British patents of invention 1711- 2011: A corpus-based diachronic generic structure analysis

Nicholas Groom (University of Birmingham, UK)

This paper has three interrelated aims. The first aim is to present a novel corpus-based methodology for the diachronic analysis of generic structure. The second aim of the paper is to present the results of an empirical study which used this methodology to identify and study changes in the generic structure of a diachronic corpus of British patent texts spanning three centuries. The third aim of the paper is to consider the implications of my diachronic analysis of patent texts for a key theoretical question in current diachronic genre studies: is genre change best understood as a Darwinian, evolutionary process, or is it better understood as a series of Kuhnian, revolutionary paradigm shifts?

The methodology to be described in the first part of this paper follows Biber and Conrad (2009) in regarding generic structure as being principally marked by overt textual features such as section headings and subheadings, fixed phrases, special formatting and so on. The first stage of the methodology therefore involves identifying these overt textual features. The function(s) of the generic stages marked out by these overt features are then established by qualitative analysis, and each stage is given a simple identifying code letter and/or number. To illustrate, Figure 1 shows a generic structure coding for patent no. 1700 ('Elin's Improvements to Shoe Buckles'), which was published in 1789:

Code	*Function of generic stage*
z	Identifying addressees of the text
p1	Identifying author of the text
y1	1st person statement that monarch has granted patent protection
x1	1st person statement of condition of grant
w1	1st person statement of fulfillment of condition of grant
v	Description of invention
u1	1st person witness and/or signature
q	Other witness signatures
te	English-language confirmation that specification has been lodged on a particular date
n	Drawings

Figure 1: Generic structure coding for patent 1700/1789

Following this method, the generic structure of each complete exemplar text is thus reduced to and represented as a relatively simple code string, such as *zp1y1x1w1vu1qten* for the patent described above. These code strings can then be processed using adaptations of standard corpus analysis techniques. The key technique for identifying patterns of diachronic change involves using

an adaptation of dispersion plotting to identify when and how particular variant forms emerge, how long they persist for, and whether and how they mutate or disappear from use altogether over time.

In the second part of the paper I will present the results of an empirical study which used this methodology to identify and study changes in the generic structure of a corpus of British patent specification texts between 1711 (the year in which the world's first patent specification was published) and 2011. The corpus was compiled from the complete collection of over 2 million historical UK patent documents held at The British Library, and (for more recent texts) from the European Patent Office's *Espacenet* online patent search interface (http://www.epo.org/searching-for-patents/technical/espacenet.html#tab1). On the basis of this analysis I will argue that there have been five major transformations in the structural form of the British patent specification genre in its three hundred years of continuous existence. I will interpret these generic changes in social and functional terms, showing how they can be related to concurrent changes in intellectual property law and its conceptual underpinnings, to developments in science and technology, to the growth of manufacturing industry and other forms of commercial activity during the period, and to broader developments in British society and politics as a whole.

In the final section of my talk I will discuss the implications of my empirical analysis of patents for a current theoretical question in diachronic genre studies: is the process of genre change best understood in Darwinian terms, as a matter of constant and gradual evolution (e.g. Gross et al, 2002), or is it better understood in Kuhnian terms, as a series of relatively stable periods of activity punctuated by occasional and abrupt revolutionary shifts (e.g. Berkenkotter, 2009)? In general, the results of my study lean more towards a revolutionary than an evolutionary account of genre change. However, I will also point out a number of features which seem to be better described in evolutionary terms. Ultimately, I will caution against an overly literal Kuhnian interpretation of my data. I will also suggest that the aptness of either an evolutionary or a revolutionary interpretation of the results of a diachronic genre analysis may also depend to a considerable extent on the function and status of the genre in the society in which it operates, and on the level of the genre analysis itself.

References

Berkenkotter, C. (2009). *Patient tales: Case histories and the uses of narrative in psychiatry*. Columbia, SC: University of South Carolina Press.

Biber, D. & Conrad, S. (2009). *Register, genre, and style*. Cambridge: Cambridge University Press.

Gross, A. G., Harmon, J. E., & Reidy, M. (2002). *Communicating science. The scientific article from the 17th century to the present.* Oxford: Oxford University Press.