

Connected Communities

Orkney and Shetland Dialect Corpus Scoping Study

Ragnhild Ljosland



Background

Executive Summary

The main objective of the research is to undertake a scoping study and research review with a view to developing a larger corpus-based project on Orkney and Shetland dialect grammar. The Orkney and Shetland dialects are different from other Scottish dialects for two reasons: (1) The Northern Isles are relict areas where certain phonological and grammatical forms from older stages of Scots still survive and (2) These dialects contain a substratum of the extinct language Norn, which was spoken up until c. 1750 (Barnes 1998). For four hundred years or more, a dialect of Old Norse was in contact with dialects of Scots in the Northern Isles. During this period (15th – 19th century) Scots gained currency as the language of Orkney and Shetland, while the local dialect of Old Norse gradually went out of use. Sometime during this period, local dialects of Scots formed in Orkney and Shetland, incorporating words and structures from the Old Norse dialect. Therefore, research on Orkney and Shetland dialect is of particular interest because it sheds light on the grammar of two dialects which are the products of a language contact situation. Orkney and Shetland dialect grammar is currently an under-researched area.

The research review undertaken as part of the project facilitates further grammatical studies of Orkney and Shetland dialect grammar by bringing together the research which has already been undertaken in the field. The project also lays the foundation for developing a digital corpus of naturally occurring dialect texts from Orkney and Shetland, which is intended to be a future resource for the study of these dialects.

Researchers and Project Partners

Principal investigator

Dr Ragnhild Ljosland

University of the Highlands and Islands

Key words

Language contact
Dialect grammar
Historical linguistics
Orkney
Shetland
Scots
Corpus linguistics
Language shift

The Orkney and Shetland Dialect Corpus Scoping Study

The main objective of the research is to undertake a scoping study and research review with a view to developing a larger corpus-based project on Orkney and Shetland dialect grammar. The Orkney and Shetland dialects are different from other Scottish dialects for two reasons: (1) The Northern Isles are relict areas where certain phonological and grammatical forms from older stages of Scots still survive and (2) These dialects contain a substratum of the extinct language Norn, which was spoken up until c. 1750 (Barnes, 1998). For four hundred years or more, a dialect of Old Norse was in contact with dialects of Scots in the Northern Isles. During this period (15th – 19th century) Scots gained currency as the language of Orkney and Shetland, while the local dialect of Old Norse gradually went out of use. Sometime during this period, local dialects of Scots formed in Orkney and Shetland, incorporating words and structures from the Old Norse dialect. Therefore, research on Orkney and Shetland dialect is of particular interest because it sheds light on the grammar of two dialects which are the products of a language contact situation. Orkney and Shetland dialect grammar is currently an under-researched area.

The research review undertaken as part of the project facilitates further grammatical studies of Orkney and Shetland dialect grammar by bringing together the research which has already been undertaken in the field. The project also lays the foundation for developing a digital corpus of naturally occurring dialect texts from Orkney and Shetland, which is intended to be a future resource for the study of these dialects.

Activities

The scoping study proposed to (1) write a research review, (2) establish contacts with other relevant corpus projects and researchers, (3) plan the development of a digital corpus of Orkney and Shetland dialect texts, (4) hold a workshop with potential future collaborators to discuss draft outcomes of the review, pointing up potential areas of developmental interest, and list possibilities for the next steps of action and (5) identify sources for funding the future digitalisation of the corpus. All these aims have been achieved.

1. A research review was written, covering the areas of Norn research and the history of language shift in Orkney and Shetland, Orkney and Shetland dialect syntax and morphology, Orkney and Shetland dialect phonology, prosody and intonation, and the sociolinguistics of Orkney and Shetland. This review has allowed us to identify and evaluate previous research relevant to the project. The research review has also enabled us to identify specific areas of Orkney and Shetland dialect grammar where there is work to be done. At present, there is much work to be done in many areas, such as in historical linguistics, both regarding the Scots and the Norn elements of the Orkney and Shetland dialects.
2. Contact has been established both with individual researchers and with other relevant corpus projects. The principal investigator, Ragnhild Ljosland, has visited and established contact with the ScanDiaSyn project in Norway, the Scottish Corpus of Texts and Speech in Glasgow, and the Gaelic corpus, also in Glasgow.

An informal network has been established, consisting of researchers and students who have either published on aspects of Orkney and/or Shetland dialect linguistics in the past, are currently working on relevant research projects, or have expressed an interest in undertaking such research in the future. This informal network also has an online discussion forum on Facebook with 34 members: www.facebook.com/#!/groups/107550722702645/

3. Plans for how to develop a digital and searchable text corpus have been laid. The following issues have been considered:

a) Identifying sources of text.

In dialogue with the informal network of researchers which has been established, the scoping study has considered whether to use oral or written dialect texts or both, and identified sources of such texts dialect literature, existing in collections and archives in Orkney and Shetland. An Excel database has been created for each of Shetland and Orkney, listing potential texts with information about title, source location, copyright information, publication year, genre, text length and whether it is fully or partly in dialect.

b) Identifying and considering the need for IT software.

Consultation has been sought with the ScanDiaSyn Nordic corpus project (Janne Bondi Johansen, University of Oslo, and Leif Inge Aa, Norwegian University of Science and Technology), with the Scottish Corpus of Texts and Speech, University of Glasgow (Wendy Anderson and Marc Alexander) and the Gaelic Corpus (Stephen Barrett).

The ScanDiaSyn corpus has developed a variant spelling recognition programme for Scandinavian dialects, which can be trained to understand any dialect and may potentially be

useful for Scottish dialects also. They offered to host the Orkney and Shetland Dialect Corpus in their database if necessary, but would need a separate space to keep the Scottish texts separated from the Scandinavian ones. This will have a cost implication which is unknown at this stage.

Converting from PDF or image formats to text format

For this, there is a programme called OmniPage which takes PDF, jpg and tif. However, the resulting texts need to be corrected by hand, so time for an assistant to do this job should be built into the next grant application.

Tagging and presenting the corpus

The simplest solution is to provide text only. Part-of-speech tagging is labour intensive and time-consuming, but highly desirable. A simple tagging telling simply whether this is Orkney dialect, or Shetland dialect, or neither, is easily feasible.

For simple tagging, the best solution, which would make the corpus compatible with other digitalisation initiatives, is the The Text Encoding Initiative.

The Text Encoding Initiative (TEI) is a consortium which collectively develops and maintains a standard for the representation of texts in digital form. Its chief deliverable is a set of Guidelines which specify encoding methods for machine-readable texts, chiefly in the humanities, social sciences and linguistics. Since 1994, the TEI Guidelines have been widely used by libraries, museums, publishers, and individual scholars to present texts for online research, teaching, and preservation. In addition to the Guidelines themselves, the Consortium provides a variety of supporting resources.

More information available on:
www.tei-c.org/index.xml

For presenting the corpus and making it accessible online, the best solution seems to be the free, open source IMS Open Corpus Workbench (CWB).

The IMS Open Corpus Workbench (CWB) is a collection of open-source tools for managing and querying large text corpora (ranging from 10 million to 2 billion words) with linguistic annotations. Its central component is the flexible and efficient query processor CQP.

This is available from:
<http://cwb.sourceforge.net/>

3. A workshop was held on the 5/10/2012, discussing areas for future research, as pointed up by the research review, and technical requirements of the corpus database. In addition, those members of the informal research network who were not able to participate in the workshop, have been consulted through the online discussion forum. The next step of action will be to submit a grant application to get the corpus up and running, firstly using written texts, but since there was a strong wish in the research network to include spoken texts as well, audio resources should be incorporated in the near future.
4. Two sources of further funding have been considered, and at present the best option seems to be to submit the Dialect Corpus project as part of a joint initiative for the digitalisation of Orkney and Shetland texts. Another group of researchers are currently working towards the digitalisation of mediaeval text material from Orkney and Shetland, and have expressed an interest in including the dialect text corpus as part of a wider online portal.

Academic outputs

1. Journal article (refereed)

The following article has been submitted for publication, and is currently undergoing evaluation:

Ljosland, Ragnhild (under evaluation): "I'll cross dat brig whin I come til him: Grammatical gender in the Orkney and Shetland dialects of Scots." Submitted for an edited volume *Nordic Language Variation: Grammatical, Sociolinguistic, and Infrastructural Perspectives*, edited by Þórhallur Eythórsson, and Øystein Vangnes.

2. Journal article (non-refereed)

Ljosland, Ragnhild (2012): "The Establishment of the Scots Language in Orkney." Published in the *New Orkney Antiquarian Journal* 6, 2012, pp. 65-80.

3. Research review

To be submitted to the AHRC Outputs system.

4. Presentation at Aberdeen University

7/6/2012 "I'll cross dat brig whin I come til him": Grammatical gender in the Orkney and Shetland dialects of Scots."

5. Conference presentation

"The Orkney and Shetland Dialect Corpus Project." Presented at the Nordic Research Network Conference, Edinburgh University, 23-24/2/2012.

6. Conference presentation

"The Orkney and Shetland Dialect Corpus Project". Presented at the University of the Highlands and Islands Research Conference, Inverness, 7-9/11/2012.

Community connection

1. Contact has been established with Shetland Amenity Trust and Shetland ForWirlds, both of which are organisations working to make Shetland’s cultural heritage accessible to the local community and beyond. In Orkney, contact is established with Orkney Heritage Society.

Shetland ForWirlds was given a project presentation on the 6/3/2012, and the project has been much helped by their advice about where to find dialect texts. Orkney Heritage Society has received an academic paper for their journal the New Orkney Antiquarian Journal.

2. The UHI Centre for Nordic Studies already uses a range of channels for knowledge exchange within the local communities on Orkney and Shetland and with interested parties nationally and internationally, including the provision of regular research and communication columns in the Orkney newspaper “The Orcadian” and the Shetland Museum newsletter “Unkans”, as well as maintaining an online presence on Facebook, the Centre’s website and its Posterous blog.

The study enabled expansion and enhancement of these activities. The general public has been able to follow the progress of the project on the Centre for Nordic Studies’ website and Facebook page, in addition to a dedicated Facebook group being set up for researchers interested in the Orkney and Shetland dialects. Also, the general reading public got access to information about the project as well as tasters of some of its research finds in the Centre for Nordic Studies’ column in the local Orkney newspaper

Mimir’s Well. Two articles have appeared in this column in connection with the scoping study: The first appeared on the 27/9/2012: “Masculine and feminine in dialect.” The article is also available online: www.uhi.ac.uk/en/research-enterprise/cultural/centre-for-nordic-studies/blogs/mimirs-well-articles/masculine-and-feminine-in-dialect

The second article appeared on the 6/12/2012: “Freed from the Norn connection.” The article can be downloaded from: www.uhi.ac.uk/en/research-enterprise/cultural/centre-for-nordic-studies/blogs/mimirs-well-articles/freed-from-the-norn-connection

3. The local communities of Orkney and Shetland have benefitted from the project in the form of raised awareness of the local dialects. The dissemination of grammatical research into the dialects contributes to valorising them as linguistic varieties, and helps dialect speakers as well as incomers to the islands develop a view that speaking in dialect is not in any way inferior to speaking in Standard English or other British varieties. This could enhance the quality of life for dialect speakers. More insight into the local dialects will also contribute to the study of the islands’ culture by increasing our understanding of the interplay between the Norse and Scots tongues.

Of direct benefit to the local community in Orkney was a 6-week evening class in Orkney dialect for incomers to the islands, running for 2 hours a week from the 15/2/2012 – 21/3/2012. This was very popular, with 14 students taking part and receiving a certificate at the end.

Also, the Orkney community and visitors benefited from a talk by principal investigator Ragnhild Ljosland at the 2012 Orkney International Science Festival. This talk was part of a day seminar entitled “VOICES

AROUND THE ISLANDS – PAST, PRESENT AND FUTURE” organised by the Orkney International Science Festival in collaboration with the Scapa Flow Landscape Project Partnership.

The local community in Shetland has benefitted from the project being presented in The Shetland Times and on local radio, in addition to contact with the local dialect group Shetland ForWirds and the local archive.

Recommendations for future research

As can be seen from the research review, there are currently several gaps in our knowledge about the Orkney and Shetland dialect, both in a historical and in a contemporary perspective. Having access to a digitally

searchable corpus of dialect texts (written and/or oral with transcriptions) would facilitate further research. Implementation of the text database in digital form is therefore strongly recommended.

Some specific areas where research is lacking, which have been identified in the research review, are:

1. Language shift and historical linguistics

The article “The Establishment of the Scots Language in Orkney” (Ljosland 2012) combines historical accounts with data from the more modern dialect text material which would be included in the digital corpus in order to address the question of when the Scots language became established as one of two indigenous languages in the Orkney community. Using linguistic data is an innovative approach, as

SCAPA FLOW LANDSCAPE PARTNERSHIP SCHEME
with ORKNEY HERITAGE SOCIETY and ORKNEY INTERNATIONAL SCIENCE FESTIVAL
PRESENT

Voices Around the Islands – Past, Present and Future

A day of eight lecture lectures on the Orkney dialect
 "Whar's this a' coonin' fraim?" "She's at the ground" "wath' heaving jerrin' com-
 in' it thro'ing?" "Been no' wath' but jerrin' at' moo..." "wath' 'ow' Buid'!!!"
 "Aag' the' foyner!" "I'll enjoy wath'ing about' I come' to' mair'!" wath' Ragnhild Ljosland!
 (and oob' oob'!)

Saturday 8 September
 9.30am – 3.30pm
 at the St Magnus Centre
 Palace Road, Kirkwall

Speakers
 Harvey Johnston
 Tom Rendall
 Ragnhild Ljosland
 Sarah Jane Gibbon
 Cameron Stout
 Neil Leask

FREE EVENT! Full programme and general booking at:
www.oistf.org
 or telephone 01856 652014 for information.

ORKNEY HERITAGE SOCIETY
 ORKNEY INTERNATIONAL SCIENCE FESTIVAL
 LOTTERY FUNDED

Yaas, min. A'm been here fir
 five thousand year, but the
 wey folk spilk noot Na, min.
 Hid's aa yin ferrylooper!

Orkney dialect for beginners

Have you moved to Orkney and are curious to find out more about the dialect here? Or are you a born-and-bred Orcadian who would like some of your why-questions about the dialect answered? This may be the evening class for you! Through a fun and informal course, you will come out with a better understanding of how the Orkney dialect has developed historically and how it works linguistically. It is taught by the Norwegian sociolinguist Dr Ragnhild Ljosland.

Time: Wednesday nights 7.30–9 pm
 February 15th–March 21st.
Venue: Centre for Nordic Studies, Kiln Corner, Kirkwall. Also available on Video-Conferencing.
Enrolment: Contact Tina Brown at Orkney College, tina.brown@orkneyuhs.ac.uk
 Tel: 01856 569206
More info: 01856 569 302 Centre for Nordic Studies

previous research has mainly been focused on using the historical accounts and comments that are available – see, for example, Millar (2008) and (2010). However, Ljosland (2012) only addresses one grammatical phenomenon: the lack of do-support in questions and negative imperatives, which can be seen as an archaic dialect trait. Other grammatical and phonological characteristics of the Orkney and Shetland dialects which can be deemed archaisms, for example the pronunciation of the consonant clusters *kn* and *gn*, the differentiation of verbal nouns and the present continuous tense of verbs, and the use of overt subjects in non-negative imperatives, should also be examined.

2. Contact linguistics

The contact and mutual influence between Scots and Norn at various historical stages should be investigated further. How long did the Orkney and Shetland dialects of Norn keep in touch with Norwegian dialects, and can we see any influence in the Scots dialects of Orkney and Shetland? A case in point may be the stopping of *ð* and *þ* which happened in Middle Norwegian (c. AD 1400-1450), which also occurs in contemporary Shetland Scots. As regards the Norn substrate in Orkney and Shetland Scots, the most substantial research has been done on the lexicon (Jakobsen 1928-32), (Marwick 1929), but it would be interesting to go through these lists of identified Norn words and see how Scots has influenced their pronunciation – for example, are the front rounded vowels *ø* and *y* distributed along a Norwegian or Scots pattern? There is disagreement in the research literature, see (Millar 2010) versus (Melchers 2010), but without a thorough investigation having been undertaken by either. Another

case in point would be the syllable duration system of modern Shetland Scots. It has been shown to resemble the Norwegian syllable duration system (Leyden 2002), but it remains to be shown that the Shetland system originates in Norn and is evidence for Shetland's contact with late Norse/ Middle Norwegian, subsequently being carried over to Shetland Scots.

3. Morphosyntax and syntax

Claims are often made of what is considered to be particular morphosyntactic of syntactic characteristics of Orkney and Shetland dialects (see, for example Heddle 2010), but a more thorough investigation is needed to evaluate and substantiate such claims. For example, the relative frequency of phrasal verbs is often held out as "Norse", but needs to be compared to Scandinavian and Scots corpora.

4. Sociolinguistic issues

Since the 1970s, many families have moved to Shetland and Orkney, but no work has been done specifically on the speech of children of such families. Smith and Durham (2011) and (2012) have investigated intergenerational dialect death in Shetland. However, it would be interesting to examine to which extent children of incoming families pick up the local dialects of various island communities in Orkney and Shetland, and if they do, which traits are picked up most frequently.

References and external links

- Barnes, M.** (1998). *The Norn Language of Orkney and Shetland*. Lerwick, The Shetland Times.
- Hedde, D.** (2010). The Norse element in the Orkney dialect. *Northern Lights, Northern Words*. R. M. Millar. Aberdeen: 48-57.
- Jakobsen, J.** (1928-32). *An Etymological Dictionary of the Norn Language in Shetland*. London.
- Leyden, K. v.** (2002). "The relationship between vowel and consonant duration in Orkney and Shetland dialects." *Phonetica* 59: 1-19.
- Ljosland, R.** (2012). "The Establishment of the Scots Language in Orkney." *New Orkney Antiquarian Journal* 6: 65-80.
- Marwick, H.** (1929). *The Orkney Norn*. Oxford, Oxford University Press.
- Melchers, G.** (2010). "'This unique dialect': the profile of Shetland dialect in a typology of World Englishes." *Scottish Language* 29.
- Millar, R. M.** (2008). "The origins and development of Shetland dialect in light of dialect contact theories." *English World-Wide* 29: 237-267.
- Millar, R. M.** (2010). "The death of Orkney Norn and the genesis of Orkney Scots." *Scottish Language* 29.
- Smith, J. and M. Durham** (2011). "A tipping point in dialect obsolescence? Change across the generations in Lerwick, Shetland 1." *Journal of Sociolinguistics* 15(2).
- Smith, J. a. D., Mercedes** (2012). "Bidialectalism or dialect death? Explaining generational change in the Shetland Islands, Scotland." *American Speech* 87(1).



www.connectedcommunities.ac.uk