

Micro- and Macro-Level Features of NLP-Based Writing Tools in Higher Education

Michael BURKHARD^{a*}, Sabine SEUFERT^a, Patcharin PANJABUREE^b,
Chailerd PICHITPORNCHAI^c & Christina NIKLAUS^d

^a*Institute for Educational Management and Technologies, University of St. Gallen, Switzerland*

^b*Faculty of Education, Khon Kaen University, Thailand*

^c*Institute for Innovative Learning, Mahidol University, Thailand*

^d*Institute for Computer Science, University of St. Gallen, Switzerland*

*michael.burkhard@unisg.ch

Abstract: This paper reviews tools that use natural language processing (NLP) to support writing in higher education. These NLP-based writing tools usually include various features such as grammar and spelling suggestions, but also general feedback such as the overall readability or consistency of the text. In contrast to previous studies, these feedback features are analyzed and described in detail to give a more accurate overview of the current capabilities of these tools. In a systematic analysis, we present 14 NLP-based writing tools that provide different types of micro-level feedback (e.g., grammar, spelling), macro-level feedback (e.g., argumentative structure, rhetorical moves), as well as additional services (e.g., plagiarism checks, access to writing templates). In contrast to previous findings, the state-of-the-art analysis revealed that many writing tools provide feedback at the micro-level and the macro-level. However, macro-level feedback and further additional services are often only included in paid subscription plans. Overall, the presented features of NLP-based writing tools could be helpful in designing own use cases as they may serve as a checklist and guideline to develop tools that go beyond what already exists.

Keywords: Academic writing, natural language processing, writing tools, higher education

1. Introduction

Academic writing skills are imported 21st-century skills (van Laar et al., 2017; Oberländer et al., 2020) and help students convey their understanding and critical thinking. However, in many contexts, it is challenging to promote *academic writing skills* (Thaiss & Zawacki, 2006; Stevenson & Phakiti, 2014). For example, lecturers often do not have the time to provide detailed and individual feedback to each student. The sparse feedback students receive usually provides them with inadequate guidance on how to develop their writing skills further and leads to frustration among students (Seufert & Spiroudis, 2017).

To address this issue, recent advances in artificial intelligence (AI) and natural language processing (NLP) can be used to provide students with an additional, individual, and adaptive feedback channel to improve the overall quality of the text (Zhang, 2013; Rapp & Kauf, 2018; Strobl et al., 2019). NLP refers to the processing of natural language information with the help of a computer (Gu, 2020). For this purpose, elements from the field of linguistics are combined with various AI approaches (Ding, 2019; Vajjala et al., 2020). In the context of this paper, we define *NLP-based writing tools* as tools, programs, or apps that provide NLP-based feedback on an inserted text. These *NLP-based writing tools* usually include various *features* such as grammar and spelling suggestions, but also general feedback on specific aspects of the text (e.g., readability, persuasiveness of arguments, degree of formality, consistency).

Previous research has already examined *writing tools used in education* (e.g., Allen et al., 2016; Cotos, 2015; Stevenson & Phakiti, 2014). Allen et al. (2016) suggested a classification of three different computerized writer support functions, which they called *Automated Writing Evaluation* (AWE), *Automated Essay Scoring* (AES), and *Intelligent Tutoring Systems* (ITS). The literature review of Strobl et al. (2019) analyzed 44 writing tools used in an academic context, distinguishing between *process-oriented* and *product-oriented* tools. According to Strobl et al. (2019), *process-oriented* tools

are *interactive writing platforms* (IWP) that support learners' learning process using prompts and scaffolds. In contrast, *product-oriented* tools generally include *automated writing evaluation* (AWE) and can provide formative feedback on a specific text (by using NLP).

Within AWE, Strobl et al. (2019) further distinguish between *micro- and macro-level* feedback. *Micro-level* feedback is based on factual knowledge and provides feedback on individual words or sentences (e.g., grammar, spelling). Instead, feedback on the *macro-level* refers to the text as a whole and provides feedback on the overall context of the text (e.g., argumentative structure, rhetorical moves) (Strobl et al., 2019, p. 33). In their analysis, Strobl et al. (2019, p. 44) conclude that writing tools that provide *macro-level* feedback are still rare.

Although the literature review by Strobl et al. (2019) is very valuable, it looks at writing tools in secondary and higher education together, even though the writing needs for the two audiences are different in many ways. For example, the structure of scientific papers differs from that of other text forms, which are more commonly used in secondary education (e.g., narrative text types, poetry). Additionally, citation according to academic standards plays a more important role in higher education. Students often have to write their first academic paper without knowing academic writing requirements (Kruse & Chitez, 2014). For this reason, it might be meaningful to look at writing tools intended for use in higher education. Due to the different target groups, these tools may also have different features. Regarding the review of Strobl et al. (2019), a more detailed assessment of the feedback features offered by the different writing tools could also be conducted. This assessment would provide a more accurate overview of the current capabilities of these writing tools. Moreover, since NLP research has made great progress in recent years, it could well be that writing tools have improved significantly since 2019. Therefore, a revisited analysis in the more specific context of *higher education* might be helpful. Accordingly, the objective of the paper at hand is to describe and analyze *NLP-based writing tools used in higher education*, obtain a more nuanced picture, and reflect on the meaningful use and prospects of these tools. Moreover, in light of the identified research desideratum, the following research question should be addressed: *What are the state-of-the-art features of NLP-based writing tools used in higher education?*

From a theoretical point of view, the paper at hand contributes to a better understanding of the usefulness of *NLP-based writing tools in higher education*. From a practical standpoint, the presented features of *NLP-based writing tools* could help design use cases, as they may serve as a checklist and guideline to develop tools that go beyond what already exists.

To this end, section 2 provides more detailed information about the applied method and the conducted search procedure to identify *NLP-based writing tools used in higher education*. Section 3 describes the identified *writing tools* and their features. Section 4 discusses the findings critically and section 5 concludes with some final remarks.

2. Method

Based on the literature review by Strobl et al. (2019), the goal was to identify *NLP-based writing tools for higher education*. First, inclusion criteria were defined that all tools had to fulfill. The goal was to find state-of-the-art *writing tools* that are (1) still working today, (2) focus on higher education, and (3) provide NLP-based feedback on an inserted text (*micro- or macro-level*).

Figure 1 summarizes the review process. The 44 digital writing tools described in Strobl et al. (2019) review were taken as a starting point. Additional supplementary desk research revealed six further potentially interesting writing tools for higher education that had not already been included in the literature review by Strobl et al. (2019). In total, 50 potentially relevant writing tools were identified.

In a first step, an attempt was made to access each of the 50 tools to see if they still existed and could be used. Several tools had restricted access (e.g., behind a login wall), but many tools were no longer available. In total, 15 writing tools had to be removed because these writing tools did no longer exist or were at least no longer publicly available.

In a second step, the remaining 35 writing tools were assessed in more detail and examined concerning the defined inclusion criteria. In total, 21 writing tools had to be removed because they did not meet the above inclusion criteria. Particularly, writing tools were removed that did not provide

personalized NLP-based feedback (*micro-* or *macro-level*) but supported the writing process primarily with the help of scaffolding elements (see, e.g., EssayJack, n.d.; Rapp & Kauf, 2018).

In total, 14 writing tools remained for further fine-grained analysis. These tools will now be described in detail in the next section.

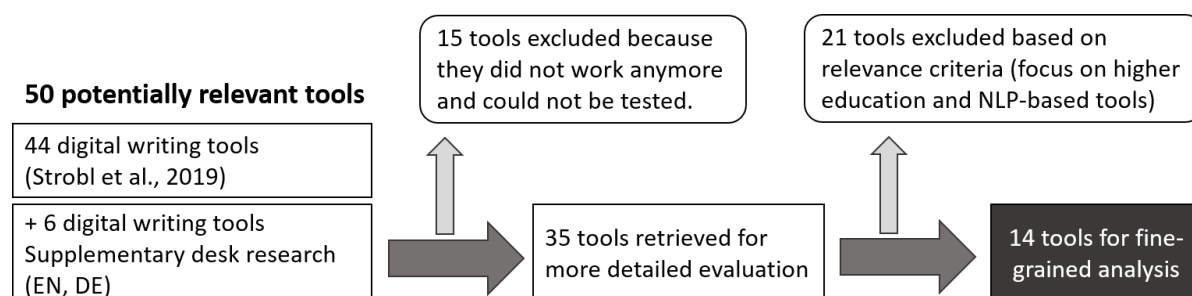


Figure 1. Review process.

3. Results: State-of-the-Art Analysis

Table 1 gives an overview of the 14 identified writing tools in higher education. The majority of the tools are freely accessible, intended to support English writing and can be tested free of charge. It is necessary for many of the tools to create a user account. The tools *AcaWriter* (1), *ArgueLearn* (2), and *CorpuScript* (3) are only accessible to students of the particular university. Through the information provided on the respective website, these tools nevertheless allow a transparent insight into their functionality and features.

Table 1. *NLP-based tools supporting student writing in higher education.*

Tool	Name	Link
(1)	AcaWriter	https://acawriter.uts.edu.au/
(2)	ArgueLearn	https://iwi.unisg.ch/en/projects/argulearn/
(3)	CorpuScript	https://www.schrijven.ugent.be/corpuscript
(4)	CorrectEnglish (Vantage Learning)	https://www.correctenglish.com/
(5)	Criterion	https://www.ets.org/criterion/about/
(6)	De-Jargonizer	http://scienceandpublic.com/
(7)	Enago	https://www.enago.com/researcher-hub/
(8)	Grammarly	https://app.grammarly.com/
(9)	Marking Mate (standard version)	http://writingtools.xjtlu.edu.cn:8080/mm/markingmate.html
(10)	PaperRater	https://www.paperrater.com/
(11)	Scientific Writing Assistant (SWAN)	https://cs.joensuu.fi/swan/
(12)	Trinka	https://www.trinka.ai/
(13)	Turnitin (Revision Assistant)	https://www.turnitin.com/
(14)	Writefull	https://www.writefull.com/writefull-for-word

Most of the tools are web-based applications. However, some tools can only be used in the form of a Microsoft-Word-Plugin (*Writefull* (14)) or must be downloaded and installed like regular software (*Scientific Writing Assistant* (11)). The various writing tools offer different services, but they also have many things in common. Table 2 provides a summary overview of the different forms of feedback offered by these tools.

Table 2. *Features of NLP-based tools supporting student writing in higher education.*

Tool	Micro-level	Macro-level	Additional features
(1)	Writing suggestions (e.g., sentence	Overall reflective report and feedback with individualized advice	-

	length, style)		
(2)	-	Feedback on readability, coherence, and persuasiveness; feedback on the structure of arguments (claims and premises)	-
(3)	-	Check of word count; auto-generated suggestions for further reading	Plagiarism check
(4)	Spelling; word suggestions	-	Format documents for MLA or APA citation standards; plagiarism check
(5)	Diagnostic feedback (e.g., grammar, style)	-	-
(6)	Highlights jargon in writing	Suitability for general audience score	-
(7)	Grammar	-	Plagiarism check
(8)	Spelling; grammar; punctuation	Checks on conciseness, readability, inclusive language, fluency, level of formality	Plagiarism check
(9)	Grammar; punctuation	Academic style rating; overall score rating with individualized advice	Reference check
(10)	Grammar; writing suggestions	-	Plagiarism check
(11)	-	Evaluates the structure of a scientific paper; gives individualized advice on different sections (e.g., abstract)	-
(12)	Grammar	Overall consistency check	Publication readiness checks (e.g., journal scope match, technical and ethical compliance)
(13)	Grammar; writing suggestions	-	Plagiarism check; similarity check; reference and citation check
(14)	Language feedback for the scientific content	Exploration of writing patterns	Title generator; paraphraser; academic phrase bank

As shown in Table 2, almost all tools provide some forms of *micro-level feedback*. This form often involves general feedback on grammar and spelling combined with alternative wording suggestions. Moreover, many tools also provide *macro-level feedback* that tries to assess or relate to the overall quality of a text. Some types of *macro-level feedback* may be quite simple such as the overall check of the allowed word count (*CorpuScript* (3)). Other types of *macro-level feedback* are more sophisticated and provide different forms of overall score ratings. For example, *De-Jargonizer* (6) highlights the words in a text according to their frequency and calculates how suitable a text is for the general audience. Based on the feedback, text passages with rarely used words (that may be difficult to understand for the general audience) can be adjusted. *AcaWriter* (1), *Grammarly* (8), and *Marking Mate* (9) provide the user with a detailed overall report and individualized advice on how to improve the text further. The overall reports usually comprise several subcategories, whose recommendations can be followed in detail. *Marking Mate* (9) has a subcategory called *academic style rating* that checks for typical deficiencies such as the use of informal expressions or colloquial words. *Grammarly* (8) gives the user feedback on different aspects of the overall quality of the text, such as the perceived fluency or the level of formality.

The *Scientific Writing Assistant* (11) can even provide detailed feedback on the text structure. For example, the *Scientific Writing Assistant* (11) can point out that certain elements that are considered important in the text structure are not mentioned in the abstract; or that certain parts of the paper (e.g., literature background) are probably too short or too long, relative to the rest of the text. However, to precisely establish and analyze these connections, the individual text elements (abstract, title, introduction, conclusion, etc.) must be entered separately, which makes the handling somewhat more complicated. *ArgueLearn* (2) also addresses the overall structure of a text and checks how well the arguments are constructed. For this purpose, *ArgueLearn* (2) identifies and marks claims in the text along with any premises that support a particular claim. In this way, the user can review the arguments made and improve them if necessary.

In addition to *micro-* and *macro-level* feedback, most tools incorporate further *additional features* that may support the user in its writing. Often, this includes plagiarism and citation standard checks (see Table 2). *Writefull* (14) includes an academic phrase bank (where typical sentences used in the academic context can be used as a template), a paraphraser, and a title-generator feature. *Trinka* (13) can run several publication readiness checks, such as a journal scope match and technical and ethical compliance checks.

In addition to the free versions, there often exist additional paid subscription plans for many writing programs (*CorrectEnglish* (4), *Enago* (7), *Grammarly* (8), *PaperRater* (10), *Trinka* (12), *Writefull* (14)). These paid subscription plans promise enhanced features (e.g., unlimited usage, faster processing, ad-free, improved *macro-level* feedback, access to writing templates and external databases, publication readiness, and plagiarism checks) and usually cost between 8 and 24 USD per month. The length of the subscription (month or year) often affects the price.

Concerning Table 2, it is noteworthy that *micro-level* feedback is generally offered free of charge, while *macro-level* feedback and other functions such as plagiarism checks often have to be paid for (at least after a certain level of usage). Considering that Microsoft Word (www.microsoft.com/word) or OpenOffice (www.openoffice.org) can already provide relatively good feedback at the *micro-level* (e.g., regarding grammar), it seems plausible that only the provision of “basis” *micro-level* feedback services generally are for free.

4. Discussion

The analysis in the last section has revealed that today several writing tools exist in higher education that can provide feedback not only at the *micro-* but also at the *macro-level*. This analysis contributes to a more nuanced picture, as Strobl et al. (2019) concluded that only a few tools provide *macro-level* feedback. However, the various tools often have a modular nature and are composed of different (*micro-* and *macro-level*) features that can add user value.

The analysis of existing *NLP-based writing tools in higher education* has also shown that in contrast to institutional providers (e.g., universities), many private providers often pursue commercial interests. In the sense of a freemium business model, basic services are usually offered free of charge, but additional improved features have to be activated via paid subscription plans.

On the one side, this raises questions of fairness since not all students may afford to access improved writing tools, and their usage can be an advantage when writing texts. On the other side, however, the motivation of the commercial providers must also be questioned, who may be pursuing different interests with their writing tools than higher education institutions (who want to use writing tools primarily in a pedagogically meaningful way). Are commercial providers concerned about what writing tools are *best* for the user in terms of learning gains or rather about what is the *most convenient* for the user? As learners may not always know best (Kirschner & van Merriënboer, 2013), commercial providers could capitalize on this and encourage a dependency on their writing tools. New approaches may be needed to ensure that users have the competencies to use writing tools to their advantage without becoming dependent on them in the long run.

5. Conclusion

This paper systematically compared fourteen NLP-based writing tools used in higher education. The state-of-the-art analysis revealed that many writing tools provide feedback at the *micro-level* (e.g., grammar, spelling) and *macro-level* (e.g., argumentative structure, rhetorical moves). However, *macro-level* feedback and other enhanced features (e.g., plagiarism checks, access to writing templates) are often only included in paid subscription plans. Overall, the presented features of *NLP-based writing tools* could be helpful in designing own use cases as they may serve as a checklist and guideline to develop tools that go beyond what already exists. Future research in NLP-based writing tools should aim to develop applications that provide pedagogical value in multiple ways. Among other things, this could be achieved through the modular and simultaneous integration of different feedback elements at the *micro-* and *macro-level*, as they were presented here in this paper.

References

- Allen, L. K., Jacovina, M. E., & McNamara, D. S. (2016). *Computer-Based Writing Instruction*. ERIC. <https://eric.ed.gov/?id=ED586512>
- Cotos, E. (2015). Automated Writing Analysis for Writing Pedagogy: From healthy tension to tangible prospects. *Writing & Pedagogy*, 7(2-3). <https://doi.org/10.1558/wap.v7i2-3.26381>
- Ding, F. (2019). *How to AI: Navigating the buzzwords of artificial intelligence*. Retresco. <https://www.retresco.com/how-to-ai-natural-language-processing/>
- EssayJack (n.d.). *Wish academic writing was easier?*. EssayJack. <https://www.essayjack.com/>
- Gu, M. (2020). *NLP Glossary for Beginners: Demystify Natural Language Processing with these terms*. Medium. <https://medium.com/analytics-vidhya/nlp-glossary-for-beginners-c3093529ee4>
- Kirschner, P. A., & van Merriënboer, J. J. (2013). Do learners really know best? Urban legends in education. *Educational psychologist*, 48(3), 169-183. <https://doi.org/10.1080/00461520.2013.804395>
- Kruse, O., & Chitez, M. (2014). Schreibkompetenz im Studium: Komponenten, Modelle, Messung. In Dreyfurst, S. & Sennewald, N. (Eds.). *Grundlagentexte zur Theorie, Didaktik und Beratung*, 107-126. Verlag Barbara Budrich.
- Oberländer, M., Beinicke, A., & Bipp, T. (2020). Digital competencies: A review of the literature and applications in the workplace. *Computers & Education*, 146, 103752. <https://doi.org/10.1016/j.compedu.2019.103752>.
- Rapp, C., & Kauf, P. (2018). Scaling Academic Writing Instruction: Evaluation of a Scaffolding Tool (Thesis Writer). *International Journal of Artificial Intelligence in Education*, 28, 590–615. <https://doi.org/10.1007/s40593-017-0162-z>
- Seufert, S. & Spiroudis, E. (2017) Wissenschaftliche Schreibkompetenz in der Studieneingangsphase fördern: Erfahrungen mit einem Rückmelde- und Bewertungsraster (Rubric). *Zeitschrift Schreiben*, 2, 1–9.
- Stevenson, M., & Phakiti, A. (2014). The effects of computer-generated feedback on the quality of writing. *Assessing Writing*, 19, 51–65. <https://doi.org/10.1016/j.asw.2013.11.007>.
- Strobl, C., Ailhaud, E., Benetos, K., Devitt, A., Kruse, O., Proske, A., & Rapp, C. (2019). Digital support for academic writing: A review of technologies and pedagogies. *Computers & Education*, 131, 33–48. <https://doi.org/10.1016/j.compedu.2018.12.005>
- Thaiss, C., & Zawacki, T. M. (2006). *Engaged writers and dynamic disciplines. Research on the academic writing life*. Boynton/Cook.
- Vajjala, S., Majumder, B., Gupta, A. & Surana, H. (2020). *Practical Natural Language Processing: A Comprehensive Guide to Building Real-World NLP Systems* (1. Ed.). O'Reilly Media. <https://www.oreilly.com/library/view/practical-natural-language/9781492054047/>
- van Laar, E., van Deursen, A.J. A. M., van Dijk, J. A. G. M., & de Haan, J. (2017). The relation between 21st-century skills and digital skills: A systematic literature review. *Computers in Human Behavior*, 72, 577-588. <https://doi.org/10.1016/j.chb.2017.03.010>
- Zhang, M. (2013). Contrasting automated and human scoring of essays. *Educational Testing Service R&D Connections*, 21, 1–11. https://originwww.ets.org/Media/Research/pdf/RD_Connections_21.pdf