Corpus Analysis of Gobbledygook

In July 2013, the British Civil Service issued a new style guide for is online documents, *Government Digital Service Content Principles*. In itself, this is nothing remarkable: all forms of publishing require style guides to ensure consistency across authors, text types and topics. What is remarkable, however, is the peculiarly prescriptive tone of the document, combining Grice's (1975) maxims with observations on undesirable language use that is reminiscent of Orwell's 1945 essay 'Politics and the English Language'. Readers of the document are urged not only to "use plain English" (§1.5), but also reminded that "... government 'buzzwords' and jargon [...] are too general and vague and can lead to misinterpretation or empty, meaningless text. *We can do without these words*" (ibid., my emphasis). In this contribution I discuss how the 37 'offending' expressions listed in the style guide are used in online policy documents, and attempt to connect this with (i) the choice of expressions listed, (ii) the definitions provided next to many of them, suggesting what their 'proper' meaning is, i.e. how they *should* be used, and (iii) if those explicitly listed as metaphors "to be avoided" are indeed metaphorical.

Corpus linguists are well aware that it is the phrasal environment of a word that fixes its meaning; and they are also familiar with the notion that the frequent use of particular phraseologies can lead to delexicalisation, i.e. a 'loss of identity' of individual words relative to the form and function of the phrase of which they form part. So it should not come as a surprise that the words on the black list often appear in phrasal sequences as well as in compositional (or "open-choice", Sinclair 1991) language; and it is in such phrasal uses that their meaning does indeed seem to vanish into "empty, meaningless text". What needs to be investigated, however, is which factors cause meaning to slip from our grasp – whether it is found in all phrasal uses, or if it is possible to identify more precise indicators as to the *kinds* of phrase (e.g. syntactic role, composition, length, variability) that are associated with a loss of meaning, at least as far as this text genre is concerned.

The data studied comprises the policy documents published in the year 2013, available from https://www.gov.uk/government/policies . WordSmith Tools (Scott 2008) was used to compile word lists from the documents' titles as well as from the complete corpus, as a means to verify the presence and frequency of the style guide's list of 'offenders' (the frequencies vary considerably from term to term, and some only occur in inflected forms). Once their presence was verified, each term was subjected to 'traditional', KWIC-concordance analysis in order to identify their meaning(s) and use(s) in the UK government policy documents; these were in turn compared with patterns in general language use using the BNC as a reference corpus. Findings to date indicate that civil service gobbledygook is a phraseological phenomenon that (i) is register-and genre-specific, and (ii) delexical rather than metaphorical.

References

Grice, H.P. 1975. Logic and Conversation. In Cole &. Morgan (eds) *Syntax and Semantics vol. 3*, pp. 41-53. New York: Academic Press.

Orwell, G. 1945. Politics and the English Language.

Scott, M. 2008. WordSmith Tools version 5. Liverpool: Lexical Analysis Software.

Sinclair, J.M. 1991. *Corpus, Concordance, Collocation*. Oxford: Oxford University Press.

Government Digital Service Content Principles. Available: https://www.gov.uk/design-principles/style-guide (accessed 30 July 2013).