# Using R to determine the specialisation of press discourse

Caroline Peynaud[*][1]

[1]Institut des Langues et Cultures dÉurope, Amérique, Afrique, Asie et Australie (ILCEA4) – Université Stendhal - Grenoble 3 : EA7356, Université Grenoble Alpes – ILCEA4 - Université Grenoble Alpes UFR Langues étrangèresCS 40700 - 38058 GRENOBLE CEDEX 9, France

**Résumé**

The specialisation of a discourse may be determined based on many different criteria, such as the terminology used, the genres, as well as on cultural aspects highlighted via ethnographic study. Many of these aspects, such as the construction of genres, cannot easily be analysed by the means of corpus analysis while others, such as terminology, can be understood via the collection of textometric data.

Press discourse poses a specific problem when it comes to terminology, as the terms used by journalists belong to different specialised domains treated by the press, such as economics or health, but do not belong to the press domain and thus do not reflect the specific specialisation of this profession. However, terms borrowed from other specialised domains are integrated into press discourse and it may be assumed that the way the press uses terms certainly reflects the specialisation of this field. This paper thus aims at answering the following question: is it possible to understand the specialisation of press discourse via an analysis of the terminology it contains run with R?

Several steps are necessary to answer this question. First, a task-specific corpus has been built, composed of papers dealing with COVID-19 and published in the first period of the pandemic, from January to March 2020. This period was characterized with great uncertainty about the disease itself, which reflected in the terms used to describe it. The texts have been taken from many different journalistic sources so as to compare them and reveal whether they belong to the same specialised domain or not. Terminology and phraseology related to definitions and explanations of the main terms have been analysed as signs of the way in which the press treats specialised knowledge. The analyses have been carried out with R, the functioning of which makes it possible to avoid the "black box" effect (Pasquale, 2015) and thus to ensure reproductibility.

Pasquale, F. (2015). *The Black Box Society: the Secret Algorithms that Control Money and Information.* Harvard: Harvard University Press.

Resche, C. (2013). *Economic terms and beyond: Capitalising on the wealth of notions. Bern:* Peter Lang.

Swales, J. (1990). *Genre analysis: English in academic and research settings.* Cambridge: Cambridge University Press.

Wozniak, S. (2019). *Approche ethnographique des langues spécialisées professionnelles.* Bern : Peter Lang.

[*]Intervenant