CLARIN Annual Conference 2016
Aix-en-Provence, 26-28 October 2016
Texts, language and geography: Understanding literature using geographical text analysis

Ian Gregory
Modern “Data”
Corpus of Lake District writing

- 80 texts from 1622 to 1900
- 1.5 million words
- Includes a wide variety of texts:
  - Daniel Defoe
  - Celia Fiennes
  - Thomas Gray
  - Thomas West
  - Robert Southey
  - Samuel Taylor Coleridge
  - William Wordsworth
  - Harriet Martineau
On Sunday Augt. 1st - half after 12 I had a Shirt, cravat, 2 pair of Stockings, a little paper & half a dozen Pens, a German Book (Voss's Poems) & a little Tea & Sugar, with my Night Cap, packed up in my natty green oil-skin, neatly squared, and put into my net Knapsack / and the Knap-sack on my back & the Besom stick in my hand, which for want of a better, and in spite of Mrs C. & Mary, who both raised their voices against it, especially as I left the Besom scattered on the Kitchen Floor, off I sallied - over the Bridge Greta Bridge, Keswick, thro' the Hop-Field, thro' the Prospect Bridge at Portinscale, so on by the tall Birch that grows out of the center of the huge Oak, along into Newlands -- Newlands is indeed a lovely Place-the houses...
<table>
<thead>
<tr>
<th>Code</th>
<th>Name</th>
<th>OS Date</th>
<th>Date</th>
<th>Location</th>
<th>Longitude</th>
<th>Latitude</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>111</td>
<td>Great Gable</td>
<td>5/8/02</td>
<td>Thursday Aug 5</td>
<td></td>
<td>510500</td>
<td>321500</td>
<td></td>
</tr>
</tbody>
</table>
William Wilberforce

“Journey to the Lake District from Cambridge”

Wednesday.

Sept. 1st. Walked out in the Morning to Castle, The

It is a fine old Building spoiled by some more modern additions... it has one small Turret remaining which is discernible at a great distance.

It has been in many hands.

It formerly belonged to the Stanleys; it is now in the Possession of a great Grandson of Col. Chartres.

From the Top you see the Course of the River, which must have looked better a few years ago before so much Wood was cut down by its side.

It is a small onscape: joins the Castle, The

A short course. I believe it is only a Beck.

The Castle, The

are the remains of greater things, but it never appears from the part which is tolerably entire to have been a place of strength except from its situation.

To the North are some very extensive moors, but in general a pretty cultivated country.

The country in general very pretty.

9 miles.

The Country near the Road between Castle, The

Nelson's

open a View the finest of the kind I ever beheld though I saw it to great disadvantage in an indifferent day; I had not the prospect station.

An exceedingly rich cultivated Valley in which Villages... Gentlemen's

A view at

An exceedingly rich cultivated Valley in which Villages... Gentlemen's

£10 to

a Pied Hill

Mines near the Road between Castle, The

Hornby

Hornby

which are carried in little carts drawn by one Horse.

About 3 miles from Castle, The

Castle, The

Castle, The

Castle, The

Castle, The

The Eye (Castle, The)
39,172 coordinate-based place names (2.57% of total word tokens)
37,564 (95.9%) in the U.K.
34,530 (88.1%) in northern England and southern Scotland
23,459 (59.9%) in the modern National Park
Landscape Appreciation: Majestic, sublime and beautiful

• S.T. Coleridge (1772-1834) met a ‘lady and a gentleman’ when visiting Cora Lynn, the highest of the Falls of Clyde

• According to Wordsworth’s *Recollections of a Tour Made in Scotland, A.D. 1803* (1874, p. 37):
  – The gentleman, ‘observed that it was a majestic waterfall’.
  – Coleridge was pleased ‘he had been settling in his own mind the precise meaning of the words grand, majestic, sublime, etc.’, as he agreed
  – The lady declared that it is was ‘sublime and beautiful’ as well.
  – Coleridge, a little disheartened but not a little amused, turned away feeling ‘not very desirous to continue the conversation’.
Beautiful

- Total instances: 1,233
- Placename collocates in LD: 486
- Collocates per instance: 39.4%
Majestic

- Total instances: 173
- Placename collocates in LD: 79
- Collocates per instance: 45.7%

Kulldorf’s Spatial Scan Statistic
Sublime

- Total instances: 270
- Placename collocates in LD: 87
- Collocates per instance: 32.2%

Kulldorf’s Spatial Scan Statistic
Comparing clusters

Only 1% clusters included
## Heights - % of PNCs

<table>
<thead>
<tr>
<th></th>
<th>&lt;300m</th>
<th>300-600</th>
<th>&gt;600m</th>
</tr>
</thead>
<tbody>
<tr>
<td>Whole corpus</td>
<td>78.1</td>
<td>13.9</td>
<td>8.0</td>
</tr>
<tr>
<td>Beautiful</td>
<td>83.3</td>
<td>11.5</td>
<td>5.1</td>
</tr>
<tr>
<td>Majestic</td>
<td>50.6</td>
<td>13.9</td>
<td>35.4</td>
</tr>
<tr>
<td>Sublime</td>
<td>63.2</td>
<td>24.1</td>
<td>12.6</td>
</tr>
</tbody>
</table>

- Majestic very much associated with high places (20/28 are Skiddaw)
- Sublime also tend to be high but not to the same extent as Majestic – more diverse than majestic
- Beautiful tend to be low
Word use by Period

% of total instances
- Majestic is mainly used in the Romantic period
- Sublime is mainly used in both the Long C18th and the Romantic
- Beautiful largely follows the overall word count in the corpus

% of expected
Early Modern Excluded due to small nos.
Place-Centred Reading: Majestic

• Skiddaw
  – Mainly a term used in Guidebooks
  – Used in a wide range of ways and times

• Blencathra
  – only really appears because of proximity to Skiddaw
    • Eg. “majestic Skiddaw, with Blencathra, or Saddleback on its right”
      Mackay (1846a) or “the majestic Skiddaw and Blencathra” Anon (1852b)

• Duddon
  – 5 of the 7 texts are Victorian
  – Wordsworth (1820a) *The River Duddon: A series of Sonnets “*now expands majestic Duddon, over smooth flat sands”
  – All of the subsequent Victorian texts use this quote
Place-Centred Reading: Sublime

• Western Fells
  – Many different place-names – Only Sca Fell and Pillar used more than once
    • Waugh (1861b): “that sublime group of mountains, Yewbarrow, Pillar, Great Gable, Kirkfell, Lingmell, and Scawfell…”
    • Baines (1829a): “…on the west the mountains of Red Pike, High Stile, and High Crag rise precipitously to a sublime elevation.”

• Other ranges
  – Mackay (1846a) in two instances describes the Old Man and the Coniston fells as sublime. In one he describes “the sublime chain extending from Coniston to Ennerdale”
  – This phrase is copied by subsequent Victorians eg. Black (1853a) and Baddeley (1900a)

• Lakes
  – Cumberland (1776c) Odes says “…our own domestic lakes of Ulswater [sic], Keswick and Windermere exhibit scenes in so sublime a stile[sic]…”
    • Later guidebooks copy this, eg. Robinson (1819b), Cooke (1827), Baines (1829a), Leigh (1830a). Some also describe other lakes as sublime.
Place-Centred Reading: Beautiful

• South Lakes:
  – Driven mainly by Waugh (1860a) *Over the Sands to the Lakes* and copies in his (1861b) *Rambles in the Lake Country*

• Other clusters:
  – Wide variety of uses, authors and genres
Having left Rosthwaite in Borrowdale, on a bright morning in the first week of October, we ascended from Seatwaite to the top of the ridge, called Ash Course, and thence beheld three distinct views. On one side, the continuous Vale of Borrowdale, Keswick, and Bassenthwaite;—with Skiddaw, Helvellyn, Saddleback, and numerous other mountains;—and, in the distance, the Solway Frith and the Mountains of Scotland. On the other side, and below us, the Langdale Pikes— their own vale below them;—Windermere;—and, far beyond Windermere, Ingleborough in Yorkshire. But how shall I speak of the deliciousness of the third prospect! At this time, that was most favoured by sunshine and shade. The green Vale of Esk—deep and green, with its glittering serpent stream, was below us; and, on we looked to the Mountains near the Sea;—Black Comb pre-eminent;—and, still beyond, to the Sea itself in dazzling brightness. Turning round we saw the Mountains of Wastdale in tumult; to our right, Great Gavel, the loftiest, a distinct and huge form, though the middle of the mountain was, to our eyes, as its base.

We had attained the object of this journey; but our ambition now mounted higher. We saw the summit of Scaw Fell, apparently very near to us; and we shaped our course towards it; but, discovering that it could not be reached without first making a considerable descent, we resolved, instead, to aim at another point of the same mountain, called the Pikes, which I have since found has been estimated as higher than the summit bearing the name of Scawfell Head, where the Stone Man is built.

The sun had never once been overshadowed by a cloud during the whole of our progress from the centre of Borrowdale;—on the summit of the Pike, which we gained after much toil though without difficulty, there was not a breath of air to stir even the papers containing our refreshment, as they lay spread out upon a rock. The stillness seemed to be of this world;—we paused, and kept silence to listen; and no sound could be heard: the Scawfell Cataracts were voiceless to us; and there was not an insect to hum in the air. The vales which we had seen from Ash Course lay yet in view and, side by side with Eskdale, we now saw the sister Vale of Duddondale terminated by the Duddon Sands. But the majesty of the mountains below, and close to us, is not to be conceived. We now beheld the whole mass of Great Gavel from its base,—the Den at Wastdale at our feet,—a gulphe immeasurable; Grasmere and the other mountains of Crummock,—Ennerdale and its mountains; and the Sea beyond! Gladly would...
Conclusions

• Contrasting patterns:
  – Majestic is used for a small number of specific features
  – Sublime tends to be used for mountain ranges and lakes in a more general way
  – Beautiful is a more general word

• We can use texts to ask:
  – ‘what places are being talked about?’
  – ‘what places are associated with this word/theme?’
  – ‘what is being said about this place?’
  – Combine corpora with other data

• Need to balance large-scale quantitative with detail and nuance

• Further work
  – Early English Books Online
  – BL’s C19th Newspaper Collection
    • 30-60 billion words

wp.lancs.ac.uk/dighum