Towards Measuring Linguistic Creativity in Literary and Non-literary Text. First results and insights

Sitter, Emilie

emilie.sitter@uni-bielefeld.de
Universität Bielefeld, Deutschland

Wu, Yaru

yaru.wu@uni-bielefeld.de
Universität Bielefeld, Deutschland

Zarrieß, Sina

sina.zarriess@uni-bielefeld.de
Universität Bielefeld, Deutschland

Herrmann, Berenike

berenike.herrmann@uni-bielefeld.de
Universität Bielefeld, Deutschland
ORCID: 0000-0002-5256-0566

Introduction

Creative signs can be observed in literary and non-literary genres. Computational Linguistics and DH, however, often expect a poetic function of linguistic creativity. They tend to equate it with figurative language (Gervás 2010) or "foregrounding", a deviation from the linguistic norm or an aspect of the text brought to the fore through repetition or parallelism (see (Simpson 2014: 50; Peer et al. 2021). This conflation limits the concept of creativity to artistic genres. Following Runco and Jaeger (2012), in the presented project we therefore conceptualize the creativity of a sign as its originality and effectiveness across genres.

Using data-driven metrics and models, we operationalize linguistic creativity by disentangling originality at the level of the sign from literariness, defined as "density of specific textual features that create linguistic foregrounding" by Appel et al. (2021: 177) and from success at the level of usage, i.e. perception and evaluation by the actors in the discourse domain. This approach considers how readers' perceptions of creativity depend on genres and their internal conventions and usage patterns (Steen et al. 2010).

Objective

The overall goal of the project is to explore to what extent linguistic creativity can be measured across textual genres with quantitative methods. Previous studies on creativity assessment have shown that this seemingly subjective concept can be measured by objective, text-inherent linguistic features (DiStefano, Patterson, and Beaty 2024; Jacobs and Kinder. 2022; Weinstein et al. 2022; Zedelius, Mills, and Schooler 2019). Building on lay readers' creativity assessments of a collection of spatial descriptions obtained from both literary and non-literary texts, we approach creativity from the following perspectives:

- Rhetorics: As fixed sets of forms that *can* elicit foregrounding effects, rhetorical devices are only potentially creative. In usage, their perceived degree of creativity depends heavily on the context.
- Computational stylistics (Herrmann et al. 2015): At the level of passages or whole texts, certain forms of stylistic gestalts may be viewed as linguistically creative, e.g. when stylistic divergent features are integrated (C. Martindale and A. E. Martindale 1988) into the text as one coherent whole (Anderegg 1995) in an original and communicatively successful way. However, so far, empirical research is lacking.
- NLP: From the perspective of computational linguistics, linguistic foregrounding or stylistic deviations from a frequency norm can be interpreted as an "unusual" usage of language. This can be captured at scale using surprisal metrics, such as perplexity (Humpston et al. 2016), attention entropy (Oh and Schuler 2022), and Kullback-Leibler Divergence (KLD) (Kullback and Leibler 1951).

Project Status

At the time of the poster presentation, we will be able to report a first set of findings about linguistic creativity at the sentence-level. We are currently running a set of experts' annotations of rhetorical devices to determine the degree of potential creativity within our data set (see Amabile 1982). In parallel, our piloting of rating studies has provided first methodological insights into how to capture lay readers' perceptions of creativity, and is currently being expanded to our data set.

Our methodological work tests how conceptions of conventional language use shape assessment of originality (Veale 2012: 59). When a text is perceived as literary, linguistic signs appear systematically judged as less original, even though they may have a high level of literariness. This may be a function of exposure to genres, involving predictions from holistic schemas as well as the level of specific forms. Consider the following examples from Robert Walser's *Geschwister Tanner* [1907]:

(1) *Unten lag die Stadt, weit und wollüstig über die Ebene gebreitet, wie ein flimmernder, glitzernder Teppich [...].*

(2) *Stelle dir vor, die Luft wird ganz blau und warmfeucht in die Straßen hinuntersinken, der Himmel geht dann in Paris spazieren und mischt sich unter die entzückten Menschen.*

Example (1) constitutes no strong deviation from predicted literariness in literary texts from around 1900. It might be perceived as literary, but not particularly creative—this is very likely to be different in a non-literary context. Example (2), however, is more unpredictable even in the literary domain, yet meaningful, and can thus be perceived as creative. To operationalize creativity, we therefore need to disentangle literary style from originality across genres at points in time.

In the current phase of the project, we finetune measures to compare lay and expert readers' creativity judgment and correlate them to stylometric features and the surprisal measures of a language model.

Future Outlook

In this four-year project, next steps will be determined by the correlated measures across genres. Creativity assessment will be extended to the full text level, using feature frequencies from distinct style combined with surprisal measures to assess the degree to which a text as a holistic unit is creative and literary.

Presenting our multi-methodological and interdisciplinary approach to measuring linguistic creativity in texts and readers, we hope to engender a fruitful discussion with interested and knowledgeable colleagues.

Acknowledgement

Bibliographie

Amabile, Teresa M. 1982. "Social Psychology of Creativity: A Consensual Assessment Technique". *Journal of Personality and Social Psychology* 43 (5): 997–1013. https://doi.org/10.1037/0022-3514.43.5.997.

Anderegg, Johannes. 1995. "Stil und Stilbegriff in der neueren Literaturwissenschaft". In *Stilfragen*, edited by Gerhard Stickel, 115–27. De Gruyter. https://doi.org/10.1515/9783110622515-007.

Appel, Markus, David Hanauer, Hans Hoeken, Kobie van Krieken, Tobias Richter, and José Sanders. 2021. "The Psychological and Social Effects of Literariness: Formal Features and Paratextual Information". In *Handbook of Empirical Literary Studies*, edited by Don Kuiken and Arthur M. Jacobs, 177–202. Berlin/Boston: de Gruyter.

DiStefano, Paul V., John D. Patterson, and Roger E. Beaty. 2024. "Automatic Scoring of Metaphor Creativity with Large Language Models". *Creativity Research Journal*, March. https://doi.org/10.1080/10400419.2024.2326343.

Gervas, Pablo. 2010. "Engineering Linguistic Creativity: Bird Flight and Jet Planes". In *Proceedings of the NAACL HLT 2010 Second Workshop on Computational Approaches to Linguistic Creativity*, 23–30. Los Angeles. https://aclanthology.org/W10-0304.pdf.

Herrmann, Berenike, Karina van Dalen-Oskam, and Christof Schöch. 2015. "Revisiting Style, a Key Concept in Literary Studies". *Journal of Literary Theory* 9 (1): 25–52.

Humpston, Clara S., and Matthew R. Broome. 2016. "Perplexity". In *An Experiential Approach to Psychopathology: What Is It like to Suffer from Mental Disorders?*, edited by Giovanni Stanghellini and Massimiliano Aragona, 245–64. Cham: Springer International Publishing. https://doi.org/10.1007/978-3-319-29945-7_13.

Jacobs, Arthur M. and Annette Kinder. 2022. "Computational Analyses of the Topics, Sentiments, Literariness, Creativity and Beauty of Texts in a Large Corpus of English Literature". arXiv. https://doi.org/10.48550/arXiv.2201.04356.

Kullback, S., and R. A. Leibler. 1951. "On Information and Sufficiency". *The Annals of Mathematical Statistics* 22 (1): 79–86.

Martindale, Colin, and Anne E. Martindale. 1988. "Historical Evolution of Content of Style in Nineteenth- and Twentieth-Century American Short Stories". *Poetics* 17: 333–55.

Oh, Byung-Doh, and William Schuler. 2022. "Entropy- and Distance-Based Predictors From GPT-2 Attention Patterns Predict Reading Times Over and Above GPT-2 Surprisal". arXiv. https://doi.org/10.48550/arXiv.2212.11185.

Peer, Willie van, Paul Sopčák, Davide Castiglione, Olivia Fialho, Arthur M. Jacobs, and Frank Hakemulder. 2021. "Foregrounding". In *Handbook of Empirical Literary Studies*, 145–76. De Gruyter. https://doi.org/10.1515/9783110645958-007.

Runco, Mark A., and Garret J. Jaeger. 2012. "The Standard Definition of Creativity". *Creativity Research Journal* 24 (1): 92–96. https://doi.org/10.1080/10400419.2012.650092.

Simpson, Paul. 2014. *Stylistics: A Resource Book for Students*. 2nd ed. Routledge. https://www.routledge.com/Stylistics-A-Resource-Book-for-Students/Simpson/p/book/9780415644976.

Steen, Gerard, Aletta G. Dorst, Berenike Herrmann, Anna A. Kaal, Tina Krennmayr, and Trijntje Pasma. 2010. *A Method for Linguistic Metaphor Identification. From MIP to MIPVU*. Vol. 14. Converging Evidence in Language and Communication Research (CELCR).

Amsterdam/Philadelphia: John Benjamins Publishing Company.

Veale, Tony. 2012. *Exploding The Creativity Myth: The Computational Foundations of Linguistic Creativity*. London: Bloomsbury Academic.

Weinstein, Theresa J., Simon Majed Ceh, Christoph Meinel, and Mathias Benedek. 2022. "What's Creative About Sentences? A Computational Approach to Assessing Creativity in a Sentence Generation Task ". *Creativity Research Journal* 34 (4): 419–30. https://doi.org/10.1080/10400419.2022.2124777.

Zedelius, Claire M., Caitlin Mills, and Jonathan W. Schooler. 2019. "Beyond Subjective Judgments: Predicting Evaluations of Creative Writing from Computational Linguistic Features ". *Behavior Research Methods* 51 (2): 879–94. https://doi.org/10.3758/s13428-018-1137-1.