that the indigenous Germanic languages of Scotland are Scottish English and Scots is arguably an over-simplification in any case. While Broad Scots is undeniably distinct from Scottish Standard English, the nature of their relationship is less certain: they may be considered to be the end-points of a continuum, or, alternatively, completely distinct languages with a clear dividing line.¹⁴ In addition, there is variation within Broad Scots which must be taken into account. Aberdeenshire Scots (the 'Doric') is distinct from southern and central varieties, and the artificially created Lallans variety is different again.¹⁵ Nor is there a clear-cut functional or generic divide between Scots and Scottish Standard English: although there is a tendency for Scots to be used in speech, and Scottish Standard English to be used in written contexts, there is also a significant literature in

¹³ The 2001 census revealed a drop in the number of Gaelic speakers over the previous ten years. In 2001, just under 2% of the combined Scottish population for all areas indicated some knowledge of Gaelic. This rises to just under 9% of the Highland region, and nearly 71% of the population of the Western Isles, however. See www.scrol.gov.uk/scrol/common/home.jsp for fuller census information. Some information on ethnicity is potentially too sensitive to be freely available on this site: however, statistics for household composition by ethnic group of Household Reference Person is available, and shows that 1.4% of households are recorded of an ethnicity other than white. On Scots, see also Macafee 2000.

¹⁴ See for example Aitken and McArthur 1979 for discussion of the nature of this relationship.

¹⁵ See Görlach 2002 for more information on Lallans and its place in Scottish literary history.

Broad Scots. There is not, however, and never has been a written standard variety, for which reason spelling is very variable. This is a major problem in constructing a worthwhile, searchable, corpus of a language like Scots.¹⁶

5.1. Language contact in the past

Language contact in the past in Scotland and involving the Scots language is well-documented.¹⁷ It is the different contact situations in which both Scots and English have found themselves over the centuries that have caused these Germanic languages, with a shared origin in Anglo-Saxon, or Old English, to come to differ to the extent that they do. Scots descends from the Northumbrian form of Anglo-Saxon, with influence over the centuries from a number of sources.

Firstly, there is a strong Scandinavian influence, both from the settlement by the Vikings of northern areas of Scotland in the 9th and 10th centuries and from the Danelaw in England. It is probably this significant component which differentiates Scots most from English. Some well-known examples of Scandinavian borrowings are: *kilt*, *kenspeckle* (recognisable, familiar) and *speir* (ask, inquire).¹⁸

The influence of Dutch and Low Saxon was particularly strengthened by the shared Calvinist faith in both Scotland and the Netherlands, but its influence can be seen in Scots from the Middle Ages up to the 18th century. *Loun* (rascal, or more neutrally, lad, and especially in the variant spelling *loon*, more common than *loun* in SCOTS), *scone*, *trauchle* (to trail, drag etc.), *widdershins* (contrariwise) are all common words which demonstrate the Dutch influence.¹⁹

Closer to home, Gaelic has provided a number of loanwords in Scots (such as *sassenach*, *quaich*, the interjection *och*). This is generally accepted as a relatively small number, at least given the physical closeness. For evidence that the Gaelic contribution is more considerable than is usually appreciated, see however McClure (1986) and Dareau (2001).

A final significant influence comes from the Romance languages, both Latin, the language of religion, the law and high culture, and French. The French influence stretched over a long period of time, and was due to a number of factors. The first wave of French borrowings can be dated to the period immediately following the Norman invasion: from 1066 onwards, the French of Normandy, itself historically tied to Scandinavia, began to have an impact on Scots, owing to the beginnings of the feudal system and the establishment of burghs. By the early 14th century, the influence of Central, or Parisian,

¹⁶ See Douglas 2003 on the issue of spelling variation.

¹⁷ For an excellent general overview of the history of Scots to 1700, see the introduction by C. Macafee to the *Dictionary of the Older Scottish Tongue*, vol. 12, pp. xxvix-clvii. In particular, this introduction includes a discussion of the sources of the vocabulary, a table indicating the relative proportions of loanwords from different sources, and an overview of the grammar and phonology of Older Scots.

¹⁸ See Murison 1979 for a brief overview of the Norse influence on Scots.

¹⁹ See for example Murison 1971 on lexical borrowing from Dutch.

French also began to be felt, as English possessions in France extended, and Norman French began to seem old-fashioned (Murison 1977, pp. 50-51). As Murison has noted, feudal barons spoke French, while their retinue spoke Anglo-Norman (ibid, p. 3).

A second wave of French influence arrived with the Auld Alliance, which was formalised in 1295 (see Corbett et al. 2003, p. 9), and lasted until the Reformation in 1560. While earlier French influence for the most part percolated up through England to reach Scotland, the Auld Alliance allowed borrowings to bypass England: there is therefore often no parallel item in English (see Murison 1977, p. 52). This is the case, for example, for *cordisidron* (from *écorce de citron*, lemon peel)²⁰, *dittay* (from Old French *dité*, indictment in law), *vaig* (to wander idly, French *vaguer*), *gean* (wild cherry, compare modern French *guigne*), all of which are attested in SCOTS. Macafee (1997, p. 206) sees the Auld Alliance as a period prolonging the phase of popular borrowing begun in the 11th century. Indeed, this has not necessarily ended even today, for the lure of France, in cuisine, wine, high culture, and so on, continues. It is true, however, that the pull since 1700 is less strong, now that there is no particular political connection, and French language “[stands] for the very opposite of down-to-earth vigour that was seen as one of the essential qualities of writing in Scots” (Tulloch 1997, p.392). That is not to say that earlier borrowings are not comfortable in down-to-earth registers of Scots today: consider the following examples of the items *fash* (to bother, annoy, French *fâcher*), and *manky* (to spoil, mutilate, now also dirty, disgusting, French *manquer*, to lack or be wanting), and their derivatives, from SCOTS:

Example 1: Fash

“A trolly like Ben Nevis! / Michty fit a fash! / I'm scunnert an I'm foonert. / Grip, skyte, flash.”²¹

Example 2: Manky

“Wheesht Johnny. Dicht yer face. Yer chikks is manky.”²²

Even these few examples suggest the strongly negative semantic prosody which the words have acquired since being borrowed from French, indeed a new veneer of meaning not yet exemplified in dictionaries.

The SCOTS corpus will be able to provide evidence of the subsequent fate of these and many other foreign borrowings over the centuries. While these are documented in dictionaries, such works cannot provide more than a couple of citations to show how the word has settled into the language, and its current behaviour. A large corpus can show subtleties of usage, and fuller context.

5.2. Current day language contact and research questions

²⁰ See the Dictionary of the Scots Language for definitions: <http://www.dsl.ac.uk/dsl/>

²¹ Example from Sheena Blackhall, ‘The Check Oot Quine’s Lament’, in *The Wizard o the North*.

²² Example from Sheena Blackhall, ‘Overheard at the Chippie: Big brither, wee brither’.

We have seen that SCOTS can provide evidence of the continued usage, and continuing semantic change of items borrowed in the past from other languages. We should also consider briefly the more dynamic question of ongoing language contact involving Scots and Scottish English.

For the last two centuries in Scotland, Scots has been predominantly the language of everyday informal speech, and Scottish Standard English the language of written texts. Does this imply that the influence from Standard English is creeping into Scots through the written language? The corpus will be able to answer such questions regarding reciprocal influence between these two language varieties. An up-to-date body of material will also enable an investigation of the influence of the mass media – mass communication may lead to the loss of non-cognate Scots terms in favour of more Anglicised forms. Alternatively, it might encourage pride in Scots.

It hardly needs stated that the corpus will have applications for lexicography. Consider the word *owersett* (to translate): while attested in SCOTS, and indeed on the Internet, the word is not listed with this sense in the authoritative *Dictionary of the Scots Language*.²³ The headword *overset* in the *Dictionary of the Older Scottish Tongue* notes a number of meanings, including: to press hard; to overpower or overwhelm; to oppress; to overthrow, cast down; but there is no suggestion of the sense of ‘translate (a work from one language into another)’.²⁴ The word with this sense appears to be a Scandinavian or Germanic borrowing.²⁵ The more modern *Scottish National Dictionary* and its supplement do not list *owersett*, or any variant.²⁶ *Translate* alone covers this sense. SCOTS itself testifies to the usage of both, indeed in the same sentence:

Example 3: Owerset

“A Scots *owersettin* o the poem 'Lost' frae the North West American Indian tradition, *translatit* inno Inglis bi David Waggoner” (emphasis my own)²⁷

Example 4: Owerset

“This wad mak it easier for writers tae evyte wrangly *owersettin* shared words intae an orra kinna pseudo Scots.”²⁸

Clearly both words exist in current-day Scots. SCOTS, when larger, will hopefully be able to suggest a subtlety of usage or meaning, which may serve to differentiate them. If this is a very recent acquisition, as the evidence suggests, then it has spread quickly.

²³ See <http://www.dsl.ac.uk/dsl/> for the searchable *Dictionary of the Scots Language*, incorporating both the *Dictionary of the Older Scottish Tongue* (DOST) and the *Scottish National Dictionary* (SND).

²⁴ Compare the entry for *overset* in the Oxford English Dictionary (<http://dictionary.oed.com>).

²⁵ Consider Swedish *översätta*; Norwegian *oversette*; German *übersetzen*, Dutch *overzetten*.

²⁶ ‘Owerset’ does however appear as a headword in the more recent *Scots School Dictionary* (1996), with the sense of ‘translate’ alone, and in the *Concise English-Scots Dictionary* (1993). My thanks to Dr. Maggie Scott, of Scottish Language Dictionaries for her advice on this word.

²⁷ Example from Sheena Blackhall, ‘Tint’ from *The Humpty Dumpty Man*. The sentence above precedes a poem, and clarifies its origin.

²⁸ Exemple from Robert Fairmie, ‘Written Scots’, a non-fiction article on the Scots language.

A final possibility which should not be overlooked is the evidence that SCOTS may be able to give regarding the influence of non-indigenous community languages (Italian, Polish, Chinese, Pakistani).²⁹ In what contexts do these languages come into contact with Scots and Scottish English? Can any lexical or grammatical borrowing, or other influence – in either direction - be discerned?

6. Conclusion

Scotland and her languages offer a wealth of as yet under-exploited possibilities for the study of language contact, and the SCOTS project seeks to facilitate such research. The trace of contact in past centuries is still abundantly, and often surprisingly, clear in the Scots language: as on a palimpsest, layers of Scandinavian languages, Dutch, Gaelic, Latin and French are all visible. Ongoing situations of language contact suggest that in future, similar layers of influence from non-indigenous community languages, like Punjabi, Urdu and Chinese, and the languages of incoming asylum seekers, such as Portuguese, may join them. It is not only the lexical stock of Scots and Scottish English which may bear witness to language contact: the phonological level (accent, pronunciation), grammatical and phraseological levels are also affected. Over time, the SCOTS archive will allow researchers the opportunity to track such changes.

In addition to the possibilities which SCOTS opens up as a self-contained corpus, it is also a snapshot, or set of snapshots, which will enable future comparative work. First and foremost, however, the corpus seeks to provide a record of the present-day linguistic situation in Scotland, including languages under threat. As such, it is a significant resource for future research, in linguistics and sociolinguistics, and cultural studies.

References

- Aitken, A. J. and Tom McArthur (eds.) (1979): *Languages of Scotland*. Edinburgh: Chambers.
- Atkins, Sue, Jeremy Clear, and Nicholas Ostler, (1992): 'Corpus design criteria'. *Literary and Linguistic Computing*, Vol 7, no. 1, pp. 1-16.
- Corbett, John (1997): *Language and Scottish Literature*. Edinburgh: Edinburgh University Press.
- Corbett, John (1999): *Written in the Language of the Scottish Nation: A History of Literary Translation into Scots*. Clevedon: Multilingual Matters.
- Corbett, John, J. Derrick McClure, and Jane Stuart-Smith, (eds.) (2003): *The Edinburgh Companion to Scots*. Edinburgh: Edinburgh University Press.
- Craigie, William A. (1931): *Dictionary of the Older Scottish Tongue*. (Vol. 12, 2002). London: Oxford University Press.
- Dareau, Margaret G. (2001): 'The Scots / Gaelic Interface: A New Perspective.' In John M. Kirk and Dónall P. Ó Baoill (eds.), *Language Links: The Languages of Scotland and Ireland*. Belfast Studies in Language, Culture and Politics 2, pp. 237-254.
- Douglas, Fiona (2003): 'The Scottish Corpus of Texts and Speech: Problems of Corpus Design'. *Literary and Linguistic Computing*, Vol 18, no. 1, pp. 23-37.
- Görlach, Manfred (2002): *A Textual History of Scots*. Heidelberg: Universitätsverlag C. Winter.

²⁹ See for example the writing of Suhayl Saadi, for a literary perspective on Scots-Asian contact.

- Jones, Charles (ed.) (1997): *The Edinburgh History of the Scots Language*. Edinburgh: Edinburgh University Press.
- Leech, Geoffrey (1991): 'The state of the art in corpus linguistics'. In Karin Aijmer and Bengt Altenberg (eds.) (1991): *English Corpus Linguistics: Studies in Honour of Jan Svarvik*. London: Longman.
- Macafee, Caroline I. (1997): 'Older Scots Lexis'. In Jones, C. (ed.) (1997), pp. 182-212.
- Macafee, Caroline I. (2000): 'The Demography of Scots: The Lessons of the Census Campaign.' *Scottish Language*, 19, pp. 1-44.
- Macleod, Iseabail (1993): *Concise English-Scots Dictionary*. Edinburgh: Chambers.
- Macleod, Iseabail and Pauline Cairns (1996): *Scots School Dictionary*. Edinburgh: Chambers.
- McClure, J. Derrick (1986): 'What Scots owes to Gaelic'. In J. Derrick McClure (1995), pp. 68-85.
- McClure, J. Derrick (1988): *Why Scots Matters*. Saltire Pamphlets.
- McClure, J. Derrick (1995): *Scots and its Literature*. Amsterdam/Philadelphia: John Benjamins.
- Meurman-Solin, Anneli (1995): 'A New Tool: The Helsinki Corpus of Older Scots (1450-1700)'. *ICAME Journal*, Vol. 19, pp. 49-62.
- Meurman-Solin, Anneli (2003): 'Corpus-based Study of Older Scots Grammar and Lexis'. In John Corbett, J. Derrick McClure, and Jane Stuart-Smith, pp. 170-196.
- Murison, David (1971): 'The Dutch element in the vocabulary of Scots'. In A. J. Aitken *et al.* (eds.): *Edinburgh Studies in English and Scots*, London: Longman, pp. 159-176.
- Murison, David (1977): *The Guid Scots Tongue*. Edinburgh: James Thin/The Mercat Press.
- Murison, David (1979): 'Norse influence on Scots'. *Lallans*, 13, pp. 31-34.
- Price, Glanville (ed.) (2000): *Languages in Britain and Ireland*. Oxford: Blackwell.
- Smith, Jeremy (2000): 'Scots'. In G. Price (ed.) (2000), pp. 159-170.
- Trotter, D. A. (2000): 'Anglo-Norman'. In G. Price (ed.) (2000), pp. 197-206.
- Tulloch, Graham (1997): 'Lexis'. In C. Jones (ed.) (1997), pp. 378-432.