

Gilquin, G. 2019. 'Review of Georgeta Cislaru & Thierry Olive, 2018, Le processus de textualisation: Analyse des unités linguistiques de performance écrite'. *Discourse Studies* 21.5: 603-605.

Georgeta Cislaru and Thierry Olive, *Le processus de textualisation: Analyse des unités linguistiques de performance écrite*, Louvain-la-Neuve: De Boeck Supérieur, 2018; 271pp., €37.00 (pbk).

**Reviewed by:** Gaëtanelle Gilquin, *Université catholique de Louvain, Belgium*

Georgeta Cislaru and Thierry Olive's book, whose title can be translated as "The process of textualisation: Analysis of the linguistic units of written performance", investigates the online process through which expert and semi-expert texts are written, by relying on keystroke logging data, which record the keys struck on a keyboard while the texts are being typed. The focus is on 'bursts of writing' (*jets textuels*) as linguistic units, corresponding to sequences of text produced between pauses of over two seconds. The study of these bursts at different levels of analysis provides crucial insights into the incrementality of the written performance and contributes to developing a theoretical and methodological framework to deal with this new type of linguistic unit. The book is made up of five chapters, preceded by an introduction and followed by a conclusion, as well as an appendix (the instructions of a writing task), a glossary of twenty-three important terms, and an index.

The **introduction** of the book outlines the general background of the study and presents the corpus on which it relies. The corpus was collected during the production of social reports on at-risk children written by French social workers (expert writing) and of academic texts written by French university students within the frame of a seminar (semi-expert writing). As explained in **Chapter 1**, while research into the writing process was carried out in the past, for example through verbal protocols or by comparing different versions of a text (cf. genetic editing), digital writing and the development of keylogging software have made it possible to capture this process more comprehensively and more precisely. Two different approaches to the keylogging data are identified: macro-segmentation, which looks at the different working sessions over which a text is written, and micro-segmentation, the segmentation into bursts. **Chapter 2** reviews several theoretical frameworks that could accommodate the analysis of bursts. Parallels are drawn with the study of speech, in which sequences can be demarcated by pauses (just like bursts) and which is characterised by disfluencies (similar to editing in the writing process). Chunks, which correspond to groups of words strongly associated with each other, and constructions, which in Construction Grammar are defined as pairings of form and function, are also considered in relation to bursts. The relevance of different frameworks for the analysis of bursts leads Cislaru and Olive to adopt an approach drawing upon several of them, including those just mentioned, but also functional and cognitive theories, corpus linguistics, and discourse analysis. **Chapter 3** is the first chapter of analysis. It provides an overview of the bursts found in the social reports and academic texts (over 12,000 bursts in total), distinguishing between saturated bursts (i.e. those that correspond to complete entities like phrases or clauses) and non-saturated bursts (i.e. unfinished sequences). As much as 60% of the bursts are non-saturated. Among these, some patterns are investigated more thoroughly, including the *X Prep* and *X Det* bursts, ending with a preposition or a determiner, respectively. The semantic function of the sequence is argued to be more important than its syntactic structure in accounting for the presence of pauses. **Chapter 4** discusses how certain elements within bursts can create links that contribute to the status of bursts as units. Four such types of elements, called *attracteurs de junction* ('junction attractors'), are examined,

namely 'strong' punctuation marks (full stops and colons), connectors, coordinating conjunctions, and noun phrases with a possessive or demonstrative determiner. Bursts thus turn out to have a stronger internal logic than is suggested by their boundaries (cf. non-saturated bursts in Chapter 3). **Chapter 5** explores the interface between the writing process and the product, by comparing aspects of the bursts and of the finished texts (e.g. repeated segments). It also demonstrates that bursts have a role to play in textual organisation, for instance through the development of macro-sequences of the type PROBLEM-SOLUTION, and it investigates how these sequences get realized through the different writing sessions. The **conclusion** summarizes the main contributions of the book and points to some possible avenues for research.

Cislaru and Olive offer one of the very first book-length treatments of the writing process as investigated through a linguistic analysis of keylogging data. Although it deals with French, it should be of interest to anyone who wants to know more about the construction of the text (provided they can read French). It comes with a solid theoretical background and includes a rich literature review, showing how concepts from several (linguistic and other) fields can be useful to the topic at hand. While it is thus strongly anchored in the established views of language, and writing in particular, it has to move away from the traditional analyses of grammatical constituents when tackling bursts. This is because segmentation into bursts takes place spontaneously, through pausing in the writing process, rather than being imposed post hoc by the analyst, and bursts therefore do not necessarily reflect traditional grammatical structures. Although dealing with such self-segmented units is thus not always easy, as Cislaru and Olive themselves acknowledge, this bottom-up approach allows for a more objective description of the data, avoiding normative biases. The data are described with the right balance of quantitative and qualitative information. Readers are provided with figures, but do not get lost in them, and while they are mostly shown the global results per corpus, sometimes they also get to see the specificities of certain individual writers. The abstract patterns are described through (some of) their concrete realisations, and the many examples in the book give a good sense of what the actual data look like. Multiple-level constraints are taken into account in the analysis, from grammatical to discursive constraints, through semantic and pragmatic ones. One can perhaps regret that the nature of the corpus (social reports written by experts and academic texts written by semi-experts) does not really make it possible to distinguish between the influence of the genre constraints and that of the degree of expertise, and that the former seems to be favoured over the latter in the explanations. However, given that the comparison between the two types of texts is not central to the study, this should rather be seen as an avenue to explore in future studies. Indeed, the fact that one closes the book with so many ideas for further research proves that Cislaru and Olive have managed to share their enthusiasm for the topic and to convince readers of its relevance and value.