

Towards a Taxonomy of Writing Activities

Yoram M Kalman

The Open University of Israel
yoramka@openu.ac.il

Laura K Allen

University of New Hampshire
Laura.Allen@unh.edu

ABSTRACT: This workshop paper describes efforts to develop a taxonomy of writing activities that are involved in the production of a new text. The taxonomy under development is a first step towards the automated characterization of the typing and pointing activities which go into the writing and editing of texts on digital devices.

Keywords: Writing analytics; Taxonomy; Logger; Editing

1 WRITING AND EDITING ON DIGITAL DEVICES

The advent of digital devices and applications and the consequent proliferation of computer-based writing opens the opportunity for writing analytics scholars to study not only the product of the writing process (Shermis, Burstein, & Zechner, 2010), but also the process of writing (Van Waes, Leijten, Lindgren, & Wengelin, 2016) This endeavor has important implications for both theories of writing as well as real-world applications. A better understanding of how texts are produced and edited can improve our theoretical understanding of the cognitive processes involved in writing, the stages of text creation and development, individual differences amongst writers, and the linguistic properties associated with high-quality writing. These insights can in turn be applied to improve writing education, through the diagnosis and monitoring of specific writing features to the delivery of real-time feedback, and much more.

The primary tools used for studying writing processes on digital devices are loggers which track the keystrokes and pointing events that occur as individuals create texts. Basic loggers track only the keystrokes – i.e. the keys individuals hit as they create the text. More advanced loggers also record the evolving text in a detailed manner (*Draftback*, 2019; Kalman, Adam, & Blau, 2019). These loggers track every writing and editing activity, and their subsequent impact on the evolving text. Analyses of the output of these loggers have raised the need to create a taxonomy of the writing activities that individuals enact as they produce texts. Such a taxonomy will simplify the results of these loggers by interpreting the captured information and transforming it into specific named activities (e.g., typo correction). Consider cooking as an example -- it is simpler to state that "the onion was chopped" rather than describing how the cook took a knife, cut the onion top, split it into two halves, peeled the onion, lay the onion half on a cutting board, made many vertical slices etc.

In this workshop we describe our initial steps in creating a taxonomy of writing and editing activities. In particular, we will present the evolving taxonomy and discuss its usefulness, limitations, and our anticipated future directions.

2 A PRELIMINARY TAXONOMY OF WRITING ACTIVITIES

An ideal taxonomy of writing and editing activities would unambiguously classify all activities of a writer. Our work on developing this taxonomy relies on input from other taxonomies (e.g. in biology) and from classifications used in the teaching and analysis of writing (Van Waes, Leijten, Lindgren, & Wengelin, 2016). The development of the taxonomy has been iterative, relying on a trial and error method where classification criteria have been applied to essay writing samples (in English) in an effort to identify both redundancies and gaps. The work is still ongoing, and our most current version of the taxonomy will be presented in the workshop. Below we describe some of the primary features of the developing taxonomy.

The taxonomy characterizes activities that occur at three levels of the text: word, sentence and document. It classifies activities that occur at each level, while recognizing that as long as a writer is modifying a text at a lower level, it is futile to classify the activity at a higher level. For example, as long as I am still modifying a word (e.g. changing the word *home* to the word *house*), it is unnecessary to attempt to classify the modification I made to the sentence that contains that word. The taxonomy creates three "histories" which describe the full evolution of each unit (word, sentence, document) as it is created and as it evolves.

In the workshop we will demonstrate how the "histories" produced by the taxonomy help to characterize the evolution of each of the units. We will specifically present our work in progress on the taxonomy, let participants experiment with using the taxonomy, and discuss limitations and directions for future research.

REFERENCES

- Draftback*. (2019). Retrieved from <http://draftback.com/>
- Kalman, Y. M., Adam, E., & Blau, I. (2019). A Novel Writing Analytics Approach to Study Multiple Information Sources Integration by Students. *Companion Proceedings 9 Th International Conference on Learning Analytics & Knowledge (LAK19)*, 9–10. Retrieved from <http://wa.utscic.edu.au/wp-content/uploads/2019/03/Kalman-Adam-Blau-2019-A-Novel-Writing-Analytics-Approach-to-Study-Multiple-Information-Sources-Integration-by-Students.pdf>
- Shermis, M. D., Burstein, J., & Zechner, K. (2010). Automated essay scoring: Writing assessment and instruction. In *International encyclopedia of education* (pp. 20–26). Elsevier Science.
- Van Waes, L., Leijten, M., Lindgren, E., & Wengelin, Å. (2016). Keystroke logging in writing research: Analyzing online writing processes. In *Handbook of writing research, 2nd ed.* (pp. 410–426). New York, NY, US: The Guilford Press.