The Conceptual Foundations of the Modern Idea of Government in the British Eighteenth Century: A Distributional Concept Analysis

Peter de Bolla, Ewan Jones, Paul Nulty, Gabriel Recchia, John Regan

HCM 7: 619–652

DOI: 10.18352/hcm.575

Abstract

This essay sets out a new method for the history of ideas. Using a mixed approach combining computer assisted reading methods with more traditional close reading, the essay tracks the evolution of a set of terms over the eighteenth century that have become central to how we think about government in particular and political concepts in general. The essay is offered as an example of how data mining very large digital archives allows us to see trends and patterns that are invisible at the granular level of human scale reading, and it proposes that these large-scale observations can both complement and complicate our hitherto analogue histories of ideas. The findings of this mixed approach indicate that ‘despotism’ functioned as a type of gate in an electronic circuit, sometimes allowing the connection to liberty and government and on others blocking those connections. Most significantly ‘despotism’ is shown to be an essential ingredient in the conceptual foundations of a theory of rights, liberty and government in the period and that this structure underpins contemporary theories of government.

Keywords: conceptual history, despotism, distant reading, government, liberty, rights, text mining

Introduction

Over many years, the traditions of scholarship aimed at tracking the foundations of modernity in its political dimension have directed our attention to a tight cluster of words in English which seem to have particular weight or resonance. These terms effectively circumscribe the ‘theory of government’ as it emerged out of the long Anglophone eighteenth century, and include: democracy, aristocracy, monarchy, anarchy, republic, liberty, freedom, despotism and tyranny. This list is not intended to be exhaustive but it does comprise what we might think of as the fundamentals through which actors in the period understood existing political forms, or sought to create new ones. We can think of these words as some of, perhaps even the key basic elements of a language or discourse of political modernity as it has come to be formulated in the substantial scholarship since Quentin Skinner set out the foundations of modern political thought.¹

Within this tradition there is no paucity of close reading of the canonical texts. One familiar form of the history of ideas builds a genealogy of ‘great thinkers’ who hand down ideas to each other, each putting their own complexion on what came before. Ideas move from a thinker, or perhaps one ‘great text’, to another in a chronological chain of transmission: Montesquieu, Harrington, Hobbes, Locke, Rousseau and so on. But if we loosen the grip of the ‘great texts’ or ‘great thinkers’ and begin to trace the emergence and persistence of ideas through other lenses, can we track transmission across the entire archive? Can we begin to use the new digital tools at our disposal for the construction of data-driven accounts of the evolution of those political ideas that underpin and continue to inflect our own contemporary ideas of politics and government?

The objective of this essay is to demonstrate one way of doing this. It uses techniques borrowed from computational linguistics alongside more bespoke statistical methods for interrogating the history of ideas, allowing us to see in a different optic the ways in which the underlying political concepts formulated in the Anglophone eighteenth century coagulate into clusters and networks – what we call ‘bundles’ of conceptual forms – that provide the building blocks for the long extension of ‘modern political thought’, up through the nineteenth and twentieth centuries into our current time.² Since the methodology is entirely new

we shall spend some time exploring its salient features, but we begin with a set of observations on the understanding of government in the Anglophone eighteenth century.

The Modern Idea of Government

Within the Anglophone world of the Enlightenment the political consensus arrived at in 1688 has turned out to be very stable and durable: to this day in Britain we live in a parliamentary monarchy. Over time the lines of force in this conjunction have been subject to scrutiny and attenuation with the first term, parliament, slowly asserting its priority and unimpeachable authority. Unlike the United States, however, as is repeatedly noted, Britain lacks a document, a written constitution, to which it might refer when parliamentary or governmental authority comes under stress with regard to its legitimacy: all we have is custom and usage. This in itself might make us pause and direct our close attention to *words*. We shall begin with a brief account of some of these words in their historical constellations, but we wish to indicate that our main purpose is to go beyond mere lexis and to uncover a region of activity of behaviour that we think of as more properly conceptual. We shall have more to say about this below.

The Anglophone Enlightenment's slow cooking of the political theory of parliamentary monarchy entailed pretty consistent attention to the nature of a 'mixed regime' of government that brought into conjunction (with lines of force and attraction that our digital methodology clearly uncovers) the *words* monarchy, aristocracy and democracy. Not only words, of course, since their referents are agents in the distributions of power, privilege and politics. We contend that one of the legitimating supports for such action or agency is the conceptual conjuncture that was the 'idea' of government, the sign or *techne* of legitimate rule for the time. As a way of getting a quick grip on this 'mixed regime' we might note that Edmund Burke spent his entire career explaining and promoting its benefit whereby the Crown, Lords and the Commons were jointly and collectively understood to be sovereign. Each could serve as a check on the others with the elected lower house providing one leg of the stool in the form of democracy, the hereditary aristocracy as another, and the crown as the monarchical element.

Other writers and thinkers took a different view on some aspects of this triangulation of powers but our point here is not to follow the intricate debates between individuals or even across texts. Rather, we want to observe that this triangulation of words in the first instance (monarchy, aristocracy and democracy) in fact turns out to have a very distinctive signature (which we shall argue is conceptual through and through) in the patterns of distribution of lexis within the culture of the Anglophone eighteenth century. Moreover, this apparently commonplace association of the words supervenes upon a deeper conceptual architecture which becomes clear when we subject the massive dataset of Eighteenth Century Collections Online (ECCO) to computational scrutiny. This will be remarked upon below with respect to our data mining of that archive. Here we make the initial observation that ‘democracy-aristocracy-monarchy’ comprises a conceptual ‘bundle’ that has very deep foundations in the Anglophone tradition of political thought from the Enlightenment on.

Our second brief excursus into the uses and histories of words notes this: the word ‘despotism’ is, until around the decade of the 1740s, almost completely absent from the English language. Although the first use of the word in English was in a handbook for instruction in writing letters penned by the Rev Thomas Cooke in 1708, in fact the word appears only one hundred and eighty-nine times in all English printed text up until 1750.³ During the last decade of the century it appears over fourteen thousand times. This fact is remarkable given the load the word bears with respect to the geometries of thinking undertaken by the Anglophone Enlightenment around the nature of government. Although other European traditions of political thought had long pondered the distinction between ‘tyranny’ and ‘despotism’, English, it appears, found little use for the word before the 1750s.⁴

We should note, however, that once the explosive force of the French revolution against a venal and corrupt aristocracy and a negligent and self-satisfied monarchy occurred, the British were all over the word ‘despotism’, applying it not only to the French regime that had been tumbled (a monarchical despotism that Edmund Burke felt the division of powers mitigated) but also to the ‘despotism of the people’ that ensued. This observation is clearly supported by the fact of the word’s frequency, but we shall have a rather more fine-grained analysis to offer which asks us to consider the conceptual knot within which the term operated.

Our third introductory set of remarks seeks to bring to the fore a recurring feature of eighteenth-century conceptual mapping which tests the limits of an idea's internal coherence: how can we account in structured ways for the complex relationship between a given word (say 'despotism') and its apparent antonyms ('liberty' and 'equality', for instance) and synonyms ('tyranny')? At what point can we conclude that a concept fragments, begins to lose coherence, or extends so far as to devolve into its opposite, as when John Brooks, a newly made American, claims that 'the extremes of liberty border on despotism'?⁵ We contend that one way of answering this question (which applies not only to political discourse but also to conceptual usage in general) is again to isolate patterns of lexical co-association, through computational analysis of large datasets; accordingly, the following sets out to put such methodology into tension with more established forms of intellectual history.

Data Mining as a Means for Establishing Conceptual Architecture

In this section, we introduce some of the techniques and methodology developed by The Cambridge Concept Lab for establishing the outlines or shapes of conceptual forms – what we call their 'architecture'. Computational and statistical techniques are commonly employed in research into natural language processing where projects are directed at tasks such as machine translation, analogy solving, question-answering and natural language inference systems, and considerable success in these domains has been achieved by researchers using distributional semantic models derived from large text corpora.⁶ Furthermore, some work in digital humanities is beginning to investigate the utility of vector semantics or 'vector space models' for understanding concepts.⁷ Our own modelling of concepts follows a similar direction and the tools the Cambridge Concept Lab has created allow one to transparently operate arithmetical procedures (addition, subtraction, multiplication and division) with the vectors we derive from the co-association of individual lexical items.⁸

Irrespective of theoretical motivations, the computational implementations of these methods have much in common. Word, phrase, or

document meanings are approximated by deciding on a word-distance window within which to count word co-occurrences or compare the paradigmatic context, and counts of word or context co-occurrence are tabulated into a vector. The vectors can be compared directly to measure word associations, or combined into a matrix to measure the similarity of documents. Such an approach diverges somewhat from topic modelling, a document-based method that also discovers groups of associated words by modelling a process in which documents are generated by selecting from subsets of words with varying probabilities.⁹ The resulting clusters of associated words are intended to capture topics in the corpus, and to allow comparison of documents by topic and measurement of diachronic change in topic emphasis. This technique has been widely applied outside of computer science, for example to measure political attention and literary style.¹⁰

Applications that focus on measuring word rather than document association often use paradigmatic similarity of contexts shorter than document length.¹¹ The meaning of a word in such models is represented as a point embedded in a high-dimensional space defined by the contexts in which its occurrences are counted. The context may be a short window of words around the target word or defined by a grammatical relation discovered by automatic syntactic parsing of the text.¹² Since the number of possible contexts may be many times the size of the vocabulary, dimension-reduction methods such as singular value decomposition are often used to reduce the complexity of the model, with the drawback that the resulting dimensions are no longer directly interpretable. Word-context models with dimensionality reduction result in a ‘word vector’, usually of the order of a few hundred dimensions – these are commonly referred to as ‘word embedding’ models. An efficient and widely used implementation of this kind of model is the Word2Vec package, which uses a neural network trained to predict word contexts, and encodes the meaning of the word in its parameters rather than explicitly counting word contexts.¹³

Our primary measure is ‘lexical co-association’: our code searches through massive historical datasets of language use (here the over two hundred thousand volumes in ECCO) in order to first identify effectively the co-associations between each and every term in the corpus.¹⁴ This produces over 594 million data points. Our code then uses a statistical measure to ascertain the likelihood of one term co-associating with

another over a notional baseline of a random distribution of the terms within the dataset. This allows us to construct a metric, what we call a ‘distributional probability factor’ or dpf, which gives us an indication of the strength of probable co-association between every term.¹⁵

There are numerous well-established approaches in computational linguistics for identifying statistical associations among words on the basis of their patterns of co-occurrence in text corpora.¹⁶ Pointwise mutual information (PMI) is frequently used either on its own as a measure of lexical association, or as a starting point for the construction of more complex measures. PMI is given as

$$I(x, y) = \log_2 \frac{P(x, y)}{P(x)P(y)}$$

where $P(x)$ and $P(y)$ can each be approximated as the frequency of x and y (respectively) divided by the total number of words in the corpus, and $P(x, y)$ as the number of times that x and y co-occur divided by the total number of words in the corpus. $P(x, y)/P(x)P(y)$ has an intuitive interpretation: it simply expresses the ratio between the number of times x and y co-occur, divided by the number of times one would expect them to appear if their appearances throughout the corpus were randomly distributed. What qualifies as a ‘co-occurrence’ is up to the user of the measure and the most popular choice is appearance in the same window or document.

Although most practitioners define co-occurrence in terms of appearance in the same window or document, this is more a matter of convention than necessity. Window-based measures can be readily adapted to ‘distance-based’ measures: one can specify that in order to count as a co-association with x , a word y must appear some distance d words away from x , plus or minus a word or two. This is the approach we have taken because we wish to capture data on patterns of lexical distribution that moves from close up, where semantic or syntactic coherence is strongest, to far away, where it is weakest. By taking this wider view we mean to discover relations of binding that go beyond or underpin strictly local semantic ties. Thus, we use what could be called a ‘sliding window’, as it involves a window of a fixed size that, rather than being centred on x itself, is centred on a word d words away.

To measure lexical co-association, we have adopted a simple variant of PMI, which drops the logarithm and introduces a smoothing exponent

in the denominator in order to compensate for the tendency of PMI to give higher scores to infrequent terms. We are primarily interested in the rank order of words as they are co-associated with a focal word x . Since the logarithm does not affect this rank ordering it can be dropped without loss of generality. Doing so highlights the fact that our measure is essentially a small modification to the measure sometimes referred to simply as ‘observed over expected’: the number of times two words are observed in conjunction, divided by the number of times one would expect to see them together by chance. This measure, which we refer to as the distributional probability factor (dpf) since it expresses the extent to which a word is predicted to be distributed across the dataset, is as follows:

$$dpf(x, y) = \frac{P(x, y)}{(P(x)P(y))^\alpha}$$

When we calculated the value of α that eliminated the inverse correlation between rank and frequency described above, we found the optimal value for our corpus to be 0.78, very close to the value of 0.75 reported by previous work into smoothing for PMI scores.¹⁷ Higher dpf values corresponded to a negative correlation between rank and frequency, while lower values corresponded to a positive correlation. In other words, with an appropriate value of alpha, dpf is sensitive to data sparsity as well as frequency, and retains the simplicity and transparency of PMI. Subsequent references in this work to the dpf between two words therefore refer to formula (2) with an alpha of 0.78.

We consider dpf to capture the strength of binding between terms since it gives us a data-dependent analysis of the probability of one term co-associating with another, and our computational tool constructs rank ordered lists of co-associated terms according to the measure of dpf (a ‘binding list’) for any word in the dataset. This strength of binding can also be calculated across lexical distance: hence the likelihood of, say, liberty co-associating with licentiousness at a proximity of five words will have one dpf (a value of 35), and at another distance, say seventy words away, a different dpf (a value of 8). Since our dataset is chronological, we can also plot the binding profile for every word in the corpus diachronically.

The Language of Government

Let us take an example. The ‘language of government’ in the Anglophone eighteenth century can initially be understood for the purposes of our

argument as a network of terms, each of which have distinct semantic values. Such networks often contain antonyms, for example, because we frequently make distinctions between senses through opposition or negation (say freedom and slavery). Equally we may find synonyms or near-synonyms in a network whose semantic purpose might be to create fine distinctions between broadly similar values, adding granularity to a larger semantic set (here one might think of the similarities and differences between liberty and freedom).

But the ties between words in a network are also determined by larger units of sense-making, call these the topics which enable us to rapidly move across and within a cognitive environment (this is one way of describing what we call thinking). An example here would be the binding between 'liberty' and 'dissenters' that we find at distance 70, in the year spread 1710–1730. There is no strictly speaking semantic connection between these two words, but there is a topic or 'idea' connection: dissenters in the period were constrained as to their participation in the political and social life of early eighteenth-century Britain. Hence talk of dissenters was likely to run into the issue of liberty. When mapping these larger networks of bound terms we speak of the 'company' a term keeps: its proclivities with respect to the number of terms it is predicted to co-associate with, or the preservation of a common set of strongly bound terms over time. This allows us to track how ideas are made up of smaller scale sub-networks we call 'bundles' which are connected to the overall network of considerable complexity which determines the 'language of government'. One of the tools we have constructed, outlined in the following section, enables us to visualize the complex connections between bundles in a larger network and to identify the multiple pathways at different levels of resolution. As we shall see, we can also track these pathways diachronically, thus enabling us to see how ideas change over time.

We first turn to some of the data our custom-designed algorithm produces. The following example presents the result of our searches through the second half of the ECCO corpus (1751–1800) for the top twenty-one words most tightly bound at a distance of one hundred words away from the word 'government' (Fig. 1).¹⁸

In this chart we can see that our metric of dpf ranks the term 'government' as most tightly bound with itself.¹⁹ The terms which follow comprise the twenty next most highly co-associated words with

Figure 1. Dataset: ECCO. Top twenty-one terms on ‘government’ dpf list, 1751–1800, D:100.

government at this distance.²⁰ Getting a sense of the strength of binding between a target word, here ‘government’, and its bound pairs is a good place to begin if we want to assess the coherence and stability of a network over time. But we can also develop analyses of the different structures we find in these binding lists and compare them. One such structure we call ‘mutual association’ whereby the terms very highly bound to one term also feature very close to the top of another term’s binding list.

Terms held in a mutual association binding structure are rarely abstractions, but there is one significant outlier to this general observation: across the eighteenth century we find the triad ‘democracy/aristocracy/monarchy’ in a mutual association structure of binding.²¹ Here are (Table 1) the top six bound terms for each:

Table 1. Dataset: ECCO. Top six bound terms, 1751–1800, D:100.

Democracy	Aristocracy	Monarchy
Democracy	Aristocracy	Monarchy
Aristocracy	Democracy	Monarchical
Monarchical	Impeded	Democracy
Patricians	Monarchy	Monarchies
Republics	Republics	Aristocracy
Monarchy	Monarchical	Republican

In effect this data is telling us that in the period when using the term ‘democracy’ there was an underlying structure of binding that tied the word to ‘aristocracy’ and ‘monarchy’. And the same observation applies to the other terms: that is its mutual binding. Let us say, then, that when thinking about the nature of government in this period one would have been very likely to bump into one of the terms in this mutual association structure.

Of course, there are other gateways or apertures through which one might ‘think government’ (as our first chart indicates: one such term was ‘executive’), but if we return to Fig. 1 we can see that the terms most co-associated with ‘government’ itself include our triad of ‘democracy/aristocracy/monarchy’. Although this surely would come as little surprise to intellectual historians of the period, what is noteworthy is the underlying structure or architecture: it is not that one would have been likely to summon up ‘democracy’ when thinking with or about the idea of government, but that summoning this term inevitably brings into view the other two. And this holds true for both of the other terms. Since each of these terms (as government itself) also has a much longer list of relatively strongly bound other terms the mapping or intermeshing of all these binding lists quickly becomes extremely complex. One way of getting a view onto that complexity is to set a threshold for the number of ranked terms on a binding list and then to compare the lexis on each list so as to construct a common set. If we do that for the top sixteen terms on each binding list in our triad, once again with the date slice being the second half of the century, we arrive at this common set:

Democracy, aristocracy, monarchical, monarchy, republics, monarchies, republican, government, governments, despotic, republic, hereditary, despotism, anarchy, tyranny, constitution.

One can immediately see a pattern emerging in the co-association data presented so far: the lexis we find in the orbit of the idea of government in the second half of the eighteenth century has considerable stability and coherence and the mutuality of the binding across the three terms we have identified as a triad persists over a considerable stretch of time.²² Our aim, however, is to dig deeper into the data so as to more carefully inspect the complexities of these lexical-conceptual behaviours. We now turn to some of the techniques we have established for analyzing this complexity through the use of a visualization method.

Visualization of Networks as an Analytic Method

The following figures are derived from the same calculations we have made to discern the dpf of terms in our dataset (ECCO). In this case, however, we have exported the data into a network environment which helps us to identify complex patterns of similarity and connection between data points, in this case represented as words in our dataset. The figures are based on screenshots from these network representations displayed in an interactive web app. These are ‘neighbourhood’ or ‘ego’ graphs of order two; that is, they show the graph containing nodes within at most two steps from a specific focal node or nodes entered by the user. An edge exists between two nodes if their dpf association is above a user-specified value. It is also possible to threshold the edge list by association rank.

The network is drawn using the *threejs* R package, with a force-directed layout algorithm which models the network mechanically as repelling particles connected by spring.²³ The result is that in a graph of suitable density and degree, nodes are spaced apart enough to be distinguished, but the edges pull together nodes into clusters that share many relations. Community structure is detected using unsupervised modularity optimization and indicated by colour.²⁴ For visual clarity, the graph is displayed after removing nodes that have only a single connection.

In this first network plot we can see the lexical environment in which government operated in the 1730s (Fig. 2).

Here one can see that the term ‘government’ is connected to six other terms, based upon our code’s prediction of its co-association profile: presbyterian, presbyterians, commonwealth, monarchy, arbitrary and governments. It is to be expected that these two different ways of slicing into the co-association behaviour of the term ‘government’ across the period (the first generating the metric of dpf and listing bound words in descending order of strength of binding, and the second using a higher dimensional space of a network to visualize nodes and edges which are connected) agree to a very great extent. We should expect to find the same lexis because the underlying data is held in common. But the presentational matrix is also a diagnostic tool: we can identify patterns or behaviours by dint of the different ways in which we can inspect the information. And this helps us isolate creases or folds in the data which drive further investigation. An example of precisely this aspect

Figure 2. Dataset ECCO. Search term: government, 1720–30, $D:100$.

Figure 3. Dataset ECCO. Search terms: government, despotism, 1750–1760, $D:100$.

of the digital method can be seen in the following network plot, which explores the networks in which both ‘government’ and ‘despotism’ are embedded (Fig. 3).

This plot operates the same thresholds as our previous plots and indicates that the tightly bound company that ‘government’ keeps in the decade of the 1750s does not include ‘despotism’, the term isolated on

Figure 4. Dataset ECCO. Search term: government, despotism, 1760–1770, $D:100$.

its own at the right edge of the plot (in purple/pink). Moreover ‘despotism’ has no other terms for company. This accords with the fact that the word in English at this time is very infrequent (as we shall see below), but the algorithm is not simply driven by word frequency: the plot is telling us something about the relations or connection of these terms, and much more importantly, it is identifying the conceptual supports or architectures for the bundle of terms we are constructing as a representation of the ‘idea’ of government at the time. We can compare this to the next decade, the 1760s where we can note that ‘despotism’ now has a direct connection to the government sub-network (Fig. 4).

And if we inspect the final decade (Fig. 5) we find that the overall network much more closely accords, as it should, with the rank dpf list above (Fig. 1). Now we can see that ‘despotism’ (the sub-network in red) has its own quite substantial company including the words tyranny, anarchy and revolution, terms all present in Fig. 1, and this sub-network has both a direct connection to government and its sub-network and indirect links (via one additional node) to that same network through republican, monarchy, hereditary, anarchy and democracy. The terms in this denser network include those in its immediate vicinity: constitution, legislation, subversion, executive, burthens, discontents, Englishmen, abdication, political and magistracy. This sub-network is in turn bound to a further cluster that includes the words governments, monarchical,

Figure 5. Dataset ECCO. Search term: government, 1790–1800, $D:100$.

nation, polity, people, aristocracy, heredity, nation and democracy. The third sub-network, meanwhile, includes republic, republics, republican, monarchies, despot and despotism.²⁵

Our visualization tool has enabled us to see the steadily increasing bundle or sub-network within which ‘despotism’ appears as the century moves into its final decades. Its presence in the overall network re-calibrates or adjusts the relations between its constituent parts. One of those adjustments bears heavily upon the conceptualization of government: the centrality of ‘despotism’ as a hinge or bridge for liberty and the corresponding marginalization of ‘tyranny’ as structural element in a theory of democratic polity. In the following section we present another computational technique for digging deeper into this observation.

High Dimensional Conceptual Analytics

In the following we briefly consider what it means to represent a word as a point in a high-dimensional space. To express the location of a

point in a three-dimensional space requires three numbers, commonly denoted x , y and z . Although we cannot visualize mathematical spaces exceeding three dimensions, it is nevertheless the case that a list of 1000 numbers can be conceived of as a point in a 1000-dimensional space, and we can calculate the ‘distance’ between these points using mathematics analogous to that used to calculate the distance between two points in the physical world. If 1000 numbers are generated for two words in such a way that these numbers will be similar if the words appear in similar lexical contexts, then we can quantify how ‘close’ or ‘far’ their lexical contexts are away from each other. The standard Pearson correlation frequently used in statistics can be regarded as a measure of the ‘closeness’ between the lexical contexts of two words,²⁶ and it is in this way that we use the term ‘correlation’ in the following graphs. Specifically, we are computing Pearson correlations between lexical vectors constructed from the dpf lists, which include an element containing the dpf between the focal term and every other term in the lexicon. For example, the 22nd, 23rd, and 24th elements of the vector for ‘government’ contain the values for dpf(government, abbess), dpf(government, abbey), and dpf(government, abbot), respectively; likewise, the 22nd, 23rd, and 24th elements of the vector for ‘tyranny’ contain the values for dpf(tyranny, abbess), dpf(tyranny, abbey), and dpf(tyranny, abbot). The first bar in Fig. 8 represents the correlation between the lexical vectors for ‘tyranny’ and ‘government’.

As we have seen, the ‘language of government’, that which supports the idea of democratic polity across the eighteenth century, includes a cluster or bundle of sixteen words. We can assess the extent to which this language swerves towards one term, ‘despotism’, or the other, ‘tyranny’, over time by calculating the correlations in high dimensional space between the search term and this list of sixteen bound terms. Thus, in the early part of the century when the word ‘despotism’ is very infrequent, we can see that there is very little correlation (Fig. 6).

What this chart presents is the fact that the lexical contexts within which the word ‘despotism’ circulated in the date slice were uncorrelated with the lexical contexts for each of the sixteen words in the chart. We can compare this to the same high dimensional analysis of ‘tyranny’ in the same date spread (Fig. 7).

Here we can see that our bundle of sixteen words – what we are characterizing as the core of the idea of government in the period – shares

Figure 6. Dataset: ECCO. Pearson correlations between the full dpf list of 'despotism' and the full dpf lists of selected terms, 1720–1740.

Figure 7. Dataset: ECCO. Pearson correlations between the full dpf list of 'tyranny' and the full dpf lists of selected terms, 1720–1740.

a lexical-conceptual terrain with 'tyranny' to a far greater extent than with 'despotism'. As a way of confirming this correlation we have made the same high dimensional analysis through the search term 'government' (Fig. 8).

Figure 8. Dataset: ECCO. Pearson correlations between the full dpf list of 'government' and the full dpf lists of selected terms, 1720–1740.

Here one can see that in the period 1720–1740 the 'language of government' and its supporting conceptual bundles hardly includes the lexical-conceptual entity 'despotism'; it correlates less than the mean for every term in the dataset. By the end of the century this had been transformed (Fig. 9).

In this graph we have compared the correlations across the two date slices, 1750–70 and 1780–1800, in order to map the increasing presence of 'despotism' within the constellation of terms that supported the idea of government. Not only was 'despotism' becoming more and more central to a theory of government, it also provided a switching station, either positively or negatively charged, through which the progressive political aspiration of 'liberty' entered the network of concepts underlying that theory.

Despotism and the Modern Theory of Government

As we noted in our introduction the word 'despotism' hardly occurs in English before the 1740s. While it is the case that the word 'despot'

Figure 9. Dataset: ECCO. Differences in correlations between the full dpf list of 'government' and the full dpf lists of selected terms, 1780–1800 vs. 1750–1770. Words with higher values are more associated with 'government' from 1780–1800 than from 1750–1770.

does occur, as does 'despotic', the following graph of the relative frequencies of occurrence of all three words across the eighteenth century (Fig. 10) helps us sharpen up a diachronic account of the foundation of the conceptualization of the idea of government.

This rising use of the term 'despotism' is all the more striking, given that the English language tends, in the eighteenth century, to restrict abstractions ending in the suffix 'ism' to tightly specified contexts.²⁷ Here is a graph (Fig. 11) mapping the five most frequent 'ism' words in ECCO that indicate adherence to a particular system or ideology, which so happen to be the only five such words with a frequency in our corpus of greater than 10,000.²⁸ These terms are followed in frequency by scepticism, deism, Judaism, heathenism, libertinism, polytheism, polytheism, Arianism, Calvinism, theism, Methodism, republicanism, stoicism, Hebraism and Puritanism, which demonstrates that 'despotism' is one of the few such terms that does not indicate a religious sect.

We can clearly see just from these relative frequencies that 'atheism' for most of the century was the most common 'ism', but also that by the end of the century 'despotism' far outstripped all other contenders.

Figure 10. Dataset ECCO. Relative frequency of 'despot', 'despotic' and 'despotism' 1700–1800.

Figure 11. Dataset: ECCO. Relative frequency of the five most frequent "isms" indicating adherence to a system or ideology in the ECCO corpus, 1701–1800.

These statistical trends support what we already know about the politics of the period: the clear spike in 1793–5 can be read as a symptom of the British counter-revolutionary campaign, its 'war on terror' on home-grown republican tendencies. Yet the simple rise to prominence of the

individual word ‘despotism’, taken by itself, poses broader questions when we take into account the larger conceptual bundle that we established above. Did this new entrant re-configure the discourse on government, or was it simply drawn into that constellation’s ambit? How did it both open up and close down travel between the bundles of concepts or ideas that comprised the larger network? Our contention is that it significantly re-mapped the terrain we have been exploring, and in order to support this claim, we now want to toggle between the large computational patterns investigated above and more localized close readings that are commonly the evidence base for traditional histories of ideas.

Intellectual historians have provided extensive and persuasive accounts of concepts such as tyranny and despotism, over historical periods far broader than that which this essay surveys. Mario Turchetti, for example, has surveyed the development and subsequent eclipse of ‘tyranny’ from early antiquity, arguing for the term’s enduring relevance in our post-totalitarian age.²⁹ Melvin Richter dates the modern resurgence of ‘despotism’ to Montesquieu’s *De l’Esprit des lois* (1748), which prefers the term to the then-current ‘tyranny’.³⁰ The distinction between the two synonyms is, Richter contends, comparatively clear-cut: Montesquieu reserves ‘tyranny’ for republics, where ‘despotism’ indicates nation-states; ‘tyranny’ designates an individual ruler who may be replaced (hence the cognate ‘tyrannicide’), while despotism entails a broader notion of government. Yet the latter half of the eighteenth century, he continues, sees this distinction evaporate. Montesquieu, despite being a pragmatic reformist, unwittingly bequeaths his term to revolutionaries and counter-revolutionaries of all colours and stripes:

By the end of the 18th century, *despotisme* had been used in so many ways, positively and negatively, that it was at once omnipresent and without any single distinctive meaning. Although some theorists continued to distinguish it from *tyrannie*, as we shall find Condorcet doing in his treatment of human rights, these two terms in less specialized general usage ceased to be differentiated.³¹

It is worth pausing over such claims, both for substantive and methodological reasons. In order to substantiate this argument (that ‘despotism’ travelled from comparative conceptual clarity to increasing fuzziness), we would need to show, firstly, that tyranny and despotism were in fact ‘differentiated’ even in ‘less specialized general usage’ over the earlier portion of the eighteenth century; and secondly,

that the subsequent confusion of the two terms precluded any relevant conceptual disarticulation. Furthermore, Richter has an even more eye-catching thesis, that political terms such these undergo a process of ‘whiting out’, a phrase that Richter borrows from the founder of *Begriffsgeschichte*, Reinhart Koselleck, which results in a loss of specificity in meaning, effectively emerging as an ungovernable polysemy.³²

Such global claims about diachronic change are not easily supported by human scale reading: Richter, outstanding intellectual historian as he is, works within the traditions and prevailing methodology of the discipline. He closely reads the ‘master texts’ such as *De l’Esprit des lois*. Now, however, we are able to read at ‘inhuman’ scale, using computational means for tracking lexical behaviour across massive datasets. We believe that such computational methodologies can help to place such claims upon a more secure footing, all the more so when they are used as a supplement to, rather than a replacement of, the ‘close’ reading of texts. When, indeed, we apply the procedures that we have outlined above, we see clear evidence that the two synonyms concerned (‘despotism’ and ‘tyranny’) do continue (or perhaps even begin) to operate in definably different contexts. These contexts are both shaped by, and in turn contribute to, the broader idea of government that we traced with the mutual association set above.

The Architecture of Government

One of the advantages of the computational tools that we have been developing is their provision for a rapid toggling between quantitative patterns of data and the individual texts of which they are comprised. When we switch optics in this fashion, we find that the quantitative trends above reflect a series of structured ways in which language-users employed the idea of government in general, and the emergent notion of despotism in particular. To be sure, there are several instances in which, just as Richter claims, the two terms were used interchangeably, as when we read that ‘all monarchy degenerates into despotism as the soft modern term is or into tyranny according to the expression of the middle ages’. Yet despite this tendency, and despite what any reader of the archive has to acknowledge as the multiplicity of (sometimes mutually exclusive) linguistic contexts in which the term emerges, ‘despotism’ reveals a number of clear structural trends.

All the examples presented below explore representative contexts in which ‘despotism’ co-associates with one or more of the terms we have established as operating within the conceptual network of government in the period, and in order to dampen a focus on the texts themselves we have removed their citational traces to the footnotes. When these thousands of linguistic contexts are inspected, decontextualized from the broader argument or theme of the works at hand, three clear structural trends emerge. Despotism, unlike its synonym tyranny, is frequently employed so as to describe i), a system of governmental balance; ii), a theory of governmental development; iii), a counter-intuitive affinity between itself and liberty. These tendencies are distinct, albeit that they are often also mutually supportive. We will treat each in turn.

Governmental balance. Despotism, unlike tyranny, is frequently drawn into a broader architecture of government. The logic often runs something like this: ‘despotism’ is to ‘government’ as ‘licentiousness’ is to ‘liberty’. The optimal (or least bad) political model, on this view, lies in the golden middle of these two extremes. Freedom is therefore not an absolute good. Countless examples could be cited. Here are but a few: ‘privilege and right which constitutes the essence of a free government distinguished on one hand from despotism and on the other from too great licentiousness and anarchy’; ‘anarchy my lords is not liberty no more than despotism is government’; ‘ill calculated for so large an empire just bursting from the shackles of extreme despotism and liable to fall into the other extreme of licentiousness and anarchy’; ‘it is true that no government can subsist in the midst of licentiousness. But, *licentiousness* and *despotism* are only different names for the same thing’; ‘natural corruption of liberty is licentiousness; and the necessary correction thereof opens the way to despotism’; ‘liberty, subject to no controul, is licentiousness, producing anarchy and confusion, and frequently ending in downright despotism’.³³

At times, this striking general trend directly permits the disarticulation of despotism from tyranny, as when the desire to remove a political evil (i.e. through tyrannicide), involves those mutineers in a travesty of liberty as despotism: ‘the people deceived by the charms and delusive attractions of an apparent liberty inadvertently plunge into the most horrid excesses and finish their violent pursuits by establishing a most hateful despotism planned by the very persons who began the tragedy by proclaiming themselves the avengers of

tyranny'.³⁴ The upshot of such examples is that despotism was understood to be a form of government that took its place next to other, competing forms such as democracy, rather than a type of tyranny in which the arbitrary rule of an individual was imposed upon the polis. It is for this reason that we also find in abundance the identification and suspicion of despotism as a lurking or concealed presence: 'the Americans judge from facts they have seen an uniform lurking spirit of despotism pervade every administration'; 'the secret favourers of despotism lay in concealment and a government unconnected with the cabinet a constitutional parliament'.

The association of despotism with government therefore clearly poses a potential problem for reformist and revolutionary discourses alike: there is no silver bullet to eradicate despotism as there is for tyranny: tyrannicide. Yet it also offers practical resources: if a theory of government invariably produces the tendency towards despotism, those instruments can be appropriated even by those who believe in 'liberty'. David Hume, for example, viewed political societies as necessary composites: 'we are to consider the ROMAN government under the emperors as a mixture of despotism and liberty, where the despotism prevailed; and the ENGLISH government as a mixture of the same kind, where the liberty predominates'.³⁵ More positive formulations of this general idea can also be found: 'they may endeavour to ally the peace of despotism with the delights of liberty'.³⁶

Governmental process. We have already seen that Montesquieu's conceptual disarticulation persists with remarkable force, even into a later historical period and a separate language tradition. The English rendering of *De L'Esprit des Lois* resonates with the efforts of eighteenth-century British political theorists to articulate the exceptionalism of the peculiarly effective reconciliation of competing powers that had emerged since 1688: 'Democracy therefore has two excesses to avoid, the spirit of inequality which leads to aristocracy or monarchy; and the spirit of extreme equality, which leads to despotic power, as the latter is completed by conquest.'³⁷ The trick they had discovered was to conceptualize despotism as a form of government both in a structural sense, and also as a dynamic process.

On this common view, a natural process emerges whereby, say, despotism inevitably leads to democracy; once again, this distinguishes

uses of ‘despotism’ from ‘tyranny’ since the latter was commonly thought of as a state in which one might find oneself with little chance of being extracted from it. Its temporality was effectively steady state. This should not be understood to indicate that the processual aspect of despotism always led in the same direction, in fact the processes set in motion could be mutually exclusive. They could for example, support optimistic or even utopian agendas: ‘liberty has often sprung from despotism’.³⁸ So too they could be mobilized for more conservative projects: ‘England should *dread* that attachment to liberty which produces such excesses, and consider that if not checked they will soon rise to anarchy and possibly end in despotism’; ‘this very Cromwell originally sowed all the seeds of his future despotism in the fertile but disguised hot-bed of *Equality*’.³⁹ These and many other examples confirm Richter’s claim that despotism might be used in a ‘bewildering’ variety of often-contradictory contexts, but they also betray a recognizable mode of thinking in which government is conceptualized with respect to its evolution or change over time: a necessarily unfixed or emergent form of social and political organization, a process.

Affinity of despotism and liberty. The above examples already provide evidence for the ways in which despotism and liberty intersect in complex ways. This might involve simple contradiction, affinity-through-opposition (licentiousness is to liberty as despotism is to government), or process (liberty grows out of despotism, or vice-versa). Taken together, such common formulations pave the way for a more extreme conceptual articulation: that liberty or equality is (or can be) despotism (or vice versa). When the newly-minted American John Brooks claims, for example, that ‘the extremes of liberty border on despotism’, he is not suggesting that one extreme is like another, or produces its opposite; rather, the two terms ‘liberty’ and ‘despotism’ share a substantive essence.⁴⁰ As above, it is important to state that such formulations are politically agnostic: they can be recruited to serve conservative or revolutionary agendas. We find countless examples of the former: ‘such governments according to Polybius terminate in despotism not from the abuse of royalty or aristocracy but from the licence of democracy it seems that the people can no longer be intrusted safely with the exercise of power’.⁴¹

Yet this counter-intuitive affinity between liberty and despotism can also serve very different ends. Richter reminds us of Robespierre's famous declaration that 'the government of the revolution is the despotism of liberty against tyranny', which he takes to indicate the 'whiting out' of meaning.⁴² By polemically treating despotism as a political good, Richter claims, Robespierre unwittingly contributes to its general fuzziness. As we have seen above, the phrase certainly is used in a variety of ways, which encompass the structural and the dynamic, the positive and the pejorative. Yet in light of that broader cultural sampling, it is striking just how much Robespierre retains a common form of thinking: 'the despotism of liberty' is expressly governmental, and contrasts explicitly with an ill ('tyranny') that can be directly opposed. When Robespierre equated these two concepts, it was in part because the conceptual fusion was culturally available to him.

This is not to claim that Robespierre directly read Hume's reflections on the compatibility of despotism and liberty, nor to corroborate the sometimes stronger claims made for the cohabitation of both concepts in the more marginal English-language texts that we have surveyed. Our argument rests upon the basis that such linguistic utterances proceed not simply through the desire of a Montesquieu or a Robespierre to alter a given concept decisively – which is to enthrone individuals as the drivers of conceptual change, while often relegating the whole cultural archive to a merely subsidiary or even distorting role. Rather, we insist on the far wider embedding of lexical use in the culture at large which provides us with information on conceptual change over time. By the end of the eighteenth century, we contend, despotism was not simply one satellite in the fixed constellation of the terms used for describing and understanding government, it had become indispensable within that constellation. Thus, if we toggle back from the local level to the larger computational patterns that we established above we can see this very clearly (Fig. 12).

It is important to note that our algorithm centres the search term in a plot – hence nothing should be immediately read from the fact that 'despotism' in Fig. 12 is the largest represented node. However, it is significant that 'despotism' draws both liberty and freedom into the network with government. As can be seen from the following network plot when we open the network through the search term 'liberty', we find no connection to 'government' (Fig. 13).

Figure 12. Dataset ECCO. Search term: *despotism*, 1790–1800, $D:100$.

Figure 13. Dataset ECCO. Search term: *liberty*, 1790–1800, $D:100$.

Conclusion

These computationally derived network plots indicate that ‘despotism’ is a hinge or bridge between regions in the larger constellation of concepts that inform the idea of government at the end of the eighteenth century, and we believe that its placement as such gives us clear evidence for the basis

upon which theories of the state, polis or government were constructed. As we noted above, 'despotism' operates as switching station, sometimes operating as the aperture through which 'liberty' enters the larger network and sometimes negatively, providing the counter-weight to it, an antagonist blocking the pathway. And, as Fig. 13 indicates, the liberty network pulls with it the term that will become unavoidable in our own contemporary attempts to think equality, freedom, government: rights. Here is evidence for the observation that the inter-connection of rights and liberty in the era of universal rights, symbolized by the Universal Declaration of Human Rights (1948), has its roots in the later eighteenth century admixture of government, rights and despotism. It may be worth recalling this in our contemporary efforts to establish universal human rights which have been predominantly exercised by efforts to secure rights *in* or *to* specific things, such as water or holidays. Such efforts have had mixed success, partly on account of the fact that there is no universal governmental framework within which the conjunction of rights and liberty might be upheld. Furthermore, the quasi universal institution that has been established, the United Nations, has been seen by some polities as imposing its views and values without due regard for the differences that obtain in the various socio-political entities across the world.

We might conclude from this exemplary exploration of digital methods that large-scale computer assisted reading of the archive can deepen and strengthen more long-standing traditions of enquiry into the history of ideas. 'Ideas' and the concepts which support and construct them are both distributed across a culture at large and at the same time used by, and on occasion even invented by, historical actors. Combining both close and distant reading techniques allows us to inspect the finer grain of the archive and to track larger scale movements both across texts or regions in the archive and over time. Further applications and construction of digital tools is likely to enhance this mixed approach even further and tools such as those employed in this article are intended to enable scholars to explore and refine these techniques. It feels as if the field is on the brink of a new chapter in its disciplinary journey.

Notes

- I See *inter alia* Quentin Skinner, *The Foundations of Modern Political Thought* (2 vols; Cambridge, 1978); J.G.A. Pocock, *The Machiavellian*

- Moment* (Princeton NJ, 1975); J.G. A. Pocock, *Virtue, Commerce and History: Essays on Political Thought and History, Chiefly in the Eighteenth Century* (Cambridge, 1985); J.G.A. Pocock, *Political Thought and History* (Cambridge, 2009); John Dunn, *The Political Thought of John Locke: An Historical Account of the Argument of the 'Two Treatises of Government'* (Cambridge, 1969); John Dunn, *Political Obligation in its Historical Context: Essay on Political Theory* (Cambridge, 1980).
- 2 The debate over this long tradition has been particularly intense around Skinner's life-long project. See in particular Quentin Skinner, *Liberty before Liberalism* (Cambridge, 1998); Annabel Brett and James Tully, with Holly Hamilton-Bleakley, *Rethinking the Foundations of Modern Political Thought* (Cambridge, 2006); and Andreas Kalyvas and Ira Katznelson, *Liberal Beginnings: Making a Republic for the Moderns* (Cambridge, 2008).
 - 3 This observation is based on the ECCO dataset which does not capture all lexical use across the Anglophone eighteenth century, so to this extent the claim is subject to qualification. For the first use see Thomas Cooke, *The Universal Letter-Writer; or, New Art of Polite Correspondence* (London, 1708) 123.
 - 4 Mario Turchetti, 'Despotism' and 'Tyranny': Unmasking a Tenacious Confusion', *European Journal of Political Theory* 7:2, esp. 167–76.
 - 5 John Brooks, *An Oration to the Society of Cincinnati in the Commonwealth of Massachusetts, July 4, 1787* (Boston, 1787) 11.
 - 6 C. Kelly, B. Devereux and A. Korhonen, 'Acquiring human-like feature-based conceptual representations from corpora', *Proceedings of the NAACL HLT 2010 First Workshop on Computational Neurolinguistics*, Association for Computational Linguistics (2010) 61–9; and F. Pereira, M. Botvinick, and G. Detre, 'Using Wikipedia to learn semantic feature representations of concrete concepts in neuroimaging experiments', *Artificial Intelligence* 194 (2013) 240–52.
 - 7 See Michael Gavin, 'Intellectual History and the Computational Turn', *The Eighteenth Century* 58:2 (Summer 2017) 249–53.
 - 8 The tools are available at <https://concept-lab.lib.cam.ac.uk/>.
 - 9 See D.M. Blei, A.Y. Ng and M. I. Jordan, 'Latent dirichlet allocation', *Journal of Machine Learning Research* 3 (Jan. 2003) 993–1022.
 - 10 K.M. Quinn, B.L. Monroe, M. Colaresi, M.H. Crespin and D. R. Radev, 'How to analyze political attention with minimal assumptions and costs', *American Journal of Political Science* 54:1 (2001) 209–28; and J.M. Hughes et al, 'Quantitative patterns of stylistic influence in the evolution of

- literature', *Proceedings of the National Academy of Sciences* 109:2 (2012) 7682–6.
- 11 W. Weaver, 'Translation', *Machine Translation of Languages* 14 (1975) 15–23. See also W. Lowe, 'Towards a theory of semantic space', *Proceedings of the Twenty-Third Annual Conference of the Cognitive Science Society* (Mahwah, NJ, 2001) 576–81.
 - 12 K. Lund and C. Burgess, 'Producing high-dimensional semantic spaces from lexical co-occurrence', *Behavior Research Methods, Instruments, & Computers* 28:2 (1996) 203–8; D. Lin, 'Automatic retrieval and clustering of similar words', *Proceedings of the 17th international conference on Computational Linguistics* 2, (1998) 768–74; O. Levy and Y. Goldberg, 'Dependency-Based Word Embeddings', *Association for Computational Linguistics* 2 (2014) 302–8.
 - 13 T. Mikolov et al, 'Efficient estimation of word representations in vector space', *arXiv* (2013) (preprint arXiv:1301.3781); D.R.G.H.R. Williams and G.E. Hinto, 'Learning representations by back-propagating errors', *Nature* 323 (1986), 533–6. And for a more digital humanities focus, Ben Schmidt, 'Vector Space Models for the Digital Humanities' (<http://bookworm.benschmidt.org/posts/2015-10-25-Word-Embeddings.html>; accessed 21 June 2019).
 - 14 For practicality we have excluded terms with a frequency of less than 13,567 in the corpus as a whole. This preserves over 95% of the linguistic tokens in the ECCO corpus while keeping calculations manageable. This cut-off was chosen by ranking terms by frequency and retaining the top 25,000 terms.
 - 15 We use the term 'co-association' to emphasize that we would like to measure a reciprocal (or symmetric) association: the association score for (A,B) ought to be the same as for (B,A) in type and degree.
 - 16 Note that we do not lemmatize the corpus in order to retain important distinctions in abstraction often indicated by different word forms, for example despot vs despotism, or nerve vs nerves. Punctuation is removed, and the text is converted to lower case and tokenized by whitespace.
 - 17 O. Levy, Y. Goldberg, and I. Dagan, 'Improving distributional similarity with lessons learned from word embeddings', *Transactions of the Association for Computational Linguistics* 3 (2015) 211–25, at 215.
 - 18 Note that these ranked lists of bound words can be very long, so the issue is where to set a threshold so as to screen out noise but nevertheless identify significant information. In general, the curves that result when a word's rank is plotted against its dpf are well-characterized by a power function.

One reasonable way to set such a threshold is by fitting a power function to this curve, calculating the point at which the curve's derivative is -1 (e.g., the 'bend in the curve'), and considering all words with higher ranks than this point (in the same way that statisticians consider factors preceding the 'elbow' of the curve in factor analysis). Calculating the threshold in this manner for 'government' (1751–1800) results in 271 terms. For visualization purposes, we display only the very top of this list here.

- 19 This is due to the well-characterized phenomenon referred to in computational linguistics as 'burstiness', whereby if a word does appear once, it is much more likely to appear again, i.e. words appear in bursts. See R.E. Madsen, D. Kauchak and C. Elkan, 'Modelling word burstiness using the Dirichlet distribution', *Proceedings of the 22nd International Conference on Machine Learning ACM* (2005) 545–52, at 546.; K. W. Church and W.A. Gale, 'Poisson mixtures', *Natural Language Engineering* 1 (1995) 163–90.
- 20 The list at a distance of ten words apart contains only one different word, 'tyranny'. This provides clear evidence for the very tight lexical-conceptual network within which the language of government was articulated in the period.
- 21 We have developed data on the entirety of ECCO to ascertain how common this kind of binding structure is and the extent to which it occurs more with concrete terms than abstract. The data is clear: there are 2,852 distinct cliques of three concrete words in which mutual binding occurs as opposed to 1,547 abstract.
- 22 The same analysis for the first half of the century comprises the following terms: republics, republic, consuls, plebeians, senators, domestic, senate, tribunes, tribune, national.
- 23 See Thomas M.J. Fruchterman and Edward M. Reingold, 'Graph drawing by force directed placement', *Software: Practice and Experience* 21:11 (1991) 1129–64.
- 24 The method was developed by Aaron Clauset, M.E.J. Newman and Christopher Moore, 'Finding community structure in very large networks', arXiv (<http://www.arxiv.org/abs/cond-mat/0408187>); an implementation can be found in Gabor Csardi and Tamas Nepusz, 'The igraph software package for complex network research', *InterJournal, Complex Systems* 1695:5 (2006) 1–9.
- 25 Our metric of dpf controls for background frequency, thereby in part discounting for the fact that the later decades of the century contain more textual information. The takeaway here is that the density of plot indicates

- a far richer set of co-associations within and across sub-networks: the language of government had become richer.
- 26 For example D. L. Rohde, L.M. Gonnerman and D.C. Plaut, 'An improved model of semantic similarity based on lexical co-occurrence', *Communications of the ACM* 8 (2006) 627–63.
 - 27 David Simpson argues that theoretical abstractions were inimical to English during this period and this was one of the ways in which Britain distinguished itself from its political rival, France and the French language. See David Simpson, *Romanticism, Nationalism, and The Revolt Against Theory* (Chicago, 1993).
 - 28 This criterion excludes some highly frequent words that happen to end in 'ism' such as 'criticism,' 'schism,' 'mechanism,' etc.
 - 29 Turchetti, *Despotism and Tyranny*; Melvin Richter argues, by contrast, that the semantic fuzziness of 'tyranny', 'despotism', etc, is a strong incentive not to re-appropriate such terms for the present age. See Melvin Richter, 'A Family of Political Concepts: Tyranny, Despotism, Bonapartism, Caesarism, Dictatorship, 1750–1917', *European Journal of Political Theory* 4:3 (2005) 221–48.
 - 30 Richter, 'Family of Political Concepts', 229–30.
 - 31 Ibid, 231.
 - 32 Ibid, 244.
 - 33 These extracts come respectively from William Coxe, *A letter to the Rev. Richard Price, D. D. LL. D. F. R. S. &c. Upon his "discourse on the Love of our Country," delivered November 4, 1789* (London, 1790), 44; Charles McCormick, *The history of England, from the death of George the Second to the peace of 1783* (3 vols; London, 1795) II, 53; *A letter to the Rev. Richard Price, D. D. LL. D. F. R. S.*, 41; William Laurence Brown, *An essay on the natural equality of mankind* (London, 1799) 145; *The Scotch Preacher: Or, A Collection of Sermons* (Edinburgh, 1789) 333; *Some hints in regard to the better management of the poor: in a letter to a noble lord* (London, 1784) 25.
 - 34 John Thomas Troy, *A pastoral instruction on the duties of Christian citizens, addressed to the Roman Catholics of the Archdiocese of Dublin* (Dublin, 1793) 2.
 - 35 David Hume, *Essays and Treatises on Various Subjects* (London, 1758) 7.
 - 36 Jacques-Vincent Delacroix, *A Review of the Constitutions of the Principal States of Europe* (London, 1792) 276.
 - 37 Montesquieu, *The Spirit of the Laws* (7th edition; Edinburgh, 1778) 140.

- 38 Tobias Smollett, *A Complete History of England* (4 vols; London, 1754) II, 83.
- 39 These extracts come respectively from *Experience preferable to theory* (London, 1776) 84; *Liberty and property preserved against Republicans and Levellers* (London, 1793) 10.
- 40 John Brooks, *An Oration to the Society of Cincinnati* (Boston, 1787) 11.
- 41 See also Alexander Knox, *Essays on the political circumstances of Ireland* (Dublin, 1799) 200; Madame de Staël, *A treatise on the influence of the passions, upon the happiness of individuals and of nations* (London, 1798) 178.
- 42 Richter, 'Family of Political Concepts', 234.

About the Authors

Peter de Bolla is Professor of Cultural History and Aesthetics at the University of Cambridge, where he also directed the Cambridge Concept Lab. His most recent monograph is *The Architecture of Concepts: The Historical Formation of Human Rights* (Fordham University Press, 2013) which won the Patten award in 2015. E-mail: pld20@cam.ac.uk

Ewan Jones is lecturer in nineteenth-century literature at the University of Cambridge. He is the author of *Coleridge and the Philosophy of Poetic Form* (Cambridge University Press, 2014), and numerous articles on prosody, literary history and digital humanities. He is currently writing a second book on the history of the concept of rhythm.

Paul Nulty is Marie Skłodowska-Curie Career-FIT Fellow at the Centre for Applied Data Analytics Research at University College Dublin. In 2013 he gained his PhD from the School of Computer Science in University College Dublin, on the topic of understanding and predicting lexical expressions of semantic relations between nouns. His research focuses on applications of natural language processing to questions in digital humanities and social science. Specific interests include computational models of semantic relations between entities, and methods for extracting political concepts and ideological positions from text.

Gabriel Recchia is a research associate within the Department of Pure Mathematics and Mathematical Statistics at the University of Cambridge. Until recently, he was a research associate at the Centre for Research in the Arts, Social Sciences and Humanities, where he developed distributional models for the analysis of large corpora of historical texts, and investigated conceptual change by attending to shifting statistical associations between words over time.

John Regan is Lecturer in Literature and the Digital at Royal Holloway, University of London. He is the author of *Poetry and the Idea of Progress* (Anthem, 2018) and is currently writing a book provisionally titled *How the Eighteenth Century Knew: A Digital-Conceptual History*.