Defying Expectations

Stylistically Unconventional Anger in a Contemporary Dutch Literary Novel

Julia L. Neugarten

Huygens Institute, Amsterdam

ORCID 0000-0003-3314-9445, julia_neugarten@huygens.knaw.nl

Lisanne M. van Rossum

Huygens Institute, Amsterdam

ORCID 0000-0002-5542-5757, lisanne.van.rossum@huygens.knaw.nl

Joris J. van Zundert

Huygens Institute, Amsterdam

ORCID 0000-0003-3862-7602, joris.van.zundert @huygens.knaw.nl

In 2021, Lale Gül won the popular-literary NS Publieksprijs for her novel *Ik ga leven* (eng. *I Am Going to Live),* which narrates her youth and upbringing in a strict conservative muslim household in Amsterdam. The novel became last year's best-selling Dutch novel with over 207,000 copies sold to date (Peppelenbos 2022). After being published, Gül received threats from members of the muslim community.

The book's victory in the NS Publieksprijs, where it obtained 32% of the popular vote, and its high sales figures (De Cocq 2021), stand in stark contrast to its lukewarm reception on the Dutch literary scene, where the book was critiqued for its "unnatural" and "overly complicated" style (Van Houwelingen 2021). In response, Gül herself has explicitly positioned herself and her work away from professional critiques, stating that "I have only received bad critiques… Sorry, but I really look down on critics … I speak with booksellers who tell me that people come to buy my book whom they have never seen in their stores before: young people and non-Western immigrants. That is a lot more meaningful to me than good critiques. I've never received a single negative review from a reader" (Krul 2021, our translation).

As such, Gül and her work *Ik ga leven* seem to be operating on the divide between academic and popular perception and reception of fiction. This tension has been widely studied in academia, with recent work especially centered on literary-sociological notions of prestige (see, for example, Neidhofer 2016; Sandberg 2017; Eatough 2021).

We take a computational approach to this debate by investigating stylistic aspects of *Ik ga leven* and how these interact with (non)literariness of a text. The representation of anger is central to our research: literary journalist Lieke Kézér has drawn parallels between Gül's stylistic informality and the sense of anger that pervades her work. Kézér writes:

"[Gül] does not endow her book with a lot of structure, which reinforces the idea that it has been written in a single irate sigh; a passionate account of what it means to be caught between two cultures [...] Gül is venomous, crass, and funny, but also vulnerable" (2021, our translation).

In our reading of the book we have similarly observed that whenever the narrator's tone becomes more angry, critical or derisive, her style becomes more complex and grandiloquent, with longer sentences, bigger words, and more abstract comparisons. To give an example, when describing the muslim community's response to terrorist attacks, Gül uses a long sentence, abstract terms such as 'indignation' and 'superiority,' and complex references, such as reflexive pronouns and increased semantic distance between subject and relative pronoun.

> "Je kunt na elke aanslag in naam van ons geloof de klok erop gelijkzetten dat mijn gemeenschap daad en motief zal ontkoppelen en daarna verontwaardiging zal uiten over de verontwaardiging; een blijk van morele superioriteit die zichzelf uitlokt, iedere keer als zij geuit wordt door moreel inferieure mensen." (2)

> "After every terrorist attack in the name of our faith, like clockwork, my community will decouple act and motive, and will then express indignation at the indignation; a token of moral superiority that provokes itself, each time that it is expressed by morally inferior people." (our translation)

Many paragraphs which critique the muslim community, like this one, use linguistic complexity to express anger or derision. This seems counterintuitive: although the exact stylistic properties of anger in fiction remain understudied, we assume that anger is associated with a reduction in complexity or nuance in language use, with short sentences, expletives and brief, simple outbursts. Preliminary evidence indicates that expressions of anger in fiction are relatively formulaic on the sentence level (Klégr and Šaldová 2017). Gül's anger appears distinctly different from these stylistic expectations, by deviating from existing formulas and instead employing archaic words, expressions and syntactic structures. Because anger and stylistic deviation from conventions are so interwoven in *Ik ga leven*, we hypothesize that it may be a major contributor to the work's differences in reception.

To test this, we first aim to establish whether our perceived link between anger and increased textual complexity can also be recognized computationally. Establishing this link requires the computational operationalization of two variables: the level of anger in the text, and its level of complexity. We acknowledge that our approach reduces our qualitative reading of the novel to quantifiable variables, which in the case of operationalizing emotion is a constant balancing act between "low-level linguistic feature analysis of emotional language and a rigorous high-level hermeneutic inquiry dissecting the form of the novel and its under-covered philosophical layers" (Kim & Klinger 2018).

To measure the text's anger, we used the Dutch translation by Boot et. al. (2017) of Linguistic Inquiry and Word Count or LIWC (Pennebaker et. al. 2015). LIWC was developed to measure, among other things, the level of emotionality in written statements (Tausczik and Pennebaker 2010). Although LIWC was not developed for literary analysis, parts of the software have also recently been adapted for use in literary studies, for example to measure narrative arcs of texts (Boyd, Blackburn, and Pennebaker 2020). While LIWC works only on the word-level, we use this tool, rather than other tools for sentiment analysis, because it differentiates between anger and other emotions. The software scores text based on the percentage of words it contains that occur in a dictionary relating to a particular textual feature, such as anger. For this analysis, we split Gül's novel into paragraphs and assigned each paragraph an anger-score using LIWC's anger-dictionary. We also scored each paragraph in the LIWC-category "negative emotion", a higher-level category that contains dictionaries for "anxiety", "anger" and "sadness". Taking the paragraph cited above as an example, LIWC recognizes "aanslag" (terrorist attack) as an anger-word and "inferieure" (inferior) as a negative-emotion-word. This paragraph thus scores 2.08% on anger and 4.17% on negative emotion.

We operationalized textual complexity using the GZip compression algorithm to obtain a near-perfect approximation of Kolmogorov complexity for each paragraph. Kolmogorov complexity "measures the informativeness of a given string [...] as the length of the algorithm required to describe/generate that string" (Juola 2008:91). Measuring the linguistic complexity of texts has been found to yield meaningful results for English in both works of fiction and newspaper texts (Ehret 2014), as well as for Dutch (Redactie Wetenschap 2002). We excluded paragraphs shorter than 50 words to ensure that the compression yielded a meaningful ratio and was not an artifact of compression algorithm technicalities (e.g. of storing a hashed vocabulary).

Results are visualized in the scatterplots below. Each dot represents a paragraph, with its orientation on the Y-axis indicating its percentage of words related to the relevant LIWC-category (anger and negative emotion, respectively) and its orientation on the X-axis indicating its complexity expressed as the compression ratio obtained for each paragraph. Assuming that a more complex string is more difficult to compress, a higher ratio corresponds with less compression and thus with higher complexity.

We found a weak positive correlation between the LIWC category 'anger' and complexity as visualized in Figures 1 and 2, with R^2 values of 0.15 for 'anger' and 0.03 for 'negative emotion' (p<0.01). Measuring the same values in dialogue-only paragraphs (i.e. text between quotation marks) yields a corroborating result of R^2 0.36 (p<0.01), depicted in Figure 3.

These results indicate that the levels of complexity of paragraphs of Lale Gül's novel are positively correlated with the occurrence of words relating to anger or negative emotion

in those paragraphs. In other words, paragraphs that are more angry are also more complex and vice versa.
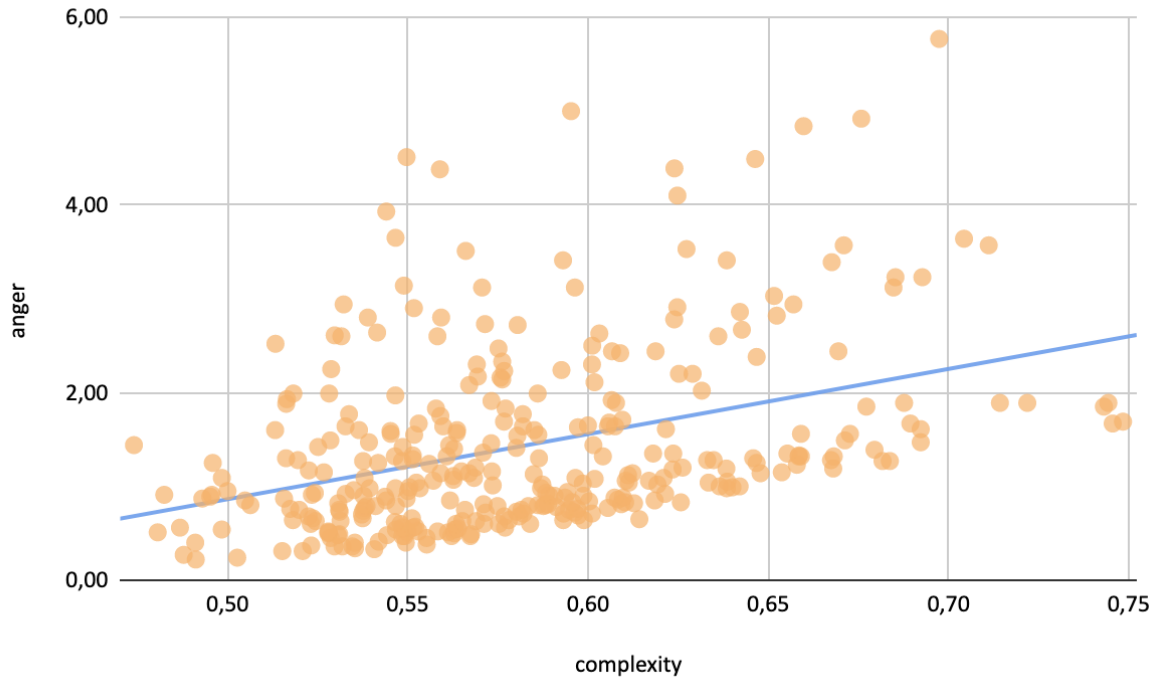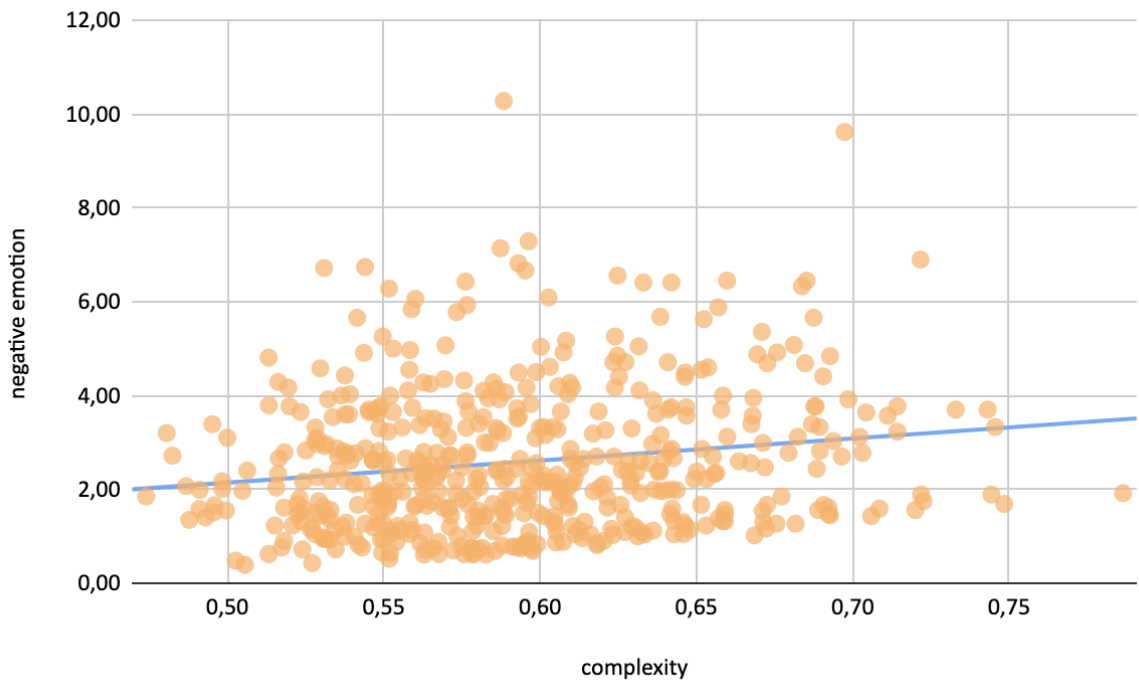

Figure 1: Kolmogorov complexity vs. LIWC category 'anger

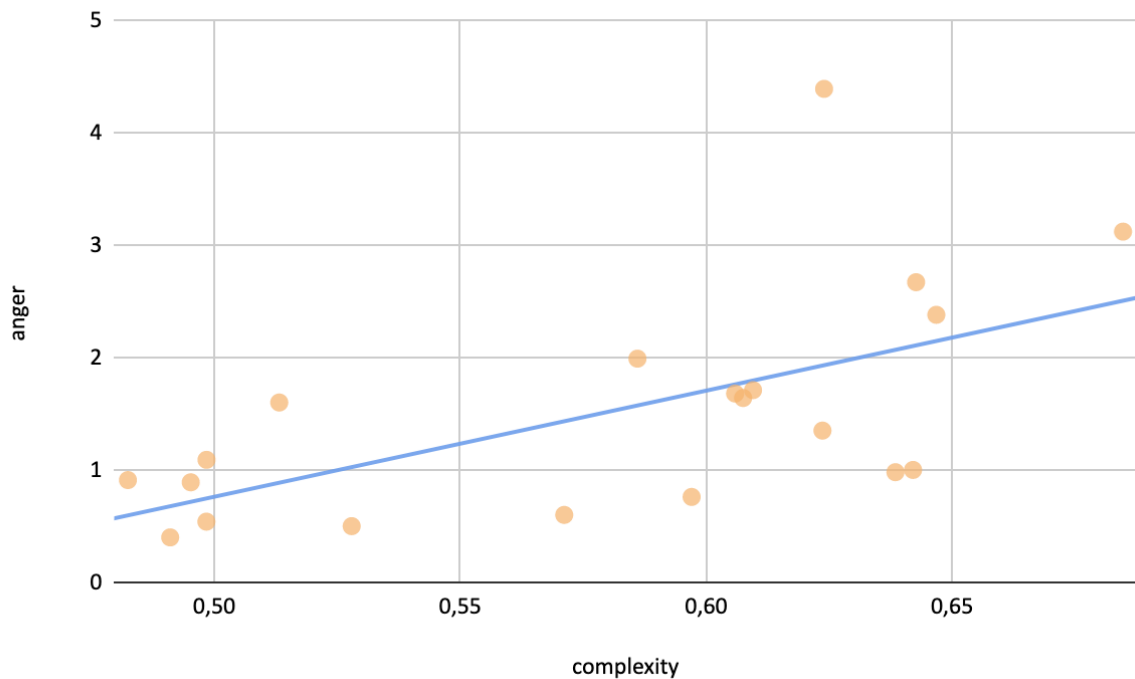Figure 2: Kolmogorov complexity vs. LIWC category 'negative emotion'.

Figure 3: Kolmogorov complexity vs LIWC category 'anger' in isolated dialogue text.

Our hypothesis was informed by intuition about an unconventional stylistic mode of expressing anger within a literary context. We favor this approach over a tool-led framework in which literature is explored as data, a pervasive tenet in computational literary studies critiqued by Da (2020:919). Further subscribing to Da's view, we have shaped our methodology to test hermeneutic interpretation and to critically reflect on tool refinement. While our results are based on conceptual models and approximate the occurrence of textual properties, they indicate that this combination of tools produces meaningful results for analyzing fiction.

Our results present a first step in examining our hypothesis that divergence from stylistic conventions associated with emotional expression inspire divergent reader responses. The results leave ample room for improvement. We are currently developing more detail-oriented measurements of the levels of anger and complexity in this text, for instance by applying T-Scan (Pander Maat et. al. 2014; Pander Maat and Dekker 2016) to generate more fine-grained measurements of textual complexity.

References

Boot, Peter, Hanna Zijlstra, and Rinie Geenen. 2017. "The Dutch Translation of the Linguistic Inquiry and Word Count (LIWC) 2007 Dictionary." *Dutch Journal of Applied Linguistics* 6 (1): 65–76. https://doi.org/10.1075/dujal.6.1.04boo.

Boyd, Ryan L., Kate G. Blackburn, and James W. Pennebaker. 2020. "The Narrative Arc: Revealing Core Narrative Structures through Text Analysis." *Science Advances* 6 (32): eaba2196. https://www.science.org/doi/10.1126/sciadv.aba2196.

Da, Nan Z. 2020. "Critical Response III. On EDA, Complexity, and Redundancy: A Response to Underwood and Weatherby." *Critical Inquiry* 46 (4): 913–24. https://doi.org/10.1086/709230.

De Cocq, Marjolijn. 2022. "Lale Güls Ik Ga Leven Bestverkochte Oorspronkelijk Nederlandstalige Boek 2021, Postuum Record Voor Lucinda Riley." *Het Parool*, January 27, 2022. https://www.parool.nl/kunst-media/lale-guls-ik-ga-leven-bestverkochte-oorspronkelijk-nederlandstalige-boek-2021-postuum-record-voor-lucinda-riley~b36dab6b/.

Eatough, Matthew. ""Are They Going to Say This Is Fantasy?": Kazuo Ishiguro, Untimely Genres, and the Making of Literary Prestige." *MFS Modern Fiction Studies*, vol. 67 no. 1, 2021, p. 40-66. Project MUSE, doi:10.1353/mfs.2021.0002.

Ehret, Katharina. 2014. "Kolmogorov Complexity of Morphs and Constructions in English." In *Linguistic Issues in Language Technology, Volume 11, 2014 - Theoretical and Computational Morphology: New Trends and Synergies*. CSLI Publications. https://aclanthology.org/2014.lilt-11.3.

Gül, Lale. 2021. *Ik ga leven*. Amsterdam: Prometheus.

Ham, Laurens, Leo Lentz, Henk Pander Maat, and Fabian Stolk. 2018. "Zijn Romans En Kranten Sinds 1950 Eenvoudiger Geworden?" *Tijdschrijft Voor Nederlandse Taal- En Letterkunde* 134 (4): 300–323.

Houwelingen, Bo van. 2021. "In brutale debuten rekenen Sofie Lakmaker en Lale Gül af met de vooroordelen uit hun jeugd." *De Volkskrant*, March 19, 2021. https://www.volkskrant.nl/cultuur-media/in-brutale-debuten-rekenen-sofie-lakmaker-en-lale-gul-af-met-de-vooroordelen-uit-hun-jeugd~bf836190/.

Juola, Patrick. 2008. "Assessing Linguistic Complexity." In *Language Complexity: Typology, Contact, Change*, 89–108. Studies in Language Companion Series 94. Amsterdam/Philadelphia: John Benjamins Publishing Company.

Kézér, Lieke. 2021 "Culturele kloof levert hartstochtelijke literatuur op." *Trouw,* March 13 2021. https://www.trouw.nl/cultuur-media/culturele-kloof-levert-hartstochtelijke-literatuur-op~bba8de96/?referrer=https%3A%2F%2Fwww.google.com%2F.

Kim, Evgeny, and Roman Klinger. "A Survey on Sentiment and Emotion Analysis for Computational Literary Studies," 2018. https://doi.org/10.48550/ARXIV.1808.03137.

Klégr, Aleš, and Pavlína Šaldová. 2017. "Coding Anger in Fiction." *AUC PHILOLOGICA* 2017 (1): 7–18. https://doi.org/10.14712/24646830.2017.1.

Neidhofer, Thilo. 2016. "Popularity and prestige. Margaret Mead and the boundaries of Science." *L'Homme,* vol. 27, no. 2, Oct. 2016, pp. 93+. Gale Academic OneFile, link.gale.com/apps/doc/A591395214/AONE?u=anon~676412cc&sid=googleScholar&xid=73ae5b65.

Pander Maat, H., R. Kraf, A. P. J. van den Bosch, N. Dekker, M. van Gompel, S. de Kleijn, T. Sanders, and K. van der Sloot. 2014. "T-Scan: A New Tool for Analyzing Dutch Text." *Http://Www.Clinjournal.Org/Sites/Default/Files/05-PanderMaat-Etal-CLIN2014.Pdf*. https://repository.ubn.ru.nl/handle/2066/134833.

Pander Maat, Henk, and Nick Dekker. 2016. "Tekstgenres Analyseren Op Lexicale Complexiteit Met T‑Scan." *Tijdschrift Voor Taalbeheersing* 38 (3): 263–304. https://doi.org/10.5117/TVT2016.3.PAND.

Pennebaker, J.W., R.L. Boyd, K. Jordan, and K. Blackburn. 2015. "The Development and Psychometric Properties of LIWC2015." University of Texas at Austin.

Peppelenbos, Coen. 2022. ""Nieuws." *Tzum*, January 27, 2022.
    https://www.tzum.info/2022/01/nieuws-ik-ga-leven-van-lale-gul-best-verkochte-nederlandst
    alige-literair-werk-met-207-000-exemplaren/.

Redactie Wetenschap. 2002. "'Grunberg is Van der Jagt.'" *NRC*, May 11, 2002.
    https://web.archive.org/web/20020607184629/http://www.nrc.nl/nieuws/kunst/1021096
    700145.html.

Sandberg, E. 2017. Hilary Mantel's Wolf Hall(s) and the Circulation of Cultural Prestige. In:
    Kennedy-Karpat, C., Sandberg, E. (eds) *Adaptation, Awards Culture, and the Value of Prestige.
    Palgrave Studies in Adaptation and Visual Culture*. Palgrave Macmillan, Cham.
    https://doi.org/10.1007/978-3-319-52854-0_4.

Tausczik, Yla R., and James W. Pennebaker. 2010. "The Psychological Meaning of Words: LIWC and
    Computerized Text Analysis Methods." *Journal of Language and Social Psychology* 29 (1):
    24–54. https://doi.org/10.1177/0261927X09351676.

Gzip 1.12. 2022. Free Software Foundation. https://www.gzip.org/.