

1.5 million words of Mary Dorothy George: a computational approach to curatorial voice and legacy descriptions of art objects

James Baker, University of Sussex

Andrew Salway, University of Sussex

Between 1930 and 1954 the historian Mary Dorothy George wrote descriptions of 12,522 satirical prints held at the British Museum. These were published as volumes 5 to 11 of the *Catalogue of Political and Personal Satires Preserved in the Department of Prints and Drawings in the British Museum* (hereafter BMCSat): 1.5 million words of text that describe satirical prints published in London between 1771 and 1832. BMCSat was a substantive work of scholarship that elevated their subjects to canonical status among historians of long-eighteenth century British history (Donald, 1996). In turn, both the prints and George's descriptions of them – made in a curatorial-cum-art historical setting – have been integral to studies of print, culture, politics, and social life in Georgian Britain.

But BMCSat is also a primary source of George's curatorial voice: of the choices made, schemas followed, and patterns set during curatorial practice. In this paper we demonstrate how new knowledge about curatorial voice can be generated by combining text classification and corpus linguistic analysis of curatorial art descriptions with the kinds of archival research more common to historical and art historical research. We ask:

- What aspects of the prints did George address?
- What aspects of curatorial voice were shaped by temporal and cultural factors?
- How did her practice change over time?

We find that George's descriptions contain a veneer of clear, neutral, and confident voice. This was a choice. By combining corpus methods (word lists, collocation, keyness measures) with archival research, we tease out the overarching and subtle patterns of choice in George's work, both how and why a consistent lexicon performed rigour, made the reader of an individual or a hand-curated selection of descriptions confident in George's authority, and permitted them to build from her prose a mental picture of the image being described. In this paper we examine in detail the ways in which her descriptions use spatial language, a feature of George's voice that emerged from an inductive corpus approach (machine classification). Originally intended for use both with and without the printed images they refer to, we describe how George's descriptions navigate the reader across the two-dimensional plane of the printed page and into the (imagined) three-dimensional space of the satirical print. We examine how the structure of her descriptions moved between spatial and non-spatial sentence parts, and transcribed texts like captions, speech bubbles, and labels. And we demonstrate the ways in which these insights can be used to interpret George's interpretations of late-Georgian satirical prints.