Chapter 10: Natural Language Processing - Writing Analytics

Andrew Gibson,[1] Antonette Shibani[2]

[1]Science and Engineering Faculty, Queensland University of Technology, Brisbane, Australia
[2]University of Technology Sydney, Sydney, Australia

DOI: 10.18608/hla22.010

ABSTRACT

Writing analytics uses computational techniques to analyse written texts for the purposes of improving learning. This chapter provides an introduction to writing analytics, through the discussion of linguistic and domain orientations to the analysis of writing, and descriptive and evaluative intentions for the analytics. The chapter highlights the importance of the relationship between writing analytics and good pedagogy, contending that for writing analytics to positively impact learning, actionability must be considered in the design process. Limitations of writing analytics are also discussed, highlighting areas of concern for future research.

Keywords: Writing Analytics, natural language processing, NLP, linguistics, pedagogy, feedback

Writing analytics (WA) is a sub-field of learning analytics (LA) that uses natural language processing (NLP) technologies to analyse written text for the purposes of improving learning. WA may be directed at obtaining insight on the writer, the writer's thinking, or the writing itself. While theoretically it could be applied to any written text, the majority of WA research has focused on learning contexts and associated writing artefacts. Common applications include analysing student writing ability, providing feedback to students and teachers on writing content and style, researching learning, investigating interactions through analysis of dialogue, analysing opinion, and automated grading. The significance of language to LA can be seen in the extent of language related chapters in this handbook, such as social network analysis [43], reading [1], multi-party interaction [16], and writing on which we focus here.

1 THE PURPOSE OF WRITING ANALYTICS

Many of the different applications of WA make use of the same computational NLP techniques. For example, tokenising, parsing, and vectorisation are used in many WA applications. Other NLP approaches such as word embedding, information extraction or topic modelling may only be used for specific applications. NLP is an extensive and fast-moving computer science field which has recently adopted contemporary machine learning practices to address many language problems [54]. Most WA work takes advantage of relatively mature NLP techniques and uses readily available open source software. Therefore, in this chapter we focus more on how WA uses NLP to obtain analytics relevant to learning, than on the NLP technologies themselves.

Although WA has been used for purposes tangential to learning (e.g. curriculum document analysis, or sentiment analysis of student surveys), here we concentrate on applications which are pedagogically driven and closer to student learning contexts.

The way in which learning relates to writing is important, particularly when WA is mostly focused on features of the writing artefact, as opposed to characteristics of the learner or aspects of the learning context. When students are learning to write, they are learning the vocabulary and the syntactic and stylistic rules for writing for a specific purpose (e.g. writing a persuasive essay as opposed to a piece of short fiction). When WA is applied to such "learning to write" tasks, the analytics are generally constructed around the key requirements of a style of writing (e.g. well-formed sentences and appropriate use of domain vocabulary). By contrast, many tasks involve "writing to learn". In these tasks, the rules of the writing may be of secondary importance, or only necessary for comprehensibility, with the primary focus being the content of the writing. With this kind of writing task, WA is used more to ascertain whether the writer has grasped important domain concepts and is expressing them in a way that demonstrates their learning (e.g. expressing relevant domain knowledge appropriately).

The difference between learning to write and writing to learn is not always clearly distinguished in writing tasks, however these different emphases tend to require selection of different NLP techniques and a difference in orientation to the analysis. The *orientation* of learning to write tasks tends to be more linguistic where formalities of language are primary, whereas writing to learn tends to demand more of a domain orientation where meaning associated with a topic is primary. Regardless of the orientation adopted, the *intention* for the WA tends to be a mix of description and evaluation. Descriptive analytics identify the presence or absence of certain features, the extent to which they occur, and how they interrelate. Evaluative analytics make judgements about the nature of the writing and the extent to which is or is not fit for purpose. The relationship between orientation to the writing and intention for the writing analytics is described in more detail next.

2 ORIENTATION TO ANALYSIS OF WRITING

Writing is a complex activity involving skilful management of cognitive, social, and affective processes [18, 24]. A writing artefact not only includes information about the subject of writing, but also incorporates information about the writer's skill in writing, stylistic characteristics of their writing, and can at times reveal information about how the writer thinks about the subject , as well as personal information about themselves.

The analysis of writing precedes the use of computational tools [20], and WA has been extensively influenced by non-computational approaches to analysis. Fundamentally, analysis tends to be approached from a mixture of two orientations, one from a *linguistic* standpoint and another that comes from a *domain* standpoint.

2.1 Linguistic Orientation

A linguistic orientation to writing draws heavily on the field of linguistics, and when applied to WA makes extensive use of computational linguistics to inform the analysis. This orientation is typically concerned with the technical aspects that apply to most writing and is heavily influenced by theories of language including Universal Grammar [11], and Functional Grammar [15]. This orientation can also result in division of analysis approaches that align with areas of linguistics, and which value the structural characteristic of natural language and associated features.

Linguistically oriented WA tends to focus on language features that are formally defined. These can include, but are not limited to: parts of speech, grammatical rules of sentence structure, word relationships in the form of syntactic dependencies or phrase structures, and vocabulary use including spelling and lexical diversity. Tools such as Linguistic Inquiry and Word Count or LIWC [42], Stanford CoreNLP [34], and Coh-Metrix [38] derive such linguistic measures, which can further be used to generate higher order writing analytics. Formal definition of language features allow WA to build on the linguistic rules and definitions towards a meaningful representation of the writing. Linguistic features can usually be determined with an accepted level of accuracy by NLP algorithms, and as long as the writing follows accepted language conventions.

2.2 Domain Orientation

Until recently, almost all of NLP technologies were heavily influenced by linguistics. With the advent of neural machine learning approaches to NLP, many contemporary NLP technologies use models trained on large corpuses of billions of words (or more) and are able to converge on the conventions of language without explicit coding of the formal structures. Some of the most successful language technologies are essentially based on complex computational models which require little or no understanding of linguistics. This shift in NLP approach resulted in significant debate between linguists and computer scientists [41] as to whether a linguistic understanding was actually necessary in order to computationally work with language [53]. Differences in perspective along these lines have not only informed the direction of NLP research, but have also influenced the WA community, particularly in the choice of underlying NLP technologies.

An ability to analyse writing without recourse to linguistics increased the relevance of taking a domain orientation to the analysis of writing. In contrast to the linguistic orientation which is interested in the technical and formal aspects of language, a domain orientation focuses more on the purpose of the written text and its content. Characteristics of writing that are of primary importance include but are not limited to: the topic or subject of the text, the meaning of text with particular relevance to the domain context, the tone and use of emotion and affect, and overall characteristics related to style and genre - the "abstract, socially recognized ways of using language" [26, p. 149].

For example, a domain orientation may be more interested in whether the writing addresses a domain topic in a required style or genre, as opposed to whether it follows language conventions or specific linguistic rules. This can be seen in machine learning classifiers that predict user-defined categories for a text (Say, predicting if an argumentative written text is a premise or a claim, built on human-annotated data sets and no linguistic rules). The context in which the writing is generated plays a major role in determining which features are of interest in a domain orientation, making them hard to generalise to different contexts.

As WA is practical field, approaches to WA tasks tend to be a mixture of the above orientations. For WA that is strongly aligned with pedagogy, this mixture of orientations can be determined by directing the analysis towards the task Knight, Gibson & Shibani [31], following a pragmatic approach towards a practical effect in the learning.

3 INTENTION FOR WRITING ANALYTICS

Regardless of the orientation towards analysing the writing, writing analytics tend to reflect an intention that blends *describing* features of the writing, and *evaluating* the writing with respect to some expected utility. These are not distinct categories, but rather ways of understanding the nature of analytics and its relationship to the written text and/or the writing task. The extent to which one is more or less prominent is governed by the intended purpose and the alignment with relevant pedagogy.

3.1 Descriptive Writing Analytics

Descriptive WA makes visible and summarizes features of the writing to inform the writer or other stakeholders (e.g. teacher). Such analytics are based on technical constructs at word-level, sentence-level, paragraph-level or whole document-level, and various structural features of the writing.

One basic measure that provides summary information is the frequency of occurrence (count) of a feature within a given scope. The total number of words, paragraphs, and sentences in document may be calculated to provide counts as feedback to the writer, to make sure that they are adhering to task requirements. Word processors, most text editors and grammar and spelling software provide this simple form of feedback. Although a simple form of analytics, this type of feedback can highlight aspects of the writing that do not adhere to accepted language conventions and can therefore aid the writer in editing their writing. However, feedback dominated by simple metrics associated with the mechanics of writing can lead to a focus on error correction, which may be of minimal pedagogic value for the writer [10], particularly in writing to learn tasks.

Descriptive WA can also be metrics of more complex linguistic features or a mix of simple features that are indicative of a more complex construct. These can include measures of cohesion, complexity, connectives in a text, and psycholinguistic data such as textual familiarity, sentiment, and effect. Examples of descriptive WA that bring together both simple frequency counts and more complex linguistic and psycholinguist constructs include Linguistic Inquiry and Word Count or LIWC [42], and Coh-Metrix [38]. Several studies have explored how these indices predict writing quality and writer characteristics, by finding correlations between selected indices and human ratings of essay quality [13, 37, 46].

Specific structural patterns of interest such as rhetorical moves and connectives in writing can also be identified using NLP technologies. Examples of this type of WA are found in AcaWriter [48], Research Writing Tutor [12], and AntMover [4], which identify rhetorically salient sentences associated with a given genre. Descriptive WA can also be generated based on the content of the written text. Key words, concepts and topics in the text can be identified with a range of NLP techniques from simple frequency measures to more advanced techniques such as word embedding [39] or using topic modelling algorithms like Latent Dirichlet Allocation [7]. Tools such as Glosser [52] and Essay Critic [40] use such analytics to bring to the writer's attention to key ideas in the text.

Some descriptive WA have also been employed by teachers and administrators to analyze writing. One example of this application is Quantext [36], which has been used for teacher professional development and analysis of student feedback surveys. Another example can be found in the combination of Coh-Metrix with Social Network Analytics to examine how learners engage with discourse [17].

In addition, how the analytics gets presented to the user plays a major role in how they engage with it. Descriptive analytics may be provided as a report, highlighted in text, and/ or displayed in a dashboard. They may also be graphically represented as plots and graphs to provide visual cues to the users. For examples, see word clouds and rainbow diagrams in OpenEssayist, Concept maps in Glosser [52] and dynamic Revision maps in ArgRewrite [55]. Indeed, visualizations and dashboards are an important part of the conversation in provoking thinking and self-regulated learning in learning analytics [19, 28]. Similarly, visual representations contribute to research on writing by studying products and processes using multiple sources such as drafts, writing logs, keystroke activities, and access logs. They examine writing products: Recurrence Quantification Analysis for instance, which visualizes recurrent word patterns over time [3], and the dynamics of writing processes like revision [2, 49, 50].

Descriptive WA generally leaves meaning making to the writer, teacher or researcher. For the most part, descriptive WA requires a level of understanding from the user to draw reliable and valid conclusions from the analytics. The analytics are simply representations of textual features which require further reflection to be meaningful. Hence, without an explicit meaning-making process, the pedagogic value of such analytics can be questionable. We cannot assume that the descriptive WA is useful for a learner to improve their writing simply because it has been provided to them. Actionable feedback is required for learners to make improvements in their writing, and descriptive WA usually needs to be augmented with this feedback from another source (such as the teacher, or additional materials).

3.2 Evaluative Writing Analytics

Evaluative WA for written texts involves making judgements on the writing to inform the learner about its quality with respect to the writing context. In contrast to descriptive WA, the evaluative WA aims to give students more information on the quality of their writing, rather than requiring them to make their own judgements. This intention for WA holds the potential to provide actionable feedback, informing about the next steps the writer can take. Evaluative WA implementations can vary significantly depending on the educational contexts in which they are used, and some have been developed for very specific purposes, such as analysing for metacognition in reflective writing [22].

A widely used application of evaluative WA is in the provision of automated feedback on writing. This feedback tends to be formative with the aim of supporting the writing process, and hence directly impacting learning. A significant goal is making the feedback actionable, by intervening and guiding the learner to make improvements, thereby completing a learning analytics loop where the analytics is derived from the writing and writing is improved based on the analytics. Examples of tools in this growing research area include AcaWriter [22, 31] , Research Writing Tutor [12] and Turnitin Revision Assistant [35].

Evaluative WA also includes technologies that make judgements on the quality of writing, without necessarily providing feedback to the writer. The most common of these are Automated Essay Scoring (AES) systems which grade assessment tasks that are generally summative rather than formative. This form of evaluative WA is often used in conjunction with high-stakes assessment, and has attracted significant criticism from sectors of the educational community [6] due to the way it is used to support performativity agendas rather than more directly helping learners. The potential for disconnect between summative assessment and learning is well established in the educational literature [51, 25], and AES systems tend not to address this nor other larger pedagogical issues [29, 44]. Further, encoded human judgements in the analytics can carry human biases and errors which may ultimately impact learning decisions.

With respect to WA, an improvement on AES systems are Automated Writing Evaluation (AWE) systems which supplement judgements with feedback. Examples of AWE include Criterion [5] and MyAccess! [32]. While the usefulness of such systems has been demonstrated to support writing instruction at varied levels [9, 33], they have also attracted criticism due to the reduction of writing to merely formulaic features of text, rather than a process of meaningful engagement [10]. Another form of evaluative WA can be found in Intelligent tutoring systems (ITS) such as Writing Pal [47]. ITS provide adaptive and interactive support for learning by providing strategy instruction of writing that are modelled around closed writing tasks. However, such ITS (and AWE and AES) cater to a specific task/ prompt and encourage and evaluate students based on a standard path, which does not necessarily reflect the messy process of learning nor consider outliers. As with descriptive WA, the value of these evaluative WA systems rests to a large extent on how they are situated within good pedagogy. When used to provide formative feedback, it should also provide learners with the opportunity to think critically about their writing, and to push back against the evaluative WA when required, to result in meaningful learning experiences.

4 PEDAGOGY AND WRITING ANALYTICS

Increasingly LA practitioners are attending to the need for LA to make a positive impact on learning [14]. Within the field there is a growing critique of LA approaches that are merely analysing learning related data without due consideration of how that analysis might inform improvements in learning and teaching [21]. WA is no exception to this, and so it is important to consider its relationship to pedagogy.

What constitutes good pedagogy is beyond the scope of this chapter, and in fact dominates the field of Education and Learning Sciences. What is important for an understanding of WA is that its success depends not only on the quality of the technology, but also on the quality of the pedagogy which ultimately determines how WA is put to work. Traditionally, a common educational technology approach has been to take high quality existing technologies and then investigate out how best to apply it within educational contexts. This approach largely rests on the assumption that what makes good technology and what makes good pedagogy are independent static factors. An alternative approach that is often adopted in WA, is to view WA as co-design process which includes both technological and pedagogical aspects, and where each aspect informs design in the other. When adopting this approach, WA naturally tends to be learning focused to augment existing practice, as it grows out the synergistic design of both the technology and the pedagogy. Explicit examples of this approach can be found in task centric WA which builds on both technical and social infrastructure [31].

Good pedagogy demands that WA account for the quality of feedback that is facilitated by its intervention in the learning context. What constitutes good quality feedback is well established in the literature [27], and feedback being actionable, contributing to improvements in learning is a top concern for WA. Therefore, when WA is providing analytics directly to the student, it is important that the student be able to take action based on the analytics received. For WA provided to a teacher, the teacher needs to be able to use the analytics in order to positively impact the learning. There is little that a student or teacher can do if they are presented with feedback that holds no meaning for them, although how they decide to act is a separate issue. Hence, the burden for ensuring that feedback is actionable should fall on the designers of the WA together with the practitioners that implement it. For WA, feedback actionability and quality should be designed in from the beginning, not considered as an afterthought.

Some WA researchers and developers have addressed this need for ensuring actionability by working in multidisciplinary teams [8]. This ensures that WA development is not dominated by NLP experts, but also includes experts in learning and pedagogy like teachers, learning designers, user experience specialists, and cognitive scientists. The constitution of effective WA teams depends on the context in which the WA is expected to be applied, however at a minimum the team needs to include relevant expertise in both NLP and pedagogical domains.

Participatory design has emerged as an important methodology in LA [45] which values the importance of pedagogy by including stakeholders in the design process. Gibson and Lang [23] have also highlighted the importance of pedagogy in the LA research process, recommending a pragmatic inquiry approach that gives priority to the intended practical effects of the analytics, that is, the nature of the effect on learning that is anticipated when the LA is implemented. When applied to WA, both of these methodological approaches can yield meaningful impact on learning [48].

5 LIMITATIONS AND POSSIBILITIES

Despite significant advances in recent years, NLP is still limited to relatively narrow tasks (compared with human processing of language). For WA, this means taking care that NLP technologies are used according to their initial design. For WA designers who are not NLP experts, this means being aware of assumptions that come about from human generalisation of computational processes and ensuring that these are addressed. This may be a non-trivial task as the tendency for computer scientists to co-opt general English language terms for specific computing names can complicate matters. For example, the use of "learning" in the computer science literature is very different to its pedagogical meaning, and assumptions that computers learn like people can be propagated in the WA, resulting in confusion for the user. Similarly, "topic modelling" algorithms do not automatically generate topics in the human form of a subject of interest. Topic models are statistical distributions across a vocabulary and result in a list of words and their corresponding probabilities of belonging to a "topic". These topics can be very useful in WA, but rarely correspond with human interpretation of what the topic might be. Distributional semantics used heavily in NLP aims to detect semantics of words and groups of words based on the words that surround them. However, this is a constrained view in comparison to how humans understand the meaning of words. Humans draw on a much bigger context than the lexical context in which the words are found. When undertaking the task of language understanding, people use prior knowledge and make complex connections that include experiences, emotions, and their physical environment.

Failing to properly comprehend the difference between computational processing of language and human language understanding in WA can have significant pedagogical consequences, as it is possible to design a system that is 'accurate' with respect to a computational analysis of language, but 'useless' with respect to human language understanding. A well-known example of this issue can be found in the simple descriptive analytic of word count. Word-count can correlate highly with quality in some written tasks, but asking a student to improve their writing by writing more words is rarely helpful. NLP limitations are often identified in terms of accuracy in achieving a specific task. However, WA needs to avoid being locked to computer science measures of what is good. Simply because an NLP process is accurate or effective at extracting a textual feature does not mean that it is useful for learning or even necessary for effective writing.

Limitations with NLP, although important to be aware of, do not necessarily translate into limitations in WA. Learning can occur with meaningful design of WA even in the absence of high levels of accuracy [30]. A common example from the teaching of writing underscores this point. Many teachers (over many years) have found that the process of writing drafts is critical to achieving good quality writing for incremental improvement. Often, the reasons for motivating this drafting process are less important than the drafting itself. When WA is concerned with impact on learning, designing WA that encourages writers to write drafts may be more significant than the extent to which the underlying NLP technology is accurate. Accurate NLP is also not equitable with respect to all learners. Many NLP technologies degrade significantly when the language used does not match the norms of usage. Therefore, NLP can perform very poorly on the writing of developing writers and students with language processing difficulties, who don't adhere to the conventions on which the technology depends. Issues of bias can be exacerbated in NLP due to the dominance of development in dominant languages, particularly English. NLP technologies built on English assumptions do not necessarily translate well into other languages, even if software exists. This is particularly the case with machine learning approaches to languages which lack the large corpora on which recent NLP models are trained. Care needs to be taken when designing and implementing WA in contexts of generally good English writers, that success is not assumed for other contexts. The extent to which WA caters for writers of all abilities in all languages could be seen as a measure of the field's maturity, and on this measure at this point in time, there is a long way to go.

The key to maximising the potential of WA despite its limitations, is an inextricable relationship with high quality pedagogy. For WA designers, developers and practitioners, this means working together with educators and holding a clear shared understanding of the practical learning effect that they wish to achieve.

REFERENCES

[1] Laura Allen, Sarah C. Creer, and Püren Öncel. "Natural Language Processing as a Tool for Learning Analytics - Towards a Multi-Dimensional View of the Learning Process". In: *The Handbook of Learning Analytics*. Ed. by Charles Lang, Alyssa Friend Wise, Agathe Merceron, Dragan Gašević, and George Siemens. 2nd ed. Vancouver, Canada: SOLAR, 2022. ISBN: 978-0-9952408-3-4. URL: https://www.solaresearch.org/publications/hla-22/.

[2] Laura K. Allen, Matthew E. Jacovina, Mihai Dascalu, Rod D. Roscoe, Kevin M. Kent, Aaron D. Likens, and Danielle S. McNamara. "{ENTER} ing the time series {SPACE}: Uncovering the writing process through keystroke analyses". In: *International Educational Data Mining Society* (2016).

[3] Laura K. Allen, Cecile Perret, Aaron Likens, and Danielle S. McNamara. "What'd you say again? Recurrence quantification analysis as a method for analyzing the dynamics of discourse in a reading strategy tutor". In: *Proceedings of the Seventh Inter-*

national Learning Analytics & Knowledge Conference. 2017, pp. 373–382.

[4] Laurence Anthony and George V. Lashkia. “Mover: A machine learning tool to assist in the reading and writing of technical papers”. In: *IEEE transactions on professional communication* 46.3 (2003), pp. 185–193.

[5] Yigal Attali. “Exploring the feedback and revision features of Criterion”. In: *Journal of Second Language Writing* 14 (2004), pp. 191–205.

[6] Nan Bahr and Donna Pendergast. “Let’s abandon NAPLAN – we can do better”. In: *The Conversation* (2018). URL: https://theconversation.com/lets-abandon-naplan-we-can-do-better-95363..

[7] David M. Blei, Andrew Y. Ng, and Michael I. Jordan. “Latent dirichlet allocation”. In: *the Journal of machine Learning research* 3 (2003), pp. 993–1022.

[8] Simon Buckingham Shum and T. McKay. “Architecting universities for learning analytics: Innovating for sustainable impact”. In: *EDUCAUSE Review* (2018). URL: https://er.educause.edu/-/media/files/articles/2018/3/er182101.pdf.

[9] Chi-Fen Emily Chen and Wei-Yuan Eugene Cheng Cheng. “Beyond the design of automated writing evaluation: Pedagogical practices and perceived learning effectiveness in EFL writing classes”. In: *Language Learning & Technology* 12.2 (2008), pp. 94–112.

[10] Julie Cheville. “Automated scoring technologies and the rising influence of error”. In: *The English Journal* 93.4 (2004), pp. 47–52.

[11] Noam Chomsky. *Aspects of the theory of syntax*. Cambridge, MA: MIT Press, 1965.

[12] Elena Cotos and Nick Pendar. “Discourse classification into rhetorical functions for AWE feedback”. In: *calico journal* 33.1 (2016), p. 92.

[13] Scott A Crossley, Kristopher Kyle, Laura K Allen, Liang Guo, and Danielle S McNamara. “Linguistic microfeatures to predict L2 Writing Proficiency: A case study in automated writing evaluation”. In: *Grantee Submission* 7.1 (2014).

[14] Shane Dawson, Srecko Joksimovic, Oleksandra Poquet, and George Siemens. “Increasing the impact of learning analytics”. In: *Proceedings of the 9th International Conference on Learning Analytics & Knowledge*. 2019, pp. 446–455.

[15] Simon C. Dik. *The theory of functional grammar: the structure of the clause*. Walter de Gruyter, 1997.

[16] Nia Dowell and Vitomir Kovanović. “Modeling Educational Discourse with Natural Language Processing”. In: *The Handbook of Learning Analytics*. Ed. by Charles Lang, Alyssa Friend Wise, Agathe Merceron, Dragan Gašević, and George Siemens. 2nd ed. Vancouver, Canada: SOLAR, 2022. ISBN: 978-0-9952408-3-4. URL: https://www.solaresearch.org/publications/hla-22/.

[17] Nia M Dowell, Oleksandra Skrypnyk, Srecko Joksimovic, Arthur C Graesser, Shane Dawson, Dragan Gaševic, Thieme A Hennis, Pieter de Vries, and Vitomir Kovanovic. “Modeling learners’ social centrality and performance through language and discourse”. In: *International Educational Data Mining Society* (2015).

[18] Dana Lynn Driscoll and Roger Powell. “States, traits, and dispositions: The impact of emotion on writing development and writing transfer across college courses and beyond”. In: *Composition Forum*. Vol. 34. ERIC. 2016.

[19] Erik Duval. “Attention please! Learning analytics for visualization and recommendation”. In: Proceedings of the 1st international conference on learning analytics and knowledge. 2011, pp. 9–17.

[20] Linda Flower and John R. Hayes. “A cognitive process theory of writing”. In: *College composition and communication* 32.4 (1981), pp. 365–387.

[21] Dragan Gašević, Shane Dawson, and George Siemens. “Let’s not forget: Learning analytics are about learning”. In: *TechTrends* 59.1 (2015), pp. 64–71.

[22] Andrew Gibson, Kirsty Kitto, and Peter Bruza. “Towards the discovery of learner metacognition from reflective writing”. In: *Journal of Learning Analytics* 3.2 (2016), pp. 22–36.

[23] Andrew Gibson and Charles Lang. “The pragmatic maxim as learning analytics research method”. In: Proceedings of the 8th international conference on learning analytics and knowledge. Mar. 2018, pp. 461–465.

[24] Steve Graham, Amy Gillespie, and Debra McKeown. “Writing: Importance, development, and instruction”. In: *Reading and writing* 26.1 (2013), pp. 1–15.

[25] Wynne Harlen. “Teachers’ summative practices and assessment for learning–tensions and synergies”. In: *Curriculum Journal* 16.2 (2005), pp. 207–223.

[26] Ken Hyland. “Genre pedagogy: Language, literacy and L2 writing instruction”. In: *Journal of second language writing* 16.3 (2007), pp. 148–164.

[27] David J. icol and Debra Macfarlane-Dick. “Formative assessment and self-regulated learning: A model and seven principles of good feedback practice”. In: *Studies in higher education* 31.2 (2006), pp. 199–218.

[28] Ioana Jivet, Maren Scheffel, Marcus Specht, and Hendrik Drachsler. “License to evaluate: Preparing learning analytics dashboards for educational practice”. In: *Proceedings of the 8th international conference on learning analytics and knowledge*. 2018, pp. 31–40.

[29] Mary Kalantzis, Bill Cope, and Andrew Harvey. “Assessing multiliteracies and the new basics”. In: *Assessment in Education: Principles, Policy & Practice* 10.1 (2003), pp. 15–26.

[30] Kirsty Kitto, Simon Buckingham Shum, and Andrew Gibson. "Embracing imperfection in learning analytics". In: *Proceedings of the 8th International Conference on Learning Analytics and Knowledge*. 2018, pp. 451–460.

[31] Simon Knight, Antonette Shibani, Sophie Abel, Andrew Gibson, Philippa Ryan, Nicole Sutton, Raechel Wight, Cherie Lucas, Agnes Sandor, Kirsty Kitto, et al. "AcaWriter: A learning analytics tool for formative feedback on academic writing". In: *Journal of Writing Research* 12.1 (2020), pp. 299–344.

[32] Vantage Learning. "How does IntelliMetric score essay responses". In: (2003). URL: `http://www.vantage.com/pdfs/research/RB929.pdf`.

[33] Jinrong Li, Stephanie Link, and Volker Hegelheimer. "Rethinking the role of automated writing evaluation (AWE) feedback in ESL writing instruction". In: *Journal of Second Language Writing* 27 (2015), pp. 1–18.

[34] Christopher D. Manning, Mihai Surdeanu, John Bauer, Jenny Rose Finkel, Steven Bethard, and David McClosky. "The Stanford CoreNLP natural language processing toolkit". In: *Proceedings of 52nd annual meeting of the association for computational linguistics: system demonstrations*. 2014, pp. 55–60.

[35] Elijah Mayfield and Stephanie Butler. "Districtwide implementations outperform isolated use of automated feedback in high school writing". In: (2018).

[36] Jenny McDonald and Adon Christian Michael Moskal. "Quantext: Analysing student responses to short-answer questions". In: Me, us, IT! Proceedings ASCILITE2017: 34th international conference on innovation, practice and research in the use of educational technologies in tertiary education. 2017, pp. 133–137.

[37] Danielle S. McNamara, Scott A. Crossley, and Philip M. McCarthy. "Linguistic features of writing quality". In: *Written communication* 27.1 (2010), pp. 57–86.

[38] Danielle S. McNamara, Arthur C. Graesser, Philip M. McCarthy, and Zhiqiang Cai. *Automated evaluation of text and discourse with Coh-Metrix*. Cambridge University Press, 2014.

[39] Tomas Mikolov, Ilya Sutskever, Kai Chen, Greg Corrado, and Jeffrey Dean. "Distributed representations of words and phrases and their compositionality". In: Advances in neural information processing systems. 2013, pp. 3111–3119.

[40] Anders I. Mørch, Irina Engeness, Victor C Cheng, William K Cheung, and Kelvin C Wong. "EssayCritic: Writing to learn with a knowledge-based design critiquing system". In: *Journal of Educational Technology & Society* 20.2 (2017), pp. 213–223.

[41] Peter Norvig. "Colorless green ideas learn furiously: Chomsky and the two cultures of statistical learning". In: *Significance* 9.4 (2012), pp. 30–33.

[42] James W. Pennebaker, Martha E. Francis, and Roger J. Booth. "Linguistic inquiry and word count: LIWC 2001". In: *Mahway: Lawrence Erlbaum Associates* 71.2001 (2001).

[43] Oleksandra Poquet and Srećko Joksimović. "Cacophony of Networks in Learning Analytics". In: *The Handbook of Learning Analytics*. Ed. by Charles Lang, Alyssa Friend Wise, Agathe Merceron, Dragan Gašević, and George Siemens. 2nd ed. Vancouver, Canada: SOLAR, 2022. ISBN: 978-0-9952408-3-4. URL: `https://www.solaresearch.org/publications/hla-22/`.

[44] Donald E. Powers, Jill C. Burstein, Martin S. Chodorow, Mary E. Fowles, and Karen Kukich. "Comparing the validity of automated and human scoring of essays". In: *Journal of Educational Computing Research* 26.4 (2002), pp. 407–425. DOI: `10.2190/CX92-7WKV-N7WC-JL0A`. URL: `https://doi.org/10.2190/CX92-7WKV-N7WC-JL0A`.

[45] Carlos G. Prieto-Alvarez, Roberto Martinez-Maldonado, and Theresa Dirndorfer Anderson. *Co-designing learning analytics tools with learners. Learning analytics in the classroom: Translating learning analytics research for teachers*. Abingdon-on-Thames, UK: Routledge, 2018.

[46] Zahra Rahimi, Diane Litman, Richard Correnti, Elaine Wang, and Lindsay Clare Matsumura. "Assessing students' use of evidence and organization in response-to-text writing: Using natural language processing for rubric-based automated scoring". In: *International Journal of Artificial Intelligence in Education* 27.4 (2017), pp. 694–728.

[47] Rod D. Roscoe and Danielle S. McNamara. "Writing Pal: Feasibility of an intelligent writing strategy tutor in the high school classroom". In: *Journal of Educational Psychology* 105.4 (2013), p. 1010.

[48] Antonette Shibani. "Augmenting pedagogic writing practice with contextualizable learning analytics". PhD thesis. Australia, 2019. URL: `http://hdl.handle.net/10453/136846`.

[49] Antonette Shibani, Simon Knight, and Simon Buckingham Shum. "Contextualizable learning analytics design: A generic model and writing analytics evaluations". In: Proceedings of the 9th international conference on learning analytics &. Mar. 2019, pp. 210–219.

[50] Vilaythong Southavilay, Kalina Yacef, Peter Reimann, and Rafael A. Calvo. "Analysis of collaborative writing processes using revision maps and probabilistic topic models". In: *Proceedings of the third international conference on learning analytics and knowledge*. 2013, pp. 38–47.

[51] Maddalena Taras. "Summative and formative assessment: Perceptions and realities". In: *Active learning in higher education* 9.2 (2008), pp. 172–192.

[52] Jorge Villalón, Paul Kearney, Rafael A. Calvo, and Peter Reimann. "Glosser: Enhanced feedback for student writing tasks". In: 2008 eighth IEEE international conference on advanced learning technologies. IEEE, 2008, pp. 454–458.

[53] Edda Weigand. "Dialogue and artificial intelligence". In: *Language and Dialogue* 9.2 (2019), pp. 294–315.

[54] Tom Young, Devamanyu Hazarika, Soujanya Poria, and Erik Cambria. "Recent trends in deep learning based natural language processing". In: *IEEE Computational intelligenCe magazine* 13.3 (2018), pp. 55–75.

[55] Fan Zhang. "Towards building an intelligent revision assistant for argumentative writings". (doctoral dissertation, PhD thesis. 2017.