

Text mining uncovers British reserve and US emotion

Writers' expressions of sentiment have grown apart in recent decades.

Philip Ball

21 March 2013

If you associate modern British fiction with the cool, detached tones of Martin Amis and Julian Barnes, and US fiction with Jonathan Franzen's emotional inner worlds or John Irving's sentimentality, it seems you have good reason. An analysis of the digitized texts of English-language books over the past century concludes that, since the 1980s, words that carry emotional content have become significantly more common in US books than in British ones.

The study¹, by anthropologist Alberto Acerbi of the University of Bristol, UK, and his colleagues, takes advantage of Google's database of more than 5 million digitally scanned books from the past several centuries. This resource has previously been used to examine the evolution of literary styles² and trends in literary expressions of individualism³.

Such mining of the cultural information made available by new technologies has been called 'culturomics'⁴. Its advocates think that these approaches can unearth trends in social opinions and norms that are otherwise concealed within vast swathes of data.

"Language use in books reflects what people are talking about and thinking about during a particular time, so Google Books provides a fascinating window into the past," says psychologist Jean Twenge of San Diego State University in California.

The latest results certainly seem to show that familiar narratives about social mood are reflected in the literature (both fiction and non-fiction) of the twentieth century. Acerbi and his colleagues find that, whereas words connoting happy emotions show peaks of usage in the 'roaring twenties' and the 'swinging sixties', sad words come to the fore during the years of the Second World War.

But there are surprises too: the First World War doesn't seem to register on this happy-sad index, for example. By the same measure, happiness seems to be rising since the 1990s, although it is too early to see whether the global recession will reverse that because the database extends only to 2008.

Historical trends

"The relationship between historical events and collective mood is complicated," Acerbi admits, "but just by doing a somewhat crude analysis of emotion words it is possible to find trends that resonate with what we know about history." He hopes that further analysis might reveal, for example, whether literature is ahead of its time or only slowly reflects other changes.

"This is a fascinating look at how two cultures have changed over time, especially how world events influence the expression of emotion in media," says Twenge.

Overall, the use of emotion-related words in English-language books declined over the twentieth century. But distinguishing between books written using American or British English (about 1 million and 230,000, respectively) told another story.

The authors found that, despite the overall decline, emotion words have become relatively more frequent in US texts than in British books since about 1980. Conversely, before then, any differences between books from the two sides of the Atlantic had been minor. Such changes were not seen for general words selected at random. "Our results ... support the popular notion that American authors express more emotion than the British," they write.

Shifting styles

A similar change is seen in the usage of 'content-free' words, such as pronouns and prepositions (such as you, us, about, within).



Horner Sykes/Getty

The frequency of occurrence of certain words has evolved differently in US and British books over the last few decades.

Acerbi and his colleagues interpret this as indicating that the shift in emotionality is coupled to a general shift in literary style. US texts, they say, are becoming increasingly prolix.

“The correlation with mood terms is not altogether surprising, as these longer constructions provide increased opportunity for expressing sentiments,” explains biologist David Krakauer of the University of Wisconsin–Madison, who with his colleagues has mined Google Books for changes in literary style².

“Authors tend to read their contemporaries and their competitors largely within their respective cultures,” he adds, “and so we might expect British English and American English to diverge somewhat.”

Do these shifts imply that the US population in general expresses more emotion than the British? Although that doesn’t necessarily follow — literary norms may sometimes invert rather than mirror tendencies in everyday life — Acerbi feels that the findings “may reflect a genuine cultural change, because of the size of the sample, and because Google Books is not explicitly biased towards successful or influent books”.

But Krakauer cautions that differences in literary expression don’t necessarily represent differences in the emotional mindscapes behind them. “It is a rather intriguing and open question why different cultures express the same level of feeling with different numbers of words,” he says.

Nature | doi:10.1038/nature.2013.12642

References

1. Acerbi, A., Lampos, V., Garnett, P. & Bentley, R. A. *PLoS ONE* **8**, e59030 (2013).
2. Hughes, J. M., Foti, N. J., Krakauer, D. C. & Rockmore, D. N. *Proc. Natl Acad. Sci. USA* **109**, 7682–7686 (2012).
3. Twenge, J. M., Campbell, W. K. & Gentile, B. *PLoS ONE* **7**, e40181 (2012).
4. Michel, J.-P. *et al. Science* **331**, 176–182 (2011).