
Digital *Hansard*: Politics and the Uncivil

Marc Alexander

marc.alexander@glasgow.ac.uk
University of Glasgow

Andrew Struan

andrew.struan@glasgow.ac.uk
University of Glasgow

Summary

This short paper uses the recently-completed Hansard Corpus to show the patterning of attitudes expressed by the British Parliament about things considered to be ‘uncivilized’ across the last two centuries. It starts from the lexical resource of the Historical Thesaurus of English to gain an overview of the lexicalisation of the concept ‘uncivilized’ and uses this digital data demonstrates a substantial shift (from foreign to domestic) in who Parliament considers to be uncivil.

Introduction

The ways in which the British have discussed ‘uncivilized’ peoples which travellers have encountered throughout the history of English gives a key insight into how people in the past have identified and classified the world around them. This paper uses data from the *Hansard Corpus 1803-2003* (Alexander and Davies, 2015-) alongside the *Historical Thesaurus of English* (Kay et al, 2015-) to analyse the evolution of how the English-speaking people have thought of those who they think uncivil in five different sense-families — as animals, as ill-formed people, as strange-speaking outsiders, as savages, and finally as innocents awaiting enlightenment. Only these large digital data sources can show us the patterning of who and what the British Parliament have considered to be barbarous across time.

Data

This analysis became possible following the completion of the *Historical Thesaurus of English* (HT) in 2009 and the semantically-tagged *Hansard Corpus 1803-2005* in 2015, both of which are currently directed by Alexander and were created by teams of scholars at the University of Glasgow.

The HT is a database of all the recorded words in the history of English arranged according to their

meaning; one of the world’s oldest digital humanities projects, and in progress for over 50 years, the HT database (stored on media from punch cards to tape to diskettes to networked storage to the Web) allows us an unparalleled resource for analysing the history of English. The *Hansard Corpus 1803-2005*, completed in 2015, is a digital corpus of speeches in the British Parliament between those dates, consisting of 1.6bn words across 7.6m speeches. Its contents were semantically tagged in the 2014-15 SAMUELS project (The SAMUELS Consortium, 2015) with disambiguated meaning codes from the HT, making it possible to search for semantic categories rather than words, as we do below.

The Uncivil

The category of *Civilization* in the HT gives us an indication of a non-typical pattern in the number of words available to describe a given concept (in English, categories normally grow throughout time) in the words referring to *uncivilized* and a *lack of civilization*, as Figure 1 shows.

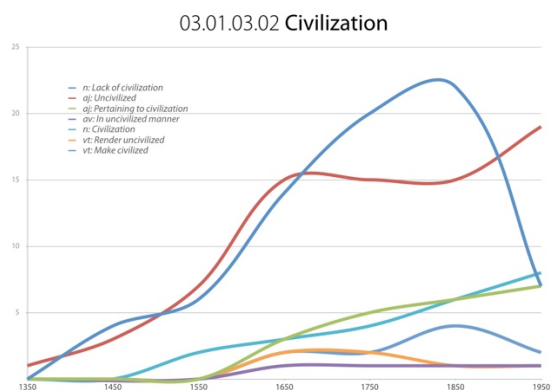


Figure 1: The size of each subcategory of *Civilization* in the HT

While the size of the *uncivilized* adjective category rises in the latter 20th century, there is a substantial fall at the same time in the size of the *lack of civilization* noun category, which we argue is connected to the shift in who has been considered to be uncivil (see below). In addition, of the [42 words](#) in the *uncivilized* category in the HT (see Figure 2), the vast majority follow a particular path of lexicalization which we describe below, with new terms reflecting the shifting conceptualization of the uncivil throughout the times at which they were coined.

03.01.03.02.01 (adj.) *Uncivilized*
 bærbære OE · elreord OE · elreordig OE · hæþen OE · ungerad OE · wild<wilde OE; 1300- · wildern 1300 · fremd 1374 · bestial 1400-1816 · savage 1420/30- · savagine 1430-1430/40 · rude 1483- · barbaric 1490-1513; 1837- · barbar 1535-1726 · barbarous 1538- · Scythical 1559-1602 · barbarious 1570-1762 · raw 1577; 1847; 1865 · incivil 1586 · barbarian 1591- · uncivilized 1607- · negerous 1609 · savaged 1611 · mountainous 1613-1703 · ruvid 1632(2) · ruvidous 1632 · incivilized 1647 · inhumane 1680 · tramontane 1739-1832 · semi-barbarous 1798- · irreclaimed 1814 · semi-savage 1833- · semi-ferine 1854-1858 (rare) · warrigal 1855 (Austral.)-1890 · sloven 1856 (US)-1882/3 · semi-barbaric 1864 · wild and woolly 1884- · woolly 1891- · jungle 1908- · medieval 1917- (colloq.) · jungli 1920- · pre-civilised 1953- **01** and *unsubdued* unatemed OE **02** *without intelligible language* ungeroord OE **03** *specifically of persons* uncivil 1553-1644 · savage 1588- **04** *pertaining to uncivilized people* savage 1614- **05** *acting/speaking as uncivilized* barbarizing 1662; 1855- **06** *rendering uncivilized* barbarized 1602; 1839 · barbarizing 1809- · decivilized 1831-1892 · barbarianized 1885 · decivilizing 1889 **07** *becoming uncivilized* barbarianizing 1859 **08** *absence of accepted social standards/values* anomic 1950-

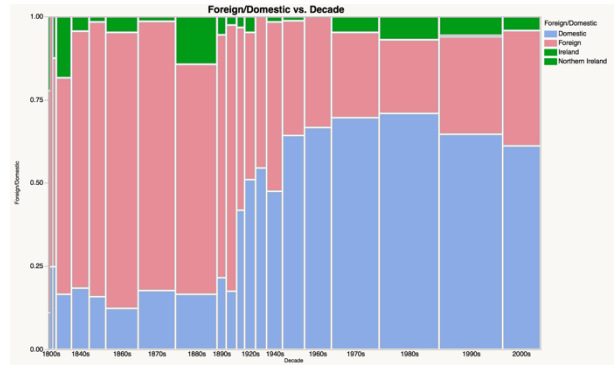


Figure 3: The proportion of uses of *uncivil* words to refer to foreign or domestic persons (thickness of bars reflects the changing amounts of text in *Hansard* in those decades); note that the status of Ireland and Northern Ireland is contested with regards to the foreign/domestic status, and so has been represented separately here

Figure 2: ‘Uncivilized’ in the HT, taken from p.1235 of the print edition.

Thus far this sort of analysis has been slow-paced and difficult to undertake. However, with the tagging in the *Hansard Corpus 1803-2005* we can investigate this sort of semantic and conceptual change in a much more rapid fashion by honing in on uses of these meanings in context across time.

Parliament

There are five families of meaning into which the words above can be categorised, as outlined above. In a past article (Alexander and Struan, 2013), we assembled some evidence for this from the history of English in a non-systematic fashion. For this short paper, we instead account for all the evidence from the *Hansard Corpus* — over 2,000 uses of the semantic category — in order to trace across recorded Parliamentary history the shifts in the cultural, political and social attitudes towards the ‘uncivilized’. This shows a substantial change in the picture which differs from the simpler five-family view of the sense evolution of *uncivil* we described in that earlier article.

Our first change to discuss is the shift, shown below, from the uncivil primarily being foreigners in the 1800s to being domestic persons in the 1900s onwards.

This is reflected in the changing discourse surrounding *barbaric* and *uncivil* things, where a majority of 20th century uses refer to barbaric practices and actions rather than persons:

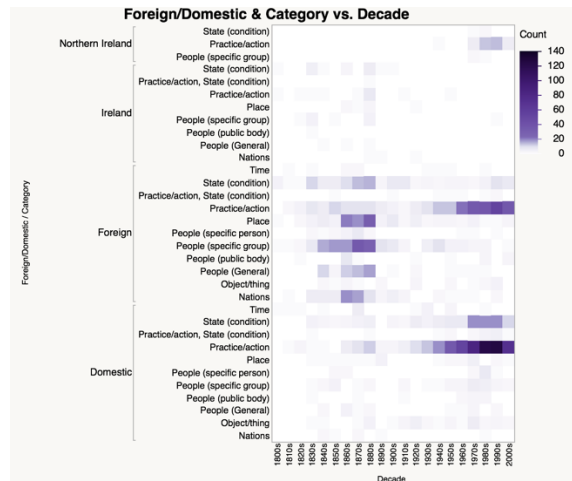


Figure 4: A heatmap of the entities (people, states, practices) considered uncivil by Parliament in the data, separated by whether the entities are foreign or domestic

Through four other graphs, we further report on the distribution of uncivil references across the globe and between the two Houses of Parliament. We also show the changes in the five evolutionary sense-families we outline above, which is key to the foreign/domestic shift we describe.

Some quotes from the corpus can briefly illustrate these changes, which here are aimed at a general body of persons, or a country:

Mr Charles Adderley, House of Commons 21 February 1865: *'...to discharge what Lord Grey described as the singular office of dispensing rude laws among uncivilized tribes.'*

Earl of Carnarvon, House of Lords 12 May 1874, on India:

'But a central government is not enough. In barbarous times and in uncivilized countries, roads are the first condition of improvement; and here it will be our first duty to open and secure the maintenance of roads and trade-paths.'

Mr Richard Cherry (Attorney-General for Ireland), House of Commons 20 March 1908: *'I never said that the people of Ireland were West African savages.'*

Lord Hylton, House of Lords 18 April 1995: *'We can now see that in dealing with Russia we are dealing with a semi-barbarous state and a society that only knew a measure of democracy for a few years before the First World War.'*

Mr Andrew Robathan, House of Commons 1 November 2001, on the pending invasion of Iraq: *'We should not allow a barbaric, mediaeval [sic] regime to succeed or last. We certainly do not want to go back to civil war.'*

As a result, we can show empirically the shift over two centuries in the ways which Members of Parliament described uncivil or barbaric entities, from foreign people or places to domestic practices. We conclude by arguing that this is the result of increased oppositionality being shown in the digital Parliamentary record, and so in this short paper we combine 'big picture' graphs of large-scale data analysis with more focused examples from the corpus record.

Bibliography

Adamson, S., Allan, K., Andrade, S., Arac, J., Davis, J., Durrant, A., Durkin, P., Heath, S., MacCabe, C., Mehl, S., Robertson, K., and Yanacek, H. (2016). The Keywords Project. University of Pittsburgh. (<http://keywords.pitt.edu/index.html>).

Alexander, M. and Davies, M. (2015–) *Hansard Corpus 1803-2005*. Available online at <http://www.hansard-corpus.org>.

Alexander, M. and Struan, A. 2013. 'In countries so unciviliz'd as those?': The Language of Incivility and the British Experience of the World. In Martin Farr & Xavier Guégan (eds.) *Experiencing Imperialism: The British Abroad since the Eighteenth Century*, volume 2. London: Palgrave Macmillan.

Kay, C., Roberts, J., Samuels, M., Wotherspoon, I., and Alexander, M. (eds.). 2015–. *The Historical Thesaurus of English*, version 4.2. Glasgow: University of Glasgow. <http://www.glasgow.ac.uk/thesaurus>.

The SAMUELS Consortium (2015). *The SAMUELS Project*. United Kingdom AHRC and ESRC. <http://www.glasgow.ac.uk/samuels>.

Williams, R. (2014). *Keywords: A vocabulary of culture and society*. Oxford University Press.