This **Discussion Paper** is an open peer review version that we do not recommend to cite. Submissions that have passed the peer review process are published as full articles on www.jfml.org.

the editors

jfml Journal für Medienlinguistik
Journal for Media Linguistics

Diskussionspapier 5 (2025):
eingereicht am 8.9.2025

Gutachten und Kommentare unter:
https://dp.jfml.org/2025/opr-no-choice/

Discussion Paper

No choice.
On the stylistics of AI-generated texts

Simon Meier-Vieracker

1 Introduction

In public discourse on generative AI, texts written by LLM applications such as ChatGPT are often assessed not only in terms of their informational content but also with regard to their stylistic qualities in the broadest sense. A common observation is that AI generated content is "too perfect […] just eerily smooth"[1] , it is said to be lacking "a distinct voice" and "emotional depth" because it is highly "repetitive" (Aster 2023). According to another statement, AI-generated texts are "not varied enough in form, too smooth and even, sometimes stiff and sometimes too cliché-laden".[2] As vague as these descriptions are, they all refer to linguistic features of texts whose analysis falls within the field of stylistics (Sandig 2006).

The aim of this paper is to sketch out the scope and limitations of stylistics of AI-generated texts as vaguely indicated in the above-mentioned everyday assessments. To date, such stylistics of AI-generated texts have so far only been partially developed. Although an increasing number of empirical studies work with the concept of (writing) style and make use of style-analytical, e. g., stylometric methods, their

1 https://www.linkedin.com/posts/manialok_sometimes-ai-generated-content-is-too-perfect-activity-7308336505846931457-mdqK/

2 „Aber die die Texte selber finde ich zu wenig abwechslungsreich in der Form, zu glatt und gleichmäßig, teilweise auch steif und manchmal zu klischeebeladen." https://www.profi-wissen.de/texte-mit-hilfe-von-ki-generieren-ein-vergleich/

theoretical foundations remain rather vague or reductionist. However, much theoretical work has been done in linguistics on the notion of style in the last decades.

In the following, I would first like to show that it is fruitful to apply concepts from sociolinguistic and pragmatic style theories to the analysis of AI-generated texts, as this highlights the similarities but also the differences between human and AI-generated styles. Secondly, I would like to show that examining the ability of LLMs to write in different styles raises interesting theoretical questions about language and style in general. The paper is theoretical in nature but will refer to empirical data for illustrative purposes.

I will first give an extensive and critical overview to the existing body of research into stylistic properties of AI-generated texts (section 2). I will then introduce a concept of style as meaningful choice as developed and elaborated in interactional sociolinguistics and pragmatic text stylistics (section 3). Against this background, I will report on an experiment in which LLM applications were prompted to write in different styles (section 4) and then point out the differences between stylistic choices in the human sense and probabilistic selections (section 5). Moreover, I will ask why LLMs do perform so well in the task of writing in different styles and will suggest a metapragmatic approach of explanation (section 6).

A note on the state of the art discussed in this article. I will focus on LLMs based on transformer-based architectures and derivative web applications for text generation such as ChatGPT (and will therefore refer to generative AI more precisely as LLMs or LLM-based text generators). Multimodal extensions of these LLMs that enable the generation of audiovisual artefacts will not be considered. Although I am aware that there are (and certainly will continue to be) more advanced ways to use LLM technologies, I would like to focus on what I believe will be the most widespread standard use in the summer of 2025, namely access via a web interface and manual prompting.

2 State of research

An important branch of empirical research into the stylistic properties of AI-generated texts stems from a practical need: detecting texts produced by or with the help of generative AI, particularly in the educational domain. Based on the assumption that AI texts have characteristic stylistic features regardless of their content, researchers have developed and tested different approaches to automatically detect texts with an AI-specific writing style.

For example, Berriche and Larabi-Marie-Sainte (2024) propose a stylometric approach "to detect ChatGPT-based plagiarism", i.e. to make them distinguishable from human written texts. In their study, an author's "writing style" is nothing but a collective term for a broad set of extractable and countable style features like the frequency of different parts-of-speech which indicates attributable authorship. Since they aim at an evaluation of different stylometric methods in the first place, Berriche and Larabi-Marie-Sainte neither focus on linguistic details of the analysed texts, nor do they reflect upon possible the effects and impact of the stylistic features used for the analysis. The same applies for a study by Ma et al. (2023) who use "style features" in terms of word length, function word frequency etc. to train a model that performs a binary classification task. Rivera Soto et al. (2024) make use of so-called style embeddings, a document embedding technique building on style features, to train a detector of generated texts. They find that style embeddings outperform semantic document embeddings not only in distinguishing generated from human-written texts but also to distinct different LLMs which therefore seem to exhibit particular writing styles. However, apart from pointing out that "writing style often comes into focus only after observing a sufficiently-large writing sample", e.g. by the observation of "repeated usage of a rare word […] discriminative of a particular author" (Rivera Soto et al. 2024: 4), they do not give a more detailed definition of style. Moreover, they do not report any concrete stylistic trait of generated texts, let alone a style effect in whatever form.

Slightly more detailed is a study by AlAfnan & MohdZuki (2023) who analyse stylistic features of ChatGPT-generated texts to ask if "artificial intelligence chatbots have a writing style" suitable for detection tasks. They report some quantitative findings about single features like the proportions of active and passive voice but still do not seek to identify interpretable stylistic patterns that could be related to (ethno-)categories of stylistic functions and effects. Opara (2024) brings even more complexity into the matter of LLM-generated content detection by a multi-layered stylometric approach compromising 31 measurable features including, among others, adverb count, emotion word count, and readability scores. They find the measure of unique word count to be most predictive and state "AI's tendency to use rare words excessively" (Opara 2024: 7). Moreover, a relatively high hapax legomena rate, i.e. "the use of words appearing only once [...] signifies rich and detailed vocabulary in human writing" (Opara 2024: 8) which cannot be emulated by AI. A more detailed reference to the functions and effects of these measurable style qualities is still missing.

However, a series of studies that employ the corpus-linguistic approach to style, proposed by Biber (1991) and Biber and Conrad (2009), at least partly fills this gap. In a quantitative, multi-dimensional approach, countable linguistic features are correlated with style axes along different dimensions like involved vs. informative production or situation-dependent vs. elaborated reference. Berber Sardinha (2024) has compared texts from different genres retrieved from the British National Corpus (BNC) on the one hand and ChatGPT-generated texts on the other. Apart from the genre (e.g. , conversation or news article), no additional information was given in the prompt in order to get a most generic response. For example, generated conversations, but also news texts, prove to be less "involved" and more "informational" than their human-authored counterparts (Berber Sardinha 2024: 4; terminology following Biber 1991). Similarly, human-authored texts "exhibit a higher degree of narrativity" (Berber Sardinha 2024: 6) as well as a higher degree of persuasiveness. In reverse direction of analysis, the features measured in the multidimensional analysis also prove as reliable predictors for authorship.

In a similar approach, Markey et al. (2024) have compared students' and ChatGPT's responses to writing assignments by conducting a style analysis across Biber's dimension I (involved vs. informative production) and III (overt vs. nonovert forms of argumentation). Moreover, published texts as examples of professional writing as opposed to the learners' texts were included into the analysis. The results show that LLM-generated responses exhibit the lowest degree of involvement and student responses the highest, while professional texts are in the middle. The same applies for the dimension of overt argumentation with students' responses exhibiting the highest degree and AI-generated responses the lowest. Moreover, all LLM-generated responses show less variance both in measures of standard deviation of the dimensional scores and in terms of repetitiveness in the use of linguistic patterns. In line with this, De Cesare observes in a study of biographic texts generated by ChatGPT in comparison to Wikipedia articles that "there is repetitio over variatio and thus also, more generally, a lack of sensitivity towards stylistic matters" (Cesare 2023: 207).

The mentioned studies make use of concise and static prompts to retrieve a kind of standard response from used LLMs. However, this approach neglects the fact that LLMs are generally able to produce texts in a variety of styles as observed in the training data during the training process. These styles can be specifically retrieved using appropriate prompts. Therefore, Reinhart et al. (2025) use a different research design and build parallel corpora of human-authored and LLM-generated texts, where the former are randomly sampled texts of similar length and of different genres from the Corpus of Contemporary American English. For the LLM corpus, different LLMs were prompted with a chunk of 500 words from the human-authored texts to complete the next 500 words in the "same style, tone, and diction" (Reinhart et al. 2025: 5). The two corpora were then contrasted with regard to the occurrence frequency of selected stylistic features according to Biber and Conrad (2009). The results show that, beyond the stylistic variation due to the variation in the prompt texts, typical stylistic features of generated texts can nevertheless be identified. For example, all analysed LLMs "have strong preferences for present participial clauses, 'that' clauses as subjects, nominalization, and phrasal co-ordination, which are typical markers of more informationally dense, noun-heavy style of writing" (Reinhart et al. 2025: 8). Also, single words like palpable or intricate show surprisingly high frequencies in the

LLM corpus which may produce a recognisable style. Finally, they find notable differences between instruction-tuned and untuned LLMs,[3] showing that some stylistic preferences might be an effect of human preferences during the fine-tuning process.

In a similar, but more sociolinguistic direction points a study by Malik et al. (2024) who instructed different LLMs to write comments to reddit posts in different styles by assigning them specific social personas across different sociodemographic categories. The results show that it is possible to 'personalize' LLMs and to retrieve significant style differences in the responses. With the help of clustering methods and automatic labelling through AI, the authors identify 8 different styles like "cheerful", "simple", "judgemental" etc., but no concrete linguistic features related to these styles are reported in the study. Buz et al. (2024), too, show that LLMs can adapt domain-specific writing styles of Reddit and generate new posts with similar lexical and syntactical profiles. From an art-theoretical perspective, Franzen (2025) diagnoses a "communalization of style" in the age of AI, since individual styles of authors can now easily be reproduced and authors begin to lose authority over their own works.

To conclude this research overview, I will briefly highlight one last type of study from reception research. Gunser et al. (2022) asked 120 participants to rate human-authored and AI-generated continuations of a few lines taken from poems by well-known German poets like Friedrich Hölderlin or Paul Celan according to different aspects of stylistic quality. Participants judged the human-authored continuations as more aesthetic, fascinating, inspiring, interesting, and well-written. They produced similar results when comparing the original poems with AI-generated continuations. Unfortunately, the study does not investigate which linguistic characteristics underlie these categorizations. One should also note that the authors relied on GPT-2, a model that is

3 In untuned LLMs like GPT3, the training is conducted solely on the basis of the training data to only fulfil the task of text completion, whereas instruction-tuned models like ChatGPT "use additional human feedback to optimize the models to follow instructions and answer questions" (Reinhart et al. 2025: 5). Of course, instruction-tuned models can fulfil the task of text completion, too, if they are prompted to do so.

defective in many respects compared to newer ones. Porter and Machery (2024), in contrast, have used ChatGPT to generate poems “in the style of” poets like William Shakespeare and Sylvia Plath and then asked participants to evaluate their poetic quality along dimensions such as beautiful, imagery or inspiring. In this case, participants rated the generated poems slightly better than the authentic ones. However, this study, too, stops short of providing a detailed analysis of the linguistic features that might explain the participants’ subjective ratings.

In this regard, both studies resemble early research on automated journalism (or robot journalism). Clerwall (2014) showed that readers perceived automated texts as more informative but also more boring, while they judged human-authored texts as more pleasant to read. Yet this study, too, makes no attempt to link these judgments to specific linguistic features. In contrast, in my own works (Meier-Vieracker 2023, 2024a) I have analysed a parallel corpus of automated and human-authored football match reports by closely looking at textual features like cohesion, coherence, and narrativity. Since the analysed automated texts were generated by rule-based algorithms with the template-based approach (Diakopoulos 2019), they prove to stand behind their human-authored counterparts in terms of variability, narrativity and suspense.

Although LLM-based text generation is not rule-based anymore and, as shown above, some studies focus on the ability of LLMs to analyze, reproduce and generate writing styles as given by the prompts, most research still builds on a rather reductionist concept of style. Most researchers treat style as a set of (typically countable) linguistic features that warrant the attribution of authorship and sometimes of stylistic labels. When they define the notion of style in more detail, they usually rely on the frequency-based approach of Biber (1991) and Biber and Conrad (2009). What remains largely absent, however, is a deeper praxeological reflection of style as choice which, as I want to argue, can be a fruitful point of comparison for better understanding LLM-based ‘style’.

3 One step back: What is style?

In their work on Register, Genre, and Style, Biber and Conrad (2009) introduce a concept of style that understands style less as a characteristic of texts and more as a perspective on text varieties. What a style perspective has in common with a register perspective on text and text analysis, is a focus on linguistic characteristics which are frequent and pervasive in samples of text excerpts. This goes without any specification of what kind of lexicogrammatical features might be typical for a certain register or style. The authors distinguish between registers and styles as follows: While register "serve important communicative functions" (Biber/Conrad 2009: 16), style "features are not directly functional; they are preferred because they are aesthetically valued" (Biber/Conrad 2009: 16). Style according to Biber and Conrad is basically "influenced by the attitudes of the speaker/writer about language" (Biber/Conrad 2009: 18) and reflects aesthetic preferences. However, stylistic choices are not functionally motivated.

This concept of style is primarily methodological in nature, as it enables a frequency-oriented approach, as outlined in their book, and can guide the interpretation of corpus linguistic results through the conceptual distinction between register and style. However, it falls behind a more interpretative, praxeological approach, as developed in sociolinguistics and in pragmatic stylistics.

In sociolinguistics, style first appeared as a category in the variationist approach of Labov (1966). Style is investigated as a result of intraspeaker variation according to different contexts and activities which still relates to intergroup-variation and the different levels of prestige attributed to group-specific varieties. For example, careful vs. casual speech as different styles in sociolinguistic interviews lead the speakers to use (or avoid) prestigious vs. stigmatized ways of speaking, thus connecting their stylistic activities to their position in a socio-economic hierarchy (Eckert/Rickford 2002: 2). While this approach paints a rather deterministic picture of stylistic variation, later approaches are more action-oriented. For example, Alan Bell in his theory of "language style of audience design" (Bell 1984) considers stylistic variation as derived from intergroup variation. Style "derives its meaning from the association of linguistic features with particular social groups" (Bell 2002: 142) which are evaluated differently. While this is still in line with a Labovian concept of style, Bell puts an additional focus on style-shifting as an adjustment towards the (real or supposed) audience to align with or distance from the addressed social group. Therefore, style serves as a strategic resource for relational work (Locher/Watts 2005) in its broadest sense. Put even more abstractly, style is a matter of (possibly intentional) choice among alternatives (Bell 2002: 139) and therefore a resource for meaning-making.

This view has been further elaborated in interactional sociolinguistics on the one hand and pragmatic text stylistics on the other. In interactional sociolinguistics, which looks at variation as a social practice, style is most generically defined as "a way of doing something" (Coupland 2007: 1) that "marks out or indexes a social difference" (Coupland 2007: 1) and therefore carries meaning. This implies that there are always alternative ways, whereby the specific choice allows for or even provokes interpretative inferences. Methodologically, studies from that paradigm look at sequences of interaction and examine

> the meaningful/significant use of co-occurring linguistic means of expression and formulation for those involved, in comparison to paradigmatic alternatives (which of course never have exactly the same meaning) in the developing interaction situation. (Selting/Hinnenkamp 1989: 5; my translation)

Rather than stylistic variation as a deterministic response to extralinguistic factors, style is a matter of choice that does not only react to, but can actively construct and shape contexts and is used as a contextualization cue (Gumperz 1982):

> 'Style' implies possible alternatives from which choices are actively and always meaningfully made, where necessary in distinction to other possible meaningful choices. (Selting/Hinnenkamp 1989: 7)

This also implies that styles, as meaning-making processes, "result from the interpretation of specific linguistic behaviour in specific language use situations in relation to paradigmatic alternatives that are deemed relevant" (Selting/Hinnenkamp 1989: 6, my emphasis).

Similarly, pragmatic stylistics as part of text linguistics emphasizes the aspect of choice as the general principle of style. Sandig (2006: 9) defines style as the "socially relevant (meaningful) way of performing an action" [sozial relevante (bedeutsame) Art der Handlungsdurchführung]. Generally speaking, style is based on "the meaning-generating [sinnerzeugend] choice between alternatives" (Sandig 2006: 23; cf. also Sanders 1988: 64–66). As in interactional sociolinguistics, this is a context-shaping activity, since styles "can in principle be chosen freely and thus also have an effect on the circumstances in which they are used" (Sandig 2006: 2) by offering guidelines for the interpretation of situational contexts.

The core idea of a conceptual link between choice and meaning, which interactional sociolinguistics and pragmatic stylistics have in common, can be further elaborated with reference to Niklas Luhmann's system-theoretical and phenomenologically based concept of meaning:

> The phenomenon of meaning appears as a surplus of references to other possibilities of experience and action. Something stands in the focal point, at the center of intention, and all else is indicated marginally as the horizon of an "and so forth" of experience and action [...]. The totality of the references presented by a meaningfully intended object offers more to hand than can in fact be actualized at any moment. Thus the form of meaning, through its referential structure, forces the next step, to selection. [...] In a somewhat different formulation, one could say that meaning equips an actual experience or action with redundant possibilities. (Luhmann 1996: 60)

Applied to speaking or writing, this means that what is actually said stands within a 'horizon' of alternatives from which something has been selected. It was said in a certain way, but could have been said differently, and this creates additional meaning for both the speaker and the listener.

With the sociolinguistic and pragmatic concept of style as interpreted yet meaningful and meaning-making choice in mind, I now reconsider the stylistic qualities of AI-generated texts and the stylistic abilities of AI-based text generators. To that end, I carried out an experiment on LLMs rewriting a given text in different styles.

4 LLMs as stylists? An experimental style exercise

4.1 Background and objectives

Older systems of text generation were rule-based algorithms. Apart from some scope for chance at defined points in the process, they were strictly deterministic (Diakopoulos 2019: 99). Thus, the 'stylistic' features of generated texts that various studies have traced are by no means the result of stylistic choices or preferences, but some sort of machine fingerprints. At most, it is the stylistic decisions of the programmers that have been incorporated into the algorithms and are replicated each time they are executed. This causes a static and repetitive quality in the texts. The analysis of these machine fingerprints is still an interesting endeavor. However, it moves most far away from the notion of style that is based in the possibility to express things in different ways and to choose between alternatives in a meaningful and interpretable way.

As shown in the state of research in sec. 2, many style-analytic studies on LLM-generated text still seem to follow the idea of tracing the machine fingerprints of LLMs in a forensic manner. But as already indicated and studied, among others, by Malik et al. (2024), this falls far behind what LLMs can do.

Michael Chollet (2023) has argued that LLMs can be viewed as program databases. Like the much older word2vec models (Mikolov et al. 2013) which allowed to retrieve transformations according to syntactic (singular to plural) or semantic relations (male to female; country to capital), LLMs contain programs to transform input into output which, however, are much more complex. Prompting, then, is the task of searching for the adequate program to process an input. As an example, Chollet cites the program "rewrite in the style of x", which allows, for example, to rewrite poems in the style of Shakespeare.

In 1947, the French poet Raymond Queneau published his book "Exercices de Style" which is based on a similar idea.

An initial narrative text is rewritten in 99 different styles like "metaphorically", "awkward" or "telegraphic". Such a style exercise can now easily be emulated with LLMs.

To this end, I ran a small experiment for the paper at hand. Using three different LLMs (ChatGPT 4o, Claude Sonnet 4 and Google Gemini 2.5 Flash), I wrote a short narrative text and prompted it together with the request to rewrite it in different styles indicated by short labels. The initial text reads as follows:

> In Dresden, a 45-year-old man boards tram line 3 heading towards Coschütz. He has forgotten his wallet, cannot buy a ticket and is promptly checked. After a long discussion with the inspectors, however, he manages to get away with just a warning and does not have to pay a fine. Sweating profusely, the man gets off at Postplatz.

The initial text was designed as a largely neutral and concise documentation of the reported events (following the example of Queneau and as a reverence to his work, I decided to let the story take place in public transport). Of course, this text includes some stylistic choices, too, and should not be misunderstood as a non-stylised template. However, some point of departure is needed.

The style labels that I used in prompts like "Rewrite this text in a ... style" include the following adjectives referring to stylistic qualities: formal, stilted, florid, ornate, emotive, clumsy, concise, conversational and crude. Additionally, I used two adjectives that refer to registers or, in structuralist terminology, functional styles: academic and officialese. Finally, two genre labels were used: stand-up comedy and tabloid. Admittedly, these labels are rather heterogeneous (as in Queneau's work, too) and refer to different levels of linguistic variation. Unlike the first group of labels, register or genre names do not specify stylistic qualities in the narrow sense. However, they refer to types of language use that can be expected to exhibit certain and relatively uniform stylistic qualities. Through queries in the English web corpus enTenTen21 as part of SketchEngine, I have checked all labels used in the prompts to ensure that they correspond to common language use. That is, it was ensured that formulations like tabloid style or clumsy style are frequently used in contemporary English.

Admittedly, this procedure moves away from the idea of social styles as developed in sociolinguistics. In a more sociolinguistically inspired approach, types of social personae (e.g., according to sociodemographic categories or social groups) or types of social situations could have been described in the prompts to see which style the LLM would use (Malik/Jiang/Chai 2024). For reasons of simplicity and controllability, however, the mentioned labels were preferred which directly designate the styles to be generated.

4.2 Results

All three models easily and mostly adequately fulfilled the task of rewriting the given text in the prompted styles. All the resulting texts can be seen in the digital appendix.[4] While Claude Sonnet 4 and Gemini Flash 2.5 simply gave back the texts, preceded by a short line like ”Here’s the text rewritten in conversational style”, ChatGPT 4o added a short characterization of that style. For example, the text written in “officialese” was described as “formal, bureaucratic, and filled with jargon and passive constructions”. Moreover, ChatGPT 4o made suggestions as to what other styles the text could be rewritten in, e.g. “Want to go surreal next? Or something deadpan, poetic, noir...?”. This suggests that the game of playful style-shifting is recognized by and therefore represented in this model.

The models’ texts significantly differ in length. At least the texts of ChatGPT tend to be shorter than the texts of Claude, while the length of Gemini's texts varies the most (Fig. 1).

4 https://osf.io/yjnb5/

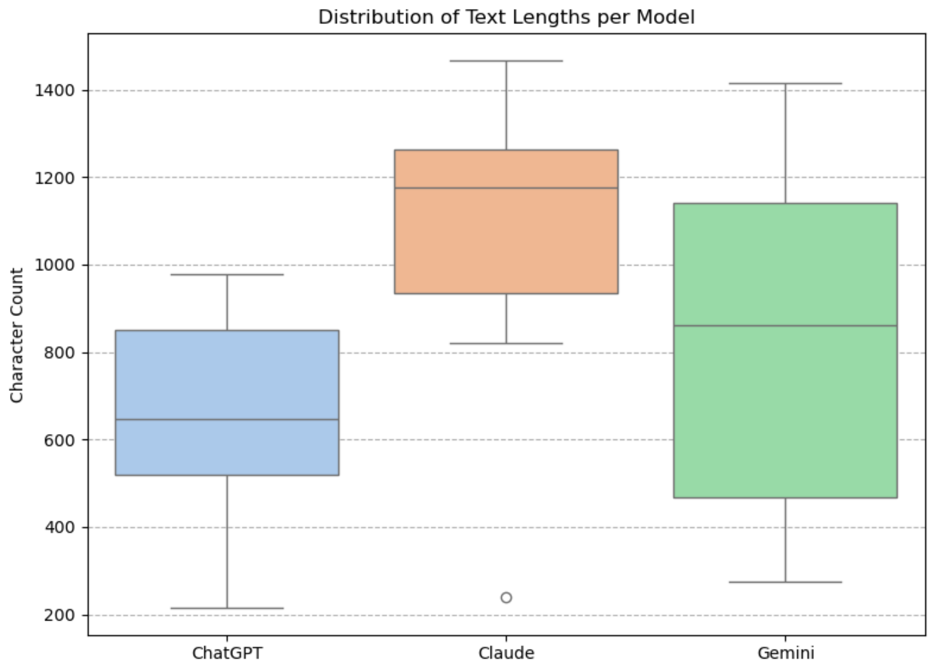

Fig. 1: Text lengths per model

To give an impression of the outputs, the texts written in a "conversational" style shall be quoted at length:

> So, there's this 45-year-old guy in Dresden, right? He hops on tram line 3, heading to Coschütz, probably thinking about a million other things — and then it hits him. He forgot his wallet. No ticket, no way to pay, nothing. And of course, that's exactly when the ticket inspectors show up. You can imagine the panic. He tries to explain, they go back and forth for a while — not angry, but definitely firm. It drags on a bit, but somehow, he gets lucky. They just give him a warning. No fine, no big scene. He gets off at Postplatz, sweating like crazy. Whether it was the stress or the close call, who knows. But yeah — could've gone way worse. (ChatGPT 4o)

> So there's this guy, right? He's 45, gets on the tram in Dresden - you know, line 3 that goes to Coschütz. And wouldn't you know it, he realizes he's forgotten his wallet! I mean, we've all been there, haven't we? You're rushing out the door and you just... forget. So he's standing there thinking "Oh crap, how am I gonna buy a ticket?" And of course, that's exactly when the ticket inspectors show up. It's like they have a sixth sense for this stuff, I swear. Anyway, they start talking to him about it, and you can imagine how that conversation went. He's probably explaining the whole situation - "Look, I'm really sorry, I

> just forgot my wallet this morning...” - and they’re doing their job, checking if he’s telling the truth. But here's the thing - he actually managed to talk his way out of it! Got away with just a warning instead of having to pay the fine. Pretty lucky, if you ask me. Those inspectors must have been feeling generous that day. By the time he got off at Postplatz, the poor guy was sweating buckets. I bet his heart was pounding the whole time! Nothing like a little scare on public transport to get your adrenaline going, right? (Claude Sonnet 4)

> So, imagine this: you’re in Dresden, and there’s this guy, about 45, just trying to get on tram line 3 to Coschütz. Totally normal, right? Except, oops, he realizes he’s forgotten his wallet. You know that sinking feeling? Like, “Oh no, how am I going to buy a ticket?” And of course, just like in a movie, the inspectors show up! So he gets checked immediately. I can just picture the whole thing: a long discussion with these serious-looking inspectors. You can bet he was trying every excuse in the book! The poor guy was apparently sweating profusely – seriously stressed out, I’m sure. But get this: he actually managed to talk his way out of a fine! They just gave him a warning. Can you believe it? He must have been really good at explaining himself. Finally, he gets off at Postplatz, probably just wanting to disappear. What a dramatic tram ride, huh? (Gemini Flash 2.5)

The examples show that the style label “conversational” carries enough information for the models to generate texts that differ greatly from the original text but are similar to each other because of common stylistic traits at the pragmatic, lexical, morphosyntactic and syntactic level. To begin with the pragmatic level, all three texts start with the discourse marker *so* typical for oral narratives (Bolden 2009).[5] Tag questions like *right* or *huh* which elicit some listener’s response (Erman 2001) and direct addresses of the listener as in *you can imagine* or *you know* consistently indicate a dialogical speech situation throughout the texts. Interjections like *oops* and *yeah* as well as exclamative constructions like *what a dramatic tram ride* (Ziem/Ellsworth 2015) indicate a high degree of emotional engagement (Caffi/Janney 1994). The texts by Claude and Gemini both enrich the narratives

5 The same observation holds for the stand-up comedy styled texts.

through reported thought and speech as a very common means of displaying affective stance in oral narrative (Günthner 1999).

On a lexical level, the neutral noun man is replaced by the more colloquial *guy*, as *to board the* tram is replaced by *to hop on the tram* or *to get on the tram*. Instead of *sweating profusely*, ChatGPT and Claude use the more expressive and figurative phrases *sweating like crazy* and *sweating buckets*. On a morphosyntactic level, clitics like *could've*, *we've* or *he's* can be found in all three texts. Finally, there are some common features between the texts at the syntactic level. For example, many instances of verbless clauses can be found: *No ticket, no way to pay, nothing […] No fine, no big scene* (ChatGPT); *Nothing like a little scare on public transport* (Claude); *What a dramatic turn ride, huh?* (Gemini). Also, anacolutha typical for spoken language can be found: *Except, oops, he realizes he's forgotten his wallet* (Gemini).

As the example shows, the conversational style generated by the various LLMs differs systematically from the original text. It has common features that correspond to what has been widely studied in conversation analysis and interactional linguistics. The same can also be demonstrated for the other styles. In the "formal" style, *to be checked* is replaced by *to be subjected to a ticket inspection* or even, from the inspectors' perspective, to *conduct their routine examination of passengers*. The neutral noun *man* is replaced by the even more objective technical term *male individual*, whereas in the "ornate" style it is replaced by *gentleman*. The "officialese" style is characterized by passive constructions like *it was adjucated that the individual would be issued a formal warning* (ChatGPT), *a formal verbal warning was issued* (Claude), *a determination was made to issue a formal warning* (Gemini). Even on a narrational level, the models use similar linguistic means. In the "emotive" style, for example, the turning point (Langenhorst/Schuppe/Frommherz 2024) of the story, i.e., the moment when the protagonist realizes that he has forgotten his wallet, is indicated by syntactic disfluency. It is typographically supported by hyphens and seems to symbolize the moment of surprise and confusion:

> It isn’t until the doors close behind him that he realizes – his wallet is gone. (ChatGPT)

> But then – oh God, the sickening realization! His wallet, his lifeline, abandoned somewhere in the chaos of his morning routine. (Claude)

> Then, a cold sickening lurch in his stomach – his wallet, gone. (Gemini)

The finding that the LLM’s texts show similar features for the different styles can further be supported by a stylometric cluster analysis (Eder/Rybicki/Kestemont 2016). This contrastive and quantitative method is not particularly well suited to identifying interpretable stylistic features. Rather, it serves to group texts according to the distribution of linguistic patterns that are “frequent and pervasive” (Biber/Conrad 2009: 16) across texts and thus may represent distinguishable styles. A comparison of the 100 most common character trigrams, presented as a dendrogram with every leaf representing a single text, yields the following result (Fig. 2). Leaves belonging to the same branch (at different levels of abstraction) are found to be stylistically similar. To make the dendrogram easier to read, texts of the same style are displayed in the same colours.

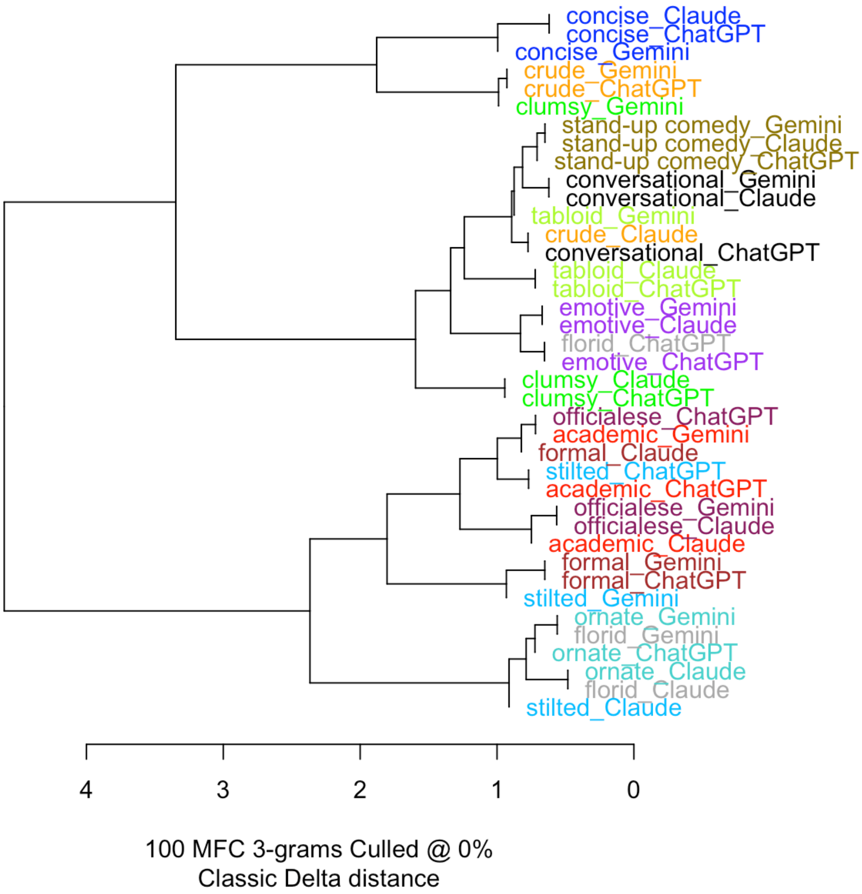

Fig. 2: Stylometric cluster analysis

As the dendrogram with its four main clusters shows, the texts in the different styles are mostly grouped together even if they come from different LLMs. Moreover, the styles as such appear to be grouped in a plausible manner: One the one side, “stilted”, “ornate” and “florid” texts are grouped together and distinguished from “formal”, “academic” and “officialese” texts. On the other side, “stand-up comedy”, “conversational”, “tabloid”, “emotive” and “clumsy” texts are grouped together and distinguished from “crude” and “concise” texts. One possible explanation for this could be that text properties like lexical and syntactic elaboration vs. signs of spontaneity and emotionality, which are reflected in the frequencies of character trigrams, were correctly recognized by the clustering algorithm. For example, the trigram “i o n” which serves as a nominalization suffix is most frequent in the “officialese”, “formal” and “academic” texts as in “On the occas**ion** of his utilizat**ion** of public transportat**ion** services within the jurisdict**ion** of Dresden” (ChatGPT). On the contrary, the trigram “i n g” used for the formation of present participles and gerunds is most frequent in the "conversational" texts as in "and they're do**ing** their job, check**ing** if he's tell**ing** the truth" (Claude). From the perspective of text generation, this means that all these style features have been generated by the LLMs in a consistent manner beforehand.

4.3 Conclusions

Three conclusions can be drawn from this experiment.

1. LLMs have remarkable abilities to generate texts in different styles. If they are prompted to do so, they can do (roughly) the same thing, i.e., telling a story, in many different ways (Coupland 2007). Therefore, studies that ask about the genuine writing style of LLMs are far too simplistic and cannot take into account the diversity of styles that are represented in the models and can also be retrieved. There may be a default style of AI-generated texts that are prompted without further specification, but this can easily be changed. Different from rule-based systems, LLMs show great flexibility.
2. Across different LLMs, texts generated it different styles share common features and consistently correlate to everyday language style labels.
3. At least retrospectively, the task of rewriting can be conceived as a series of replacements and transformations of various linguistic items. When viewed together, the items involved appear as "paradigmatic alternatives (Selting/Hinnenkamp 1989: 5) as in the set [*man*, *male individual*, *guy*, *gentleman*, *dude]*.

Taken together, one could think of LLMs to be competent stylists. However, I think there is still a significant gap between writing styles in the sense of human's language use on the one hand and LLMs writing styles on the other.

5 No choice

As introduced above, the core principle of style from a sociolinguistic and pragmatic perspective is that of choice, where choice is a meaningful and meaning-generating [sinnerzeugend] process (Sandig 2006: 23). Put even more abstractly, this can be linked to Niklas Luhmann's concept of meaning as a "surplus of references to other possibilities of experience and action" (Luhmann 1996: 60) which is still present in the intended object after selection.

As far as we know, the text generating algorithms based on LLMs are not capable of this. Instead, they predict the next word in a sequence based on the given context. This prediction is made using vector representations of the input, where each word and its surrounding context are mapped into a high-dimensional space. A key component in this process is the attention mechanism, which assigns greater weight to contextually relevant words, allowing the model to focus on important parts of the input. Based on these representations, the model assigns probabilities to all possible next words, reflecting patterns it learned during training. Finally, depending on the temperature parameter (which controls the randomness of the output), one of the high-probability words is selected (Wolfram 2023). In this process, some aspects of meaning as semantic relations and semantic similarity are captured on the basis of co-textual patterns which is sufficient for generating semantically coherent texts (Bender/Koller 2020: 5193). But this type of meaning is, to use a term coined by Bajohr, "dumb meaning [...] without any indexical relation to the world" (Bajohr 2023: 58) which is "'parasitically' dependent on a human interpreter" (Bajohr 2023: 58).

As a probabilistic device, an LLM-based text generator is not strictly deterministic, but it is still mechanistic. In other words: The text generator does select high probability words but still has no choice (not) to do so. Furthermore, there is no reason to assume that the text generator has a "horizon of an 'and so forth' of action and experience" (Luhmann 1996: 60) to accomplish its task. The significance of the LLM's probability-based selections does not go beyond dumb meaning in the sense of Bajohr. For users who can ask, say, ChatGPT to write in a certain style, it may seem as if the machine has a choice that human interpreters can make sense of, but it only chooses on demand and according to the users' specifications.

6 In the thicket of probabilities

But why, then, do LLM-based text generators succeed that well in (re-)writing in different styles if they have no choice in the full sense of the word? To clarify this question, it is worth taking a look back to the theory of style in the framework of generative grammar in the tradition of Noam Chomsky (1965). As Rosengren (1972) shows in his paper "Style as Choice and Deviation", also generativism has developed a theory of style as choice which can be reconstructed as follows: While linguistic competence is the ability to generate sentences according to grammatical rules of the language system, these rules do not conclusively determine how exactly sentences are formulated. There is some freedom for choice between alternative expressions, but this is not part of the competence but a matter of performance. According to Rosengren, this style-forming process is governed by rules, too, but these rules, which he refers to as "stylistic performance rules" (Rosengren 1972: 4), are metarules that regulate how to use the rules of grammar.

Different from grammar rules which are of general validity, stylistic performance rules are idiosyncratic, that is specific to group, occasion, or author. Moreover, Rosengren conceives the stylistic performance rules as probabilistic since in concrete styles the distinctive style features will occur with certain probabilities. A concrete style is thus seen as a "system of probabilities" (Rosengren 1972: 9), where the probabilities of multiple style features are interdependent. This is primarily intended as an analytical tool: The overall probabilities with which an author or text prefers particular formulations over other alternatives then constitutes the stylistic profile of an author or text. In fact, digital stylometry is based on precisely this idea (Horstmann 2018). But this has a generative side as well, as the knowledge of these probabilities can be used to generate texts, say, in the style of Shakespeare.

In the age of Large Language Models, the idea of a 'system of probabilities' as part of a generative process seems compatible at first glance, since LLMs appear to be precisely that: systems of probabilities. However, there is a crucial difference. According to the traditional idea of generative grammar, there is a clear division of labour between the transformational rules of grammar on the one hand and the stylistic performance rules on the other. Within this approach, an automatic generation of texts would follow a two-step procedure: First, the transformation rules would translate syntactic deep structures into surface structures of grammatically acceptable sentences. Then, within the range of grammaticality alone, the stylistic performance rules would regulate the choices of alternative formulations according to certain probabilities. But these probabilities only apply at the stylistic level and not on the level of grammar, because "[t]he [language] system itself possesses no probabilities" (Rosengren 1972: 14).

As Bubenhofer (2024) has argued, LLMs are thoroughly dispelling this idea. While older approaches to text generation using rule-based algorithms were based in some ways on ideas from generative grammar, newer systems rely exclusively on probabilities and their statistical modelling but still work much better. LLM-based text generators do not have and do not need any knowledge of grammatical rules in order to generate grammatically correct sentences (Wolfram 2023). They have no 'competence' in the traditional sense of the term, but from observing and modelling performance and its multifaceted patterns alone, LLMs have acquired the capacity to generate new sentences and even texts. Instead of being a system of abstract syntax, language appears as mere performance with "idiomacity on all its shades" (Hausmann 2008: 7) that can be statistically modelled as cooccurrence probabilities (Meier-Vieracker 2024b: 136).

In sec. 3, I have introduced the interactional sociolinguistic concept of style of interpreted and socially meaningful choice. As Selting and Hinnenkamp (1989: 6) argue, styles in this sense are "holistic communicative signs" that do not work as subsequent add-ons to grammar and lexis but rather permeate all levels of language use. LLMs, with their ability not only to generate grammatically correct sentences, but also to generate texts in various styles, provide strong evidence for this. LLMs do not interpret as humans do but process chunks of tokenized language by their transformer-based architectures. However, in the models' highly complex representations of linguistic patterns not only syntactic patterns but also styles (as well pragmatic and textual functions and other linguistic features) are apparently represented. From the perspective of LLMs, the boundaries between grammar and style are completely blurred.

The ease with which LLMs evoke different styles through targeted prompting likely stems from a phenomenon that interactional sociolinguistics has described in detail: language users themselves typify and categorize styles through metapragmatic references (Selting/Hinnenkamp 1989: 7). In everyday discourse, speakers use style labels such as "florid" or "conversational" – including those applied in the experiment discussed in Section 4 – and combine them with stereotypical descriptions and evaluations (Sandig 2006: 3). This co-occurs with linguistic patterns that can be statistically modelled during the LLMs' training process and subsequently applied in the generation of new stretches of text in these styles.[6]

Ultimately, what Schneider and Zweig (2022: 285) have pointed out about transformer-based translation tools like DeepL also applies here. These tools deliver valid translations with great sensitivity to culturally significant nuances, as the training data consists of culturally anchored translations by humans whose orientation towards these cultural nuances is also captured during training. In a later work on ChatGPT, Schneider has coined the term of "intelligible textures" as "semiotic configurations that can be read and interpreted as intelligent texts" (Schneider 2024: 15), because the LLM has been trained on intelligent texts by humans. Applied to the

6 While working on the experiment, I did some tests with the LLM based search engine you.com. Although the generated texts in different styles do not show the same quality and variation as those of ChatGPT, Claude and Gemini (and were therefore not included in the analysis), you.com offered sources that its outputs rely on. For the "stilted" style, for example, it referred to a post on Reddit "How do I Improve From Stilted to Flowing Writing" which is introduced as follows: "'The person had red hair' - Me vs 'Beta saw a splash of brilliant color above her.' - A friend of mine" (https://www.reddit.com/r/writing/comments/tyvboq/how_do_i_improve_from_stilted_to_flowing_writing/). Examples like these will appear en masse in the training data of LLMs, from which stylistic patterns can be learned without having to explicitly define stylistic rules.

topic of style, this means that LLMs appear to be stylists because humans in their language use are permanently engaged with metapragmatic categorizations of styles which are then represented in the LLMs as well and reactivated when prompted to write in these styles. To freely rephrase a quote of Asif Agha on register, which, however, can also be applied to style:

> [Large Language Models] rely on the metalinguistic ability of native speakers to discriminate between linguistic forms, to make evaluative judgments about variant forms [...] that are overtly expressed in publicly observable semiotic behavior. (Agha 1999: 216)

The constitutive role that metapragmatics plays in language use of humans is indirectly demonstrated by the fact that it is also the key to LLMs' ability to perform as stylists as well as they obviously do.

7 Concluding remarks

In this paper, I have presented some principles of what might be considered stylistics of AI-generated texts. Unlike most scholarly publications on the writing style of LLM-based applications, which use a highly reductionist concept of style suitable for quantitative approaches to authorship attribution or similar, I have drawn on a praxeological concept of style as a socially meaningful choice, as developed in interactional sociolinguistics and pragmatic text stylistics. In an experiment with three well-known LLM applications, I demonstrated that they are capable of consistently (re)writing texts in different styles. Nevertheless, I argued that there is a fundamental gap between the process of selecting the next word and stylistic choice in the human sense. Finally, I have discussed possible explanations why LLMs do perform so well in the task of writing in different styles and have pointed out the crucial role of metapragmatics in a consequently performance-oriented understanding of language.

To shortly summarize the main points of my argument: LLM-based text generators have the ability of doing (roughly) the same thing in different ways which is the core principle of style. But still, they have no choice as humans do, but humans have the choice to make them write in different and interpretable styles. This is because humans' stylistic choices, including their metapragmatic typifications and categorizations, are represented as complex patterns in LLMs.

References

Agha, Asif (1999): Register. In: Journal of Linguistic Anthropology 9 (1–2), 216–219. DOI: 10.1525/jlin.1999.9.1-2.216.

AlAfnan, Mohammad Awad/MohdZuki, Siti Fatimah (2023): Do Artificial Intelligence Chatbots Have a Writing Style? An Investigation into the Stylistic Features of ChatGPT-4. In: Jour-nal of Artificial Intelligence and Technology 3, DOI: 10.37965/jait.2023.0267.

Aster, Ruben (2023): Your generated Articles smell like AI — Let's fix that and pass the Detec-tion Tests! https://medium.com/@Ruben.Aster/your-generated-articles-smell-like-ai-lets-fix-that-b159ef7749b7.

Bajohr, Hannes (2023): Dumb Meaning: Machine Learning and Artificial Semantics. In: IMAGE 37 (1), 58–70. DOI: 10.1453/1614-0885-1-2023-15452.

Bell, Allan (1984): Language Style as Audience Design. In: Language in Society 13 (2), 145–204.

Bell, Allan (2002): Back in style: reworking audience design. In: Eckert, Penelope/Rickford, John R (ed.): Style and Sociolinguistic Variation. Cambridge: Cambridge University Press, 139–169.

Bender, Emily M./Koller, Alexander (2020): Climbing towards NLU: On Meaning, Form, and Understanding in the Age of Data. In: Online: Association for Computational Linguistics. 5185–5198. https://aclanthology.org/2020.acl-main.463.

Berber Sardinha, Tony (2024): AI-generated vs human-authored texts: A multidimensional com-parison. In: Applied Corpus Linguistics 4 (1), 100083. DOI: 10.1016/j.acorp.2023.100083.

Berriche, Lamia/Larabi-Marie-Sainte, Souad (2024): Unveiling ChatGPT text using writing style. In: Heliyon 10 (12), DOI: 10.1016/j.heliyon.2024.e32976.

Biber, Douglas (1991): Variation Across Speech and Writing. Cambridge University Press.

Biber, Douglas/Conrad, Susan (2009): Register, Genre, and Style. Cambridge: Cambridge Univer-sity Press.

Bolden, Galina B. (2009): Implementing incipient actions: The discourse marker 'so' in English conversation. In: Journal of Pragmatics 41 (5), 974–998. DOI: 10.1016/j.pragma.2008.10.004.

Bubenhofer, Noah (2024): Grammatik. In: Zeitschrift für Medienwissenschaft 16 (30), 53–55. DOI: 10.25969/MEDIAREP/21973.

Buz, Tolga/Frost, Benjamin/Genchev, Nikola/Schneider, Moritz/Kaffee, Lucie-Aimée/de Melo, Gerard (2024): Investigating Wit, Creativity, and Detectability of Large Language Models in Domain-Specific Writing Style Adaptation of Reddit's Showerthoughts. https://arxiv.org/abs/2405.01660.

Caffi, Claudia/Janney, Richard W. (1994): Toward a pragmatics of emotive communication. In: Journal of Pragmatics 22 (3), 325–373. DOI: 10.1016/0378-2166(94)90115-5.

Cesare, Anna-Maria De (2023): Assessing the quality of ChatGPT's generated output in light of human-written texts: A corpus study based on textual parameters. In: CHIMERA: Revista de Corpus de Lenguas Romances y Estudios Lingüísticos 10, 179–210.

Chollet, François (2023): How I think about LLM prompt engineering. https://fchollet.substack.com/p/how-i-think-about-llm-prompt-engineering.

Chomsky, Noam (1965): Aspects of the Theory of Syntax. Cambridge, Mass: MIT Press.

Clerwall, Christer (2014): Enter the Robot Journalist. In: Journalism Practice 8 (5), 519–531. DOI: 10.1080/17512786.2014.883116.

Coupland, Nikolas (2007): Style: language variation and identity. Cambridge: Cambridge Uni-versity Press.

Diakopoulos, Nicholas (2019): Automating the News: How Algorithms Are Rewriting the Media. Harvard University Press.

Eckert, Penelope/Rickford, John R (2002): Introduction. In: Eckert, Penelope/Rickford, John R (Hg.): Style and

Sociolinguistic Variation. Cambridge: Cambridge University Press.

Eder, Maciej/Rybicki, Jan/Kestemont, Mike (2016): Stylometry with R: A Package for Computa-tional Text Analysis. In: The R Journal 8 (1), 107–121.

Erman, Britt (2001): Pragmatic markers revisited with a focus on you know in adult and adole-scent talk. In: Journal of Pragmatics 33 (9), 1337–1359. DOI: 10.1016/S0378-2166(00)00066-7.

Franzen, Johannes (2025): Vergemeinschaftung des Stils? Werkherrschaft im Zeitalter Künstli-cher Intelligenz. In: Merkur 914, 74–82.

Gumperz, John J. (1982): Discourse strategies. Cambridge, New York: Cambridge University Press (Studies in interactional sociolinguistics, 1).

Gunser, Vivian Emily/Gottschling, Steffen/Brucker, Birgit/Richter, Sandra/Çakir, Dîlan/Gerjets, Peter (2022): The Pure Poet: How Good is the Subjective Credibility and Stylistic Quality of Literary Short Texts Written with an Artificial Intelligence Tool as Compared to Texts Written by Human Authors? In: Proceedings of the Annual Meeting of the Cognitive Sci-ence Society 44 (44), URL: https://escholarship.org/uc/item/1wx3983m.

Günthner, Susanne (1999): Polyphony and the 'layering of voices' in reported dialogues: An analysis of the use of prosodic devices in everyday reported speech. In: Journal of Pragmatics 31 (5), 685–708. DOI: 10.1016/S0378-2166(98)00093-9.

Hausmann, Franz Josef (2008): Kollokationen und darüber hinaus. Einleitung in den themati-schen Teil „Kollokationen in der europäischen Lexikographie und Wörterbuchfor-schung". In: Lexicographica 24 (2008), 1–8. DOI: 10.1515/9783484605336.1.1.

Horstmann, Jan (2018): Stilometrie. In: https://fortext.net/routinen/methoden/stilometrie.

Labov, William (1966): The social stratification of English in New York City. Washington, DC: Center for Applied Linguistics.

Langenhorst, Jan/Schuppe, Robert Cornelis/Frommherz, Yannick (2024): And then I saw it: Tes-ting Hypotheses on Turning Points in a Corpus of UFO Sighting Reports. In:

5th Con-ference on Computational Humanities Research 2024. https://ceur-ws.org/Vol-3834/paper93.pdf.

Locher, Miriam A./Watts, Richard J. (2005): Politeness Theory and Relational Work. In: Journal of Politeness Research. Language, Behaviour, Culture 1 (1), 9–33. DOI: 10.1515/jplr.2005.1.1.9.

Luhmann, Niklas (1996): Social Systems. Stanford, Calif: Stanford University Press.

Malik, Manuj/Jiang, Jing/Chai, Kian Ming A. (2024): An Empirical Analysis of the Writing Styles of Persona-Assigned LLMs. In: Proceedings of the 2024 Conference on Empirical Methods in Natural Language Processing, 19369–19388. https://aclanthology.org/2024.emnlp-main.1079.

Ma, Yongqiang/Liu, Jiawei/Yi, Fan/Cheng, Qikai/Huang, Yong/Lu, Wei/Liu, Xiaozhong (2023): AI vs. Human -- Differentiation Analysis of Scientific Content Generation. arXiv. http://arxiv.org/abs/2301.10416.

Markey, Ben/Brown, David West/Laudenbach, Michael/Kohler, Alan (2024): Dense and Dis-connected: Analyzing the Sedimented Style of ChatGPT-Generated Text at Scale. In: Writ-ten Communication 41 (4), 571–600. DOI: 10.1177/07410883241263528.

Meier-Vieracker, Simon (2023): Von Rohdaten zum Text – Themenentfaltung in automatisierten Fußballspielberichten. In: Engelken, Julian/Glund, Marc/Hensellek, Jan/Herford, La-ra/Langrock, Saskia/Poghosyan, Sargis/Schmalwieser, Susanne S./Warnke, Ingo H. (ed.): Was ist eigentlich ein Thema? Sieben linguistische Perspektiven. Bremen: Universität Bremen, 50–58. DOI: 10.26092/elib/2313.

Meier-Vieracker, Simon (2024a): Automated football match reports as models of textuality. In: Text & Talk 45 (1), 1–24. DOI: 10.1515/text-2022-0173.

Meier-Vieracker, Simon (2024b): Uncreative Academic Writing: Sprachtheoretische Überlegun-gen zu Künstlicher Intelligenz in der akademischen Textproduktion. In: Schreiber, Gerhard/Ohly, Lukas (ed.): KI:Text. Diskurse über KI-Textgeneratoren. Berlin, Boston: De Gruyter, 133–144. DOI: 10.1515/9783111351490-010.

Mikolov, Tomas/Chen, Kai/Corrado, Greg/Dean, Jeffrey (2013): Efficient Estimation of Word Representations in

Vector Space. In: arXiv:1301.3781 [cs], URL: http://arxiv.org/abs/1301.3781.

Opara, Chidimma (2024): StyloAI: Distinguishing AI-Generated Content with Stylometric Analy-sis. arXiv. http://arxiv.org/abs/2405.10129.

Porter, Brian/Machery, Edouard (2024): AI-generated poetry is indistinguishable from human-written poetry and is rated more favorably. In: Scientific Reports 14 (1), 26133. DOI: 10.1038/s41598-024-76900-1.

Reinhart, Alex/Markey, Ben/Laudenbach, Michael/Pantusen, Kachatad/Yurko, Ronald/Weinberg, Gordon/Brown, David West (2025): Do LLMs write like humans? Variation in grammati-cal and rhetorical styles. In: Proceedings of the National Academy of Sciences 122 (8), e2422455122. DOI: 10.1073/pnas.2422455122.

Rivera Soto, Rafael/Koch, Kailin/Khan, Aleem/Chen, Barry/Bishop, Marcus/Andrews, Nicholas (2024): Few-Shot Detection of Machine-Generated Text using Style Representations. ar-Xiv. https://arxiv.org/abs/2401.06712.

Rosengren, Inger (1972): Style as Choice and Deviation. In: Style 6 (1), 3–18.

Sanders, Willy (1988): Stil und Spracheffizienz. Zugleich Anmerkungen zur heutigen Stilistik. In: Rhetorik (1988) 7 (1988), 63–78. DOI: 10.1515/9783110244540.63.

Sandig, Barbara (2006): Textstilistik des Deutschen. Berlin, New York: De Gruyter.

Schneider, Jan Georg (2024): Intelligible Texturen. Welche Rolle kann ChatGPT bei der Aufsatz-bewertung spielen? URL: https://zenodo.org/records/10877034.

Schneider, Jan Georg/Zweig, Katharina A. (2022): Ohne Sinn. Zu Anspruch und Wirklichkeit automatisierter Aufsatzbewertung. In: Brommer, Sarah/Roth, Kersten Sven/Spitzmüller, Jürgen (ed.): Brückenschläge: Linguistik an den Schnittstellen. Tübingen: Narr Francke Attempto (Tübinger Beiträge zur Linguistik, 583), 271–294.

Selting, Margret/Hinnenkamp, Volker (1989): Einleitung: Stil und Stilisierung in der Interpretati-ven Soziolinguistik. In: Selting, Margret/Hinnenkamp (ed.): Stil und Stilisierung. Tübin-gen: Niemeyer, 1–24. DOI: 10.1515/9783111354637.1.

Wolfram, Stephen (2023): What is ChatGPT Doing ... and why does it work?

https://writings.stephenwolfram.com/2023/02/what-is-chatgpt-doing-and-why-does-it-work/.

Ziem, Alexander/Ellsworth, Michael (2015): Exklamativsätze im FrameNet-Konstruktikon am Beispiel des Englischen. In: Finkbeiner, Rita/Meibauer, Jörg (ed.): Satztypen und Kon-struktionen. Berlin, New York: De Gruyter, 146–191.