

WEIZENBAUM JOURNAL OF THE DIGITAL SOCIETY  
Volume 4 \ Issue 4 \ w4.4.6 \ 06-24-2024  
ISSN 2748-5625 \ DOI 10.34669/WI.WJDS/4.4.6

Information on this journal and its funding can be found on its website:  
<https://wjds.weizenbaum-institut.de>

This work is available open access and is licensed under Creative Commons Attribution 4.0 (CC BY 4.0):  
<https://creativecommons.org/licenses/by/4.0/>

---

KEYWORDS

text genre  
text genre competence  
AI  
ChatGPT  
writing didactics

RESEARCH PAPER

# ChatGPT and Its Genre Competence

## An Exploratory Study

Sarah Brommer<sup>\*1</sup> \ Karina Frick<sup>\*2</sup> \ Adriana Bursch<sup>1</sup> \  
Marina Rodrigues Crespo<sup>1</sup> \ Laura Katrin Schwerdtfeger<sup>1</sup>

<sup>1</sup>University of Bremen

<sup>2</sup>Leuphana University Lüneburg

\*Main authors, contact: [sarah.brommer@uni-bremen.de](mailto:sarah.brommer@uni-bremen.de),  
[karina.frick@leuphana.de](mailto:karina.frick@leuphana.de)

---

ABSTRACT

Being able to reciprocate and produce different kinds of texts is a key quality and a core professional competence. Therefore, genre competence is fundamental in not only educational and academic contexts but also professional environments. This paper addresses the extent to which text-generating AI tools could support the development of genre competence and how suitable they are as a tool for genre-based writing didactics. To answer this, it is necessary to examine whether AI tools, such as ChatGPT, are competent in terms of text genres. To do this, the re-search explores whether ChatGPT is capable of producing and revising genre-specific texts or identifying and analyzing genre-specific patterns and whether it produces different outputs in terms of genre. To examine these questions, we have conducted a pilot study that includes several different text types and several areas of application (generating, revising, summarizing, classifying, and analyzing). The paper's results relate to three aspects of "Education in the Digital World": a) competencies, b)

possible changes to educational and learning processes using AI tools, and c) appropriate tools for education in general.

## 1 Introduction

Being able to produce, receive, and reciprocate different kinds of texts is a key quality and a core professional language competence (Göperich, 2015; Nünning, 2008). Therefore, genre competence and its acquisition are fundamental in not only educational and academic contexts (i.e., school and university; Esterl & Krieg-Holz, 2018; Freudenberg-Findeisen, 2016; Steinhoff, 2007) but also professional environments (Trosborg, 1997; Fandrych & Thurmair, 2011). Genre competence describes the knowledge and competent handling of the typological rules necessary for the successful and appropriate production and reception of different textual genres. Importantly, what a text of a certain genre typically looks like serves as a benchmark and orientation for not only human language users but also text-generating artificial intelligence (AI), with large language models (LLMs) such as ChatGPT generating texts based on statistical probabilities (Zhao et al., 2023).

This paper concerns the extent to which text-generating AI can support the development of (human) genre competence and how suitable AI is as a tool for genre-based writing didactics, investigating the differences between positions such as that of Hutz (2021), who favors the teaching of genre or text type competence in foreign language lessons, and that of Wendt et al. (2023), who call for AI literacy (as does Bubenhofer, 2022). However, to answer these questions, it is necessary to examine whether AI tools (exemplary ChatGPT) themselves are “competent” in terms of text genres. Adopting ChatGPT as an exemplar of such tools, we ask whether the LLM is capable of producing genre-specific texts or identifying and analyzing genre-specific patterns. We further consider whether there are differences in terms of genre. To examine these questions, we conducted an exploratory pilot study that saw several tasks (classifying, analyzing, condensing/summarizing, generating) performed on six different genres to test ChatGPT’s genre competence.<sup>1</sup> The findings correspond to at least of the three aspects of “Education in the Digital World” identified in the call for contributions to this special issue, namely, competencies, possible changes to educational and learning processes using AI tools, and appropriate tools for education in general.

In the following section, we discuss the prototypical concept of “genre” as a basic classification category (Section 2). Then, we describe genre compe-

---

<sup>1</sup> Ascribing genre competence to ChatGPT is, of course, a deliberate exaggeration on our part. We are aware that competence is something that is commonly attributed to living beings and that the concept of competence cannot be comparably applied to the performance of LLMs and AI. This contrasts with the position of, for example, Schneider and Zweig (2023), who treat text grading as a human performance and discuss the extent to which AI (as an e-rater) can do something that replaces grading.

tence as a fundamental literary skill before first presenting the design (Section 3) and then the results of our exploratory study (Section 4). Using the four areas of application described above, we consider how ChatGPT performs in different genres. The most important findings can be summarized as follows. Although ChatGPT is especially successful at generating and classifying texts, it faces challenges in terms of revising and condensing texts. Furthermore, its performance depends on the genre: It manages much better with texts that follow a strongly prototypical structure and offer little scope for individual creativity than with texts that are dependent on a situation and characterized by individuality. We conclude by summarizing the results of the analysis and suggesting some possible applications in the context of education (Section 5).

## 2 Text Genres and Genre Competence

### 2.1 On the Typicality of Genres: Text-Linguistic, Pragmatic, and Corpus-Linguistic Perspectives

“Genre” can be described as the basic category of scientific as well as everyday language text classification.<sup>2</sup> A concrete text is always perceived as an example of a genre (Brinker et al., 2018, p. 133; Heinemann, 2000, p. 517).<sup>3</sup> However, “genre” is a prototypical concept (Adamzik, 2016, pp. 327, 334; Sandig, 2000, p. 108) because genres are characterized or distinguished by typical features, and individual text samples include prototypical representatives of a genre containing a high number of typical features as well as less typical and more peripheral representatives.

Numerous text-linguistic models exist for describing genres (for an overview, see Krieg-Holz & Bülow, 2016; Trosborg, 1997). To differentiate between genres, all these models consider text-external and text-internal features. Despite differing in detail, they essentially include the following descriptive dimensions of genre: situational aspects, functional aspects, thematic aspects, structural aspects, and stylistic aspects. Accordingly, genre typicality or exemplarity can manifest in various ways and is an interplay of genre indicators at different levels, ranging from, for example, the choice of medium to the ways that the text is structured and graphically designed, the orientation of the ad-

---

<sup>2</sup> Following Trosborg (1997, p. 6), we use the term “genre” – corresponding to the German term “Textsorte” – to refer to texts “used in a particular situation for a particular purpose” and “classified using everyday labels such as a guidebook, a nursery rhyme, a poem, a business letter, a newspaper article, a radio play, an advertisement, etc.”

<sup>3</sup> For a detailed conceptual discussion of “genre” – referring to concrete texts with common properties – and ‘genre pattern’ (in German, “Textsortenmuster”) – denoting the underlying mental quantity that acts as a guide to how texts of a certain genre are prototypically constituted – see Brommer (2018a, pp. 67–70), which builds upon Fix (2008).

dressee, and the typical and exemplary choice of words and word combinations on the text surface. These different kinds of exemplary features are all relevant from both a pragmatic perspective, enabling the description and analysis of genres, and a didactic perspective, enabling competent handling of genres.<sup>4</sup>

Corpus linguistics also follows this perspective of language use analysis by assuming first that patterns of language use should be located on the linguistic surface and second that patterned language use manifests in the use of individual words and word combinations that are typical for a certain section of language. This type of expression can then be operationalized as statistically measurable co-occurrence (Bubenhofer, 2009, p. 5; Feilke, 2012, p. 24). This enables empirical, corpus-linguistic descriptions of genres (Brommer, 2018a; Brommer, 2018b), building a bridge from genre via corpus linguistics to AI-based LLMs. LLMs such as ChatGPT start exactly at this point of measurable co-occurrence because they are based on statistics and probability. They calculate the probability of a word (more precisely: a token) appearing based on the words (tokens) from the input and the ones that the system has already used. Because LLMs feature countless internal parameters and have been trained on a large set of examples, they can generate natural-sounding (re)combinations of text. This means that when ChatGPT is asked to generate, analyze, or summarize a text of a certain genre, it does so based on its calculation of what texts of that genre typically look like.<sup>5</sup>

## 2.2 Genre Competence: Relevance and Acquisition

As detailed in the preceding section, texts do not stand alone but are integrated into situations of action and, as such, represent a medium of communication that is embedded in discourse domains. The production, as well as the reception, evaluation, and analysis of texts, is based on an underlying knowledge of genres and an (unconscious) comparison with conventionally anchored patterns (Lomborg, 2014; Siepmann, 2017a). Therefore, successful communication requires the competent use of these conventions to communicate effectively and appropriately in different contexts by meeting the expectations and norms of the particular genre (Brock & Schildhauer, 2017; Brommer, 2019b; Schicker & Saletović, 2023). In the course of their literalization, learners need to acquire genre competence (in German, “Textsortenkompetenz”; Heinemann, 2000, p. 517; see also the chapter on genre competence in Nünning, 2008) as a very central aspect of writing competence and language competence in general. The acquisition and, therefore, teaching of genre competence is of fundamental importance in the con-

---

<sup>4</sup> The connection between these pragmatic, text-linguistic, and didactic perspectives on genres (and also text quality) is discussed in more detail in Brommer (2018b, 2019a).

<sup>5</sup> This, in turn, means that ChatGPT is limited to a very much structure-based approach to genres, excluding, for example, situational aspects (envisioned typical addressees/authors/situations of use) and individual preferences, such as personal language style.

text of not only schooling (e.g., Augst et al., 2007; Esterl & Krieg-Holz (eds.), 2018; Fischer, 2009; Freudenberg-Findeisen (ed.), 2016; Rezat & Feilke, 2018) but also professional life (e.g., Fandrych & Thurmair, 2011; Lehnen & Schindler, 2010) and academic domains (e.g., Björk, 2003; Kaiser, 2010; Steinhoff, 2007; Swales & Feak, 2012). Genre competence enables people to compose texts according to the formal, content-related, and stylistic conventions of specific genres and, conversely, to intuitively assign texts to specific genres. Genre competence encompasses the knowledge and skills or the mastery of the text typological rules that are necessary for the successful production and reception of the widest possible range of different genres. Therefore, genre competence concerns not only knowledge of the formal, content-related, and linguistic features of specific genres but also knowledge about communicative contexts and situations. According to Siepmann (2017b), a distinction can be made between genre-related lexicogrammatical competence, which concerns the constitution and internal structure of genres, and a communicative competence in genres, which concerns the relationship between genre and the situation in which the communication takes place. A particular challenge for the acquisition and teaching of genre competence is navigating linguistic and pragmatic conventions in the context of multilingualism (e.g., Ahn, 2012; Hutz, 2021; Hyland, 2007).

As in many other sub-areas of language competence, genre competence can be systematically promoted and trained. Genre-based writing instruction (cf., e.g., Hutz, 2021) as well as the didactics of text procedures (cf., e.g., Bachmann & Feilke, 2014) draw attention to the conventional aspects of dealing with texts. Both approaches aim to reveal and convey patterns and routines of language use in specific communicative contexts.<sup>6</sup> Although the didactics of text procedures have become increasingly established in recent years, both didactic approaches share their positioning within the traditional classroom setting with a teacher. The dilemma in the acquisition and teaching of genre competence (as well as in other school teaching-learning contexts) is that learners need individualized scaffolding, but teachers have to serve the needs of an entire class at the same time.

This leads to questions about whether LLMs such as ChatGPT can support the development of genre-specific writing skills, such as by initiating debate about genres in dialogue with learners, thus providing genre awareness training and imparting knowledge about genres. In a resource-friendly way, they can support the acquisition of receptive genre competence, an important prerequisite for being able to write different genres competently and confidently. It should also be considered whether LLMs may, in a further step, also contribute to learners being able to apply and implement the explicit knowledge they have acquired about genres in general and about specific genres in their own writing practice, that is, whether learners thereby acquire active or production-oriented genre knowledge and competence.

---

<sup>6</sup> What Feilke (2012) calls “text routines” (in German, “Textroutinen”) in his essay can be more accurately conceptualized as “genre routines” (in “Textsortenroutinen”).

To investigate the didactic potential of LLMs for the promotion of human genre competence, it is necessary to first test the genre competence of the LLMs themselves. This is what our exploratory study achieves, with the following section presenting its design.

### 3 Study Design

The purpose of the current study was to test the genre competence of LLMs using ChatGPT (version 3.5) as an example. The underlying focus concerned the extent to which such LLMs might be used for the individual promotion of genre competence. Because genre competence encompasses different aspects (Chapter 2), and the goal was to obtain a broad impression of ChatGPT's competence, we asked ChatGPT to (1) classify, (2) analyze, (3) condense/summarize, and (4) generate texts (of a given genre) using different prompts in German. In this way, we covered different areas of application and tested productive as well as receptive and analytical activities.

We have chosen the following genres as examples of different socio-situational contexts: congratulatory letter (“Gratulationsschreiben”) letter of condolence (“Kondolenzschreiben”), book/film review (“Buch- oder Film-Rezension”), discursive essay (“Erörterung”), job advertisement (“Stellenanzeige”), and promotional slogans (“Werbeslogans”).<sup>7</sup> Following Brinker's (1985/2005) classification of genres, on the one hand, and Searle's (1976) classification of speech acts (especially illocutionary acts), on the other hand, the study covers texts that have (primarily) an information, appeal, and contact function within this selection of genres. This enables the comparative investigation of whether the performance of ChatGPT correlates with the text function.

Data generation took place in the context of a seminar.<sup>8</sup> To ensure that the data would be comparable, the prompts were given by the seminar instructor and adapted to the previously mentioned specific task.<sup>9</sup> Then, students were asked to choose a genre (see above) and to test the appropriate prompts with ChatGPT (the students only had to follow the instructions; the authors analyzed the data). Thus, one student at a time performed all tasks (classifying, analyzing, condensing/summarizing, generating) for a single genre. Part of the interaction with ChatGPT was the confrontation with initial exemplary texts (written by humans), namely, when ChatGPT was prompted to classify,

---

<sup>7</sup> Because the study focuses on the question of how ChatGPT can handle different genres, multimodality was not considered, even though it does represent a prominent topic in current (text) linguistics (see, e.g., Stöckl, 2015).

<sup>8</sup> This was the seminar “ChatGPT – How we (might) write texts in future” in the undergraduate linguistics program at the University of Bremen.

<sup>9</sup> The assignment and the initially proposed prompts appear in the appendix.

analyze, or summarize a given text (e.g., a letter of condolence or a book/film review).<sup>10</sup> For this purpose, text samples were selected that can be considered prototypical representatives of the respective genre. The selection is based on didactic textbooks and guidebook literature, which provide (proto)typical examples for the corresponding genre.<sup>11</sup>

Subsequently, we qualitatively analyzed the complete transcripts of these dialogues (prompts plus results from ChatGPT) for the purpose of this article. The analyses are based on a total of 40 ChatGPT transcripts, with several transcripts available for some genres.<sup>12</sup> Table 1 breaks down the distribution of analyses by genre.

Table 1: Distribution of the Analyzed Transcripts

	<b>Dialogues about classifying</b>	<b>Dialogues about analyzing</b>	<b>Dialogues about condensing/ summarizing</b>	<b>Dialogues about generating</b>
<b>congratulatory letter</b>	1	1	1	1
<b>letter of condolence</b>	2	2	2	2
<b>book/film review</b>	3	3	3	3
<b>discursive essay</b>	1	1	1	1
<b>job advertisement</b>	1	1	1	1
<b>promotional slogan</b>	2	2	2	2
<b>subtotal</b>	10	10	10	10
<b>subtotal</b>				40

After analyzing the data with regard to the different tasks (classifying, analyzing, condensing/summarizing, generating), we considered the data in terms of differences between genres. The following section examines how ChatGPT performs in terms of the different tasks.

- <sup>10</sup> Due to limited space, we can only provide a limited selection of extracts of the original texts that were given to ChatGPT, especially because they are usually quite long.
- <sup>11</sup> Thus, we only confronted ChatGPT with text samples that were atypical for the genre in a few cases. Systematically testing the performance of ChatGPT in terms of how it deals with deviations and variations would be a worthwhile topic to explore.
- <sup>12</sup> Because students were free to choose one genre, there are different numbers of transcripts for each genre. They were also free to test additional prompts (see appendix), meaning that there is no consistent number of prompts per task.



## 4 Exploratory Results

### 4.1 Classifying Texts

Text classification is an essential part of text reception and text analysis and a concern of text linguistics as well as reading and writing didactics. Depending on the context, classification can take place unconsciously, consciously, intuitively, based on genre competence, and based on text-internal or text-external criteria (see Section 2 for further discussion). Didactically, the classification of texts is located in the field of reading promotion and pertains to the metacognitive reading strategies (Philipp, 2015; Locher & Philipp, 2023) that facilitate the reception of texts. Therefore, classifying texts forms part of genre competence.

Nowadays, “text classification” is mainly understood as automated text classification (Adamzik, 2016, p. 309). The term is used to refer to the field of natural language processing and to machine learning models that – when appropriately trained with data – are able to organize large amounts of unstructured text data in different ways and extract important information from that data. For example, automated text classification is used to process customer reviews on online portals and to assign texts to predefined content categories. ChatGPT also works in this way, using training data to classify concrete text samples that are presented and attempting to generate a meaningful and appropriate response based on the context and the information provided.

Comparing the results obtained by ChatGPT in the study with what we (as humans) expect from text classification, we first observe that ChatGPT classified all submitted texts correctly and adequately. Because prototypical examples were chosen to begin with, this was unsurprising.

Later, however, we challenged ChatGPT by allowing it to classify a text that could not be clearly assigned to a genre. ChatGPT dealt with the difficulty of the concerned classification in a differentiated way, namely, by stating that more context was needed, as the example shows:

(1)

**Prompt:** Klassifiziere den folgenden Text: “Ob du’s glaubst oder nicht: Man kann auf Fleisch verzichten, aber nicht auf guten Geschmack.”

*Classify the following text: “Believe it or not, you can give up meat, but you can’t give up good taste.”*



**ChatGPT:** Der gegebene Text ist ein Beispielsatz und kann keiner spezifischen Textsorte zugeordnet werden. Es handelt sich um einen kurzen, prägnanten Satz, der eine Aussage enthält. Ohne weiteren Kontext könnte der Text als ein Zitat, ein Werbeslogan, ein Teilsatz in einem längeren Text oder eine informelle Aussage in einem Gespräch betrachtet werden. Um die Textsorte genauer zu bestimmen, wäre zusätzlicher Kontext erforderlich.

*The given text is an example sentence and cannot be assigned to any specific text type. It is a short, concise sentence that contains a statement. Without further context, the text could be considered a quote, an advertising slogan, a partial sentence from a longer text, or an informal statement in a conversation. Additional context would be required to determine the text type more precisely.*

In all cases, ChatGPT automatically classified the texts according to genre, regardless of whether the prompt only asked it to classify the text (without specifying on what basis the classification should take place; see example (1) above) or whether the task specifically asked it to classify the text by genre (see example (2) below; notably, these two conditions were not tested systematically).

To justify its classification, ChatGPT used both text-internal and text-external criteria, with content and linguistic style being the most relevant criteria and other criteria (such as the structure of a text) less frequently considered. Letters of condolence, for example, were identified primarily on the basis of their content (prototypical features mentioned were expressions of condolence, remembrance, consolation, and wishes), as well as typical formulation patterns (personal salutations, farewell formulae). In the explanatory notes, ChatGPT mostly refers to suitable examples from the text, as in the example of the classification of a book review:

(2)

**Prompt:** Ich gebe Dir gleich einen Text. Um welche Textsorte handelt es sich dabei und wie kommst Du zu dieser Einschätzung?

*I'll give you a text in a moment. What type of text is it, and how do you arrive at this assessment?*

**ChatGPT:** Basierend auf der Analyse des vorliegenden Textes handelt es sich hierbei um eine persönliche Buchrezension oder Meinungsäußerung. Hier sind die Indikatoren, die zu dieser Einschätzung führen:

- 1) Erwähnung des Buchtitels und der Autorin: Der Text bezieht sich auf das Buch "Pride and Prejudice" von Jane Austen und betont, dass es das Lieblingsbuch des Autors/der Autorin ist.

- 2) Zusammenfassung des Inhalts: Es wird eine kurze Zusammenfassung der Handlung gegeben, einschließlich der Einführung der Hauptcharaktere und des zentralen Konflikts.
- 3) Bewertung des Buches: Der Autor / die Autorin gibt seine/ihre persönliche Meinung zum Buch ab und hebt positive Aspekte hervor, wie die romantische Handlung, die liebenswerten Charaktere, den Humor und den schönen Schreibstil von Jane Austen.
- 4) Empfehlungen: Der Autor/die Autorin empfiehlt das Buch verschiedenen Zielgruppen, einschließlich Fans von Romantik, familienorientierten Geschichten, Komödien und Klassikern. Es wird auch erwähnt, dass das Buch auch für Leser geeignet ist, die normalerweise Klassiker meiden.

Basierend auf diesen Elementen würde ich diesen Text als eine Buchrezension oder literarische Meinungsäußerung einstufen.

*Based on the analysis of the text at hand, this is a personal book review or opinion piece. Here are the indicators that lead to this assessment:*

- 1) *Mention of the book title and author: The text refers to the book "Pride and Prejudice" by Jane Austen and emphasizes that it is the author's favorite book.*
- 2) *Summary of content: A brief summary of the plot is provided, including the introduction of the main characters and the central conflict.*
- 3) *Evaluation of the book: The author(s) gives his/her personal opinion about the book and highlights positive aspects, such as the romantic plot, the lovable characters, the humor, and the beautiful writing style of Jane Austen.*
- 4) *Recommendations: The author(s) recommends the book to various audiences, including fans of romance, family-oriented stories, comedies, and classics. It is also mentioned that the book is suitable for readers who usually avoid classics.*

*Based on these elements, I would classify this text as a book review or literary opinion piece.*

ChatGPT also performed satisfactorily at the task of ranking submitted texts of a genre with regard to their prototypicality, with specific prompts asking it how typical a text is for a certain genre. Although there was no evidence that classifying texts of certain genres was more difficult for ChatGPT than texts of other genres, there were several inconsistencies, such as when ChatGPT referred to prototypical features while ranking the texts but later had to put its own assessment into perspective. Nonetheless, this reflection testifies to the transparency of ChatGPT's classifications, which are also conclusive in terms of content.

In summary, ChatGPT reliably classified all the examined genres, and the justifications for the classifications were correct and plausible in terms of content. On a positive note, ChatGPT provided meta-linguistic commentary regarding the concept of prototypicality in all classifications.<sup>13</sup> It should be critically noted that genre labels were not always consistent. For example, the terms “birthday greeting,” “birthday wish,” and “birthday card or message” were used arbitrarily in the classifications. However, this can be justified by the fact that some genres are less established in terms of their designation (e.g., “congratulatory letter”) than others (e.g., ‘job advertisement’).

## 4.2 Analyzing Texts

For analyzing texts, two distinct tasks were conducted: the first primarily focused on coherence analysis and thus semantic aspects and the deep text structure (van Dijk, 1972). Therefore, ChatGPT was prompted to “analyze the text in terms of its coherence” (original German prompt: “Analysiere den Text hinsichtlich seiner Kohärenz.”). The second task, a cohesion analysis, centered on the linguistic surface and formal as well as stylistic aspects by prompting ChatGPT to “analyze the text on a stylistic level” (original German prompt: “Analysiere den Text auf sprachlich-stilistischer Ebene”). Both levels of a text – its deep structure (coherence) as well as the linguistic surface (cohesion) – must be considered when analyzing genres, and both are essential for genre competence and knowledge (Nünning, 2008, p. 93; Krieg-Holz & Bülow, 2016, p. 4). To generate and recognize coherence, the specific context of a text and the shared knowledge about it are crucial. The text refers to this by means of presuppositions, implications, and deictic elements. However, because ChatGPT has no access to the context and no background knowledge about the text, we expected that ChatGPT would focus on the linguistic surface when an-

---

<sup>13</sup> For example, when asked about the prototypical characteristics of a condolence letter, ChatGPT stated, “Der Prototyp eines Kondolenzschreibens ist ein Brief oder eine Mitteilung, die verwendet wird, um sein Beileid und Mitgefühl zum Ausdruck zu bringen, wenn jemand einen Verlust oder Todesfall erlitten hat. Es ist eine Art schriftliche Kondolenzbekundung, die oft an die trauernde Person oder an die Familie des Verstorbenen gerichtet ist. Der Prototyp eines Kondolenzschreibens enthält in der Regel folgende Elemente: [...]“ (in English, “The prototype of a letter of condolence is a letter or message used to express condolences and sympathy when someone has suffered a loss or death. It is a type of written expression of condolence that is often addressed to a grieving person or the family of a deceased individual. The prototype of a letter of condolence usually contains the following elements: [...]”)

alyzing coherence and thus not be able to clearly separate it from cohesion. In the case of linguistic and stylistic analysis, it was expected that ChatGPT would examine linguistic features and their use in the analyzed text. General linguistic features include formal, lexical, syntactical, and narrative aspects (Krieg-Holz & Bülow, 2016, pp. 90–153; Pitz-Klauser, 2019, pp. 64–73). Additionally, singular stylistic phenomena, such as figures of speech, metaphors, antithesis, parallelisms, and rhetorical questions, should also be subject to examination (Krieg-Holz & Bülow, 2016, pp. 160–169; Pitz-Klauser, 2019, pp. 74–78). Essentially, such an analysis should be able to describe the linguistic elements employed in a text and articulate their effects on its reception.

ChatGPT was first prompted to analyze the text with regard to coherence. The results were generally reasonable and consistent, albeit with variations depending on the text's typicality. Usually, the analysis began with ChatGPT providing a general assessment of the text's coherence in relation to its topic and genre. A coherence analysis focuses on the topic and topic development as well as the structure of the text; for the most part, ChatGPT combined the results with observations concerning the linguistic surface. For shorter texts, such as congratulatory letters, ChatGPT delved into detailed paragraph-level (and even sentence-level) analysis. Conversely, for longer and more intricate texts, such as book/film reviews and discursive essays, the analysis was more concise due to the space limitations of ChatGPT's answers. In these cases, ChatGPT offered an overview of the topic's development and occasionally suggested improvements to enhance coherence. The analysis culminated in a summary paragraph that considered both structural and thematic coherence, reiterating the central themes of the text. Interestingly, the coherence analysis of shorter texts sometimes extended to a greater length, with a tendency to over-interpret and over-reiterate text content compared to the more concise (and sometimes superficial) analysis of longer texts.

Thus, as the complexity and length of the texts increased, the coherence analyses became more selective, summative, and superficial, as demonstrated by the following analysis of a somewhat lengthy discursive essay:

(3)

**Prompt:** Ich gebe Dir gleich einen Text. Analysiere den Text hinsichtlich seiner Kohärenz.

*I'll give you a text in a moment. Analyze the text in terms of its coherence.*

**ChatGPT:** Der vorliegende Text behandelt das Thema der mangelnden Zeit und Aufmerksamkeit für alte, kranke und sterbende Mitbürger in der Gesellschaft. Er beleuchtet verschiedene Gründe für dieses Phänomen und gibt Einblicke in die sozialen, gesellschaftlichen und individuellen Aspekte, die dazu beitragen. Der Text ist insgesamt kohärent, da er eine klare Struktur aufweist und die Argumente logisch miteinander verknüpft sind. Hier sind einige Merkmale der Kohärenz im Text:

- 1) Einleitung: Der Text beginnt mit einer konkreten Begegnung, in der eine Krankengymnastin von ihrer Entscheidung erzählt, nicht als Ärztin zu arbeiten, aufgrund der beobachteten “Massenabfertigung” von Patienten im Krankenhaus. Dies dient als Ausgangspunkt für die folgenden Überlegungen des Autors.
- 2) Analyse der Gründe: Der Text bietet verschiedene Gründe für die mangelnde Zeit und Aufmerksamkeit für alte und kranke Menschen. Dazu gehören das reichhaltige Angebot an Freizeitmöglichkeiten in unserer Wohlstandsgesellschaft, die Scheu vor dem Thema Tod und Altern, eigene Probleme und Sorgen, sowie die ichbezogene Natur der Gesellschaft.
- 3) Beispiele und Erfahrungen: Der Autor verwendet Beispiele und persönliche Erfahrungen, um seine Argumente zu stützen. Es werden Situationen beschrieben, in denen ältere Menschen einsam sind oder von ihren Familien vernachlässigt werden.
- 4) Zusammenfassung und Appell: Am Ende des Textes wird zusammengefasst, dass Hilfsbereitschaft und soziales Miteinander wichtig sind. Es wird darauf hingewiesen, dass die Gesellschaft immer älter wird und es somit notwendig ist, sich mit den Problemen der älteren Generation auseinanderzusetzen. Der Text appelliert an die Leser, sich gegenseitig zu unterstützen, da auch sie selbst irgendwann auf die Hilfe anderer angewiesen sein könnten.

Insgesamt zeigt der Text eine klare Argumentationsstruktur und logische Verknüpfung der Gedanken, was seine Kohärenz ausmacht. Die Verwendung von Beispielen und persönlichen Erfahrungen macht die Aussagen anschaulich und verständlich.

*This text addresses the issue of the lack of time and attention given to elderly, sick, and dying fellow citizens. It highlights various reasons for this phenomenon and provides insights into the social, societal, and individual aspects that contribute to it. The text is coherent overall because it has a clear structure, and the arguments are logically linked. Here are some characteristics of coherence in the text:*

- 1) *Introduction: The text starts with a concrete encounter in which a physiotherapist talks about her decision not to work as a doctor because of the observed “mass processing” of patients in the hospital. This serves as a starting point for the author’s subsequent reflections.*
- 2) *Analysis of reasons: The text offers several reasons for the lack of time and attention given to the elderly and sick. These include the abundance of leisure opportunities in our affluent society, shyness about death and aging, our own problems and concerns, and the ego-centered nature of society.*
- 3) *Examples and experiences: The author uses examples and personal experiences to support his arguments. Situations are described in which elderly people are lonely or neglected by their families.*
- 4) *Summary and appeal: At the end of the text, it is summarized that helpfulness and social togetherness are important. It is recognized that society is getting older, and it is therefore necessary to deal with the problems of the older generation. The text appeals to readers to support each other because they themselves may need the help of others at some point.*

*Overall, the text shows a clear structure of argumentation and logical linking of thoughts, which accounts for its coherence. The use of examples and personal experiences makes the statements clear and understandable.*

It can be said that ChatGPT was largely able to analyze the coherence of the texts in a consistent and convincing manner, apart from a few weaknesses in detail (especially phrase-like repetition of content).

The second text analysis task involved a stylistic analysis. A notable aspect here was the absence of concrete examples provided by ChatGPT in approximately half of the analyses. This presents an evaluation dilemma because, for example, ChatGPT occasionally identifies the use of rhetorical questions when rhetorical questions are used but also gives examples that are not rhetorical questions. This highlights the importance of concrete examples for reliable assessment. In cases with concrete examples, ChatGPT generally produced results that aligned with expectations regarding more general stylistic aspects, including formality (e.g., formal, informal), narrative parameters (e.g., first-person, second-person), emotionality (e.g., positivity, sadness, friendliness), word choice (e.g., adjectives, adverbs), and sentence complexity (e.g., sentence length).

However, ChatGPT demonstrated challenges in recognizing and interpreting figures of speech, which often constitute more nuanced and specific text features. In more than half of cases, ChatGPT's identification of figures of speech was inaccurate. While it accurately recognized metaphors, it frequently misidentified not only rhetorical questions (as mentioned) but also antithesis and repetition, among other elements. To illustrate this, we present as an example an analysis of a film review, where assessments of language register, sentence length, and word choice align with expectations, but the identification of idioms, rhetorical questions, and intensification is inconsistent.

(4)

**Prompt:** Analysiere den Text auf sprachlich-stilistischer Ebene.

*Analyze the text on a stylistic level.*

**ChatGPT:** Auf sprachlich-stilistischer Ebene lässt sich der Text als informierender und bewertender Artikel beschreiben. Er ist überwiegend in einer sachlichen und beschreibenden Sprache verfasst, wobei gelegentlich auch subjektive Einschätzungen und Meinungen des Autors einfließen. Der Text verwendet eine recht flüssige und gut verständliche Sprache, ohne dabei zu komplex oder überladen zu wirken. Die Sätze sind größtenteils kurz bis mittellang, was zur Klarheit und Lesbarkeit beiträgt. Es werden zudem einige Fachbegriffe aus dem Film- und Literaturbereich verwendet, wie beispielsweise "Verfilmung", "Kinofassung", "Bollywood-Musical-Version" und "Schulkampagne", die dem Text eine gewisse fachliche Komponente verleihen.

Der Text enthält auch einige stilistische Elemente, um das Interesse des Lesers zu wecken und die Informationen ansprechend zu präsentieren. Dazu gehören beispielsweise die Verwendung von Adjektiven wie "überaus liebevollen Ausstattung" und "angenehme Leichtigkeit", um den Film zu beschreiben, oder die Verwendung von Vergleichen wie "auf angenehme Weise frisch und modern" und "wie auch immer man es betrachtet". Solche stilistischen Mittel verleihen dem Text eine gewisse Farbigkeit und machen ihn interessanter zu lesen.

Des Weiteren werden im Text einige rhetorische Fragen verwendet, wie beispielsweise "Doch bis der Titel gebende Stolz und die zahlreichen Vorurteile aus dem Weg geräumt sind, soll es noch eine Weile dauern...", um den Leser zum Nachdenken anzuregen oder eine gewisse Spannung aufzubauen.



Der Autor des Textes verwendet auch einige Verstärkungen, um seine Meinungen zu unterstreichen und den Film positiv zu bewerten. Beispielsweise wird Keira Knightley als “britischer Shooting Star“ bezeichnet, Donald Sutherland und Brenda Blethyn werden als “brillierend in ihren Rollen“ bezeichnet, und es wird betont, dass der Film auf “denkbar unterhaltsame Weise die Freude der klassischen britischen Literatur nahe bringt“. Solche stilistischen Mittel verstärken die positiven Eindrücke, die der Autor von dem Film gewonnen hat.

Insgesamt lässt sich sagen, dass der Text auf sprachlich-stilistischer Ebene gut strukturiert, verständlich und ansprechend ist. Er verwendet eine sachliche Sprache mit gelegentlichen stilistischen Elementen, um Informationen zu vermitteln und den Leser zu interessieren.

*On a stylistic level, the text can be described as an informative and evaluative article. It is written predominantly using factual and descriptive language, with occasional subjective assessments and opinions offered by the author.*

*The text uses relatively fluid and easy-to-understand language without being overly complex or cluttered. For the most part, sentences are short to medium in length, which contributes to clarity and readability. There is also some use of film and literary terminology, such as “film version,” “theatrical version,” “Bollywood musical version,” and “school campaign,” which give the text a certain technical component.*

*The text also contains some stylistic elements to arouse the reader’s interest and present the information in an appealing way. These include, for example, the use of adjectives such as “exceedingly lovely décor” and “pleasant lightness” to describe the film or the use of comparisons such as “fresh and modern in a pleasant way” and “however you look at it.” Such stylistic devices add color to the text and make it more interesting to read.*

*Furthermore, some rhetorical questions are used in the text, such as, “But until the pride that gives the title and the numerous prejudices are cleared out of the way, it shall still take a while...”, to make the reader think or to build up some tension.*

*The author of the text also uses some amplification to emphasize his opinions and to evaluate the film positively. For example, Keira Knightley is called a “British shooting star,” Donald Sutherland and Brenda Blethyn are described as “brilliant in their roles,” and it is emphasized that the film “brings home the joy of classic British literature in the most entertaining way imaginable.” Such stylistic devices reinforce the positive impressions the author gained from the film.*

*Overall, it can be said that the text is well-structured, understandable, and appealing on a stylistic level. It uses factual language with occasional stylistic elements to convey information and interest the reader.*

Figures of speech rely on nuances, ambiguities, and implicit intentions within a text (Krieg-Holz & Bülow, 2016, pp. 160–169), and ChatGPT appears to be limited in terms of comprehending these subtleties.<sup>14</sup> When directly questioned about its difficulties recognizing figures of speech, ChatGPT's response corroborated the findings:

(5)

**Prompt:** Ist es für dich eine Herausforderung, Redewendungen zu erkennen?

*Is recognizing figures of speech a challenge for you?*

**ChatGPT:** Das Erkennen von Redewendungen kann für Sprachmodelle wie mich eine Herausforderung sein, vor allem wenn die Redewendungen komplex oder subtil sind. Ich bin zwar auf eine Vielzahl von Texten trainiert worden und kenne viele gebräuchliche Redewendungen, aber mein Verständnis basiert auf Mustern und Beispielen, die in dem Text vorkommen, auf den ich trainiert wurde. Im Allgemeinen kann ich gängige Redewendungen wie Gleichnisse, Metaphern, idiomatische Ausdrücke und Übertreibungen erkennen und erklären. Nuanciertere oder kreativere Redewendungen erkenne ich jedoch nicht immer richtig, und meine Interpretationen erfassen möglicherweise nicht die gesamte vom Autor beabsichtigte Bedeutungstiefe.

*Recognizing figures of speech can be a challenge for language models like me, especially when the figures of speech are complex or subtle. While I have been trained on a wide range of texts and have knowledge of many common figures of speech, my understanding is based on patterns and examples present in the text I was trained on. I can generally recognize and explain common figures of speech, like similes, metaphors, idioms, and hyperbole. However, I may not always correctly identify more nuanced or creative figures of speech, and my interpretations may not capture the full depth of meaning intended by the author.*

Here, too, ChatGPT's reflections offer transparency in terms of the task performed, making them a useful possible starting point for dealing with genre (competence) in educational contexts.

---

<sup>14</sup> This observation is in line with the study by Perder & Susteck (2023) who investigated the performance of ChatGPT in analyzing poetry in literature classes. They also concluded that ChatGPT shows weaknesses in analyzing figurative language.

### 4.3 Condensing and Summarizing Texts

In a third step, we will now focus on the task of condensing and summarizing texts. Regardless of the results, this task anticipated that after the implementation, the target text would be shorter in the sense of having less linguistic material, that is, fewer words or characters. However, the condensed version should not deviate from the unabridged original in content insofar as the main message should remain the same and the text should continue to be recognizable as a sample of the corresponding genre. That is, condensing does not mean simply cutting the text short or omitting things at will but also selectively leaving out those parts that do not fundamentally affect the basic information of the text. In this sense, the task of condensing relates closely to the task of summarizing, which describes formulating something more briefly or getting to the core of the content in fewer words in the form of a résumé.

To perform this task in our study, ChatGPT was initially prompted to shorten a text sample by about 30%.<sup>15</sup> What is particularly striking about the results is that ChatGPT seems to have great difficulties with this task not at the level of content but primarily at the level of the text surface, that is, with regard to its length. Almost none of the examples saw ChatGPT able to shorten the text by the length specified due to inexact calculations of the number of words. Although this may seem surprising given that word counting tools have been implemented in text processing programs for a long time, LLMs generally still “face difficulties in numerical computation” (Zhao et al., 2023, p. 52). This struggle is especially noticeable when ChatGPT reacts to thematizations of this incapability, as the following example illustrates in the context of the condolence letter genre. The user prompted ChatGPT to shorten the original text by about 30% of its length, with ChatGPT’s implementation leading to the following dialog:

(6)

**Prompt:** Um wieviel hast du den ursprünglichen Text gekürzt?

*How much did you cut from the original text?*

**Chat GPT:** Der ursprüngliche Text hatte 40 Wörter, während die gekürzte Version nur 27 Wörter enthält. Die Kürzung beträgt also 13 Wörter, was einer Reduzierung von etwa 32,5% entspricht.

*The original text had 40 words, while the shortened version contains only 27 words. So, the shortening is 13 words, which is a reduction of about 32.5%.*

---

<sup>15</sup> Original prompt (in German): “Ich gebe Dir gleich einen Text. Bitte kürze ihn um ca. 30 Prozent auf zwei Drittel seiner Länge.” (in English, “I’ll give you a text in a moment. Please shorten it by about 30 percent to two-thirds of its length.”)

**Prompt:** Zähle den ursprünglichen Text erneut auf die Anzahl der Wörter

*Count the original text again for the number of words.*

**ChatGPT:** Der ursprüngliche Text besteht aus 8 Wörtern.

*The original text consists of 8 words.*

**Prompt:** Welches sind diese Wörter?

*What are these words?*

**ChatGPT:** Die Wörter im ursprünglichen Text sind: “Lieber Heinz, Mit großem Bedauern habe ich vom Hinschied deiner Mutter erfahren. Ich entbiete euch allen meine aufrichtige Anteilnahme und wünsche euch viel Trost und Kraft in dieser schwierigen Zeit. In stiller Trauer.”

*The words in the original text are: [...]*

**Prompt:** Das sind mehr als 8 Wörter.

*That is more than 8 words.*

**ChatGPT:** Entschuldigung für das Missverständnis. Tatsächlich besteht der ursprüngliche Text aus 20 Wörtern.

*Sorry for the misunderstanding. In fact, the original text consists of 20 words.*

The example shows that ChatGPT struggles to provide the correct number of words in the given text. It counts 40 words, then 8, then 20, and finally (not visible in the excerpt) 30, which is closest to the actual number of words (33). While the struggle to condense a text which is already very short (33 words only) is understandable, especially given that condolences contain certain elements (such as opening and closings) that can hardly be omitted because the genre requires a high level of politeness and is generally very formulaic (Fandrych & Thurmair, 2011, pp. 301–311; Bachmann-Stein, 2021, pp. 31–34), the same phenomenon can be observed for longer original texts as well as for other genres, including discursive essays or book/film reviews.

However, although condensing at the level of form is hardly successful due to the technical limitations mentioned, ChatGPT's performance at the level of content meets expectations much better. This concerns both the answers to the prompt already examined and those directly asking for a summary (with different requirements regarding the length). Likely because this seems to be a task commonly performed by ChatGPT (and also one that is often mentioned in the media and educational discourse), it also largely meets expectations (also Bubenhofer, 2022).

When prompted to "reflect" on its own performance regarding the task at hand, ChatGPT's inability to count words became apparent. At one point, the user asked ChatGPT about its shortening strategy. This example (7) concerns the book review genre, with the user prompting ChatGPT to summarize the core idea in one sentence. After the output, the user asks the following:

(7)

**Prompt:** Was hast du dabei weggelassen?

*What did you leave out?*

**ChatGPT:** Bei der Kürzung des Textes habe ich mich bemüht, die Kernaussagen und Hauptthemen beizubehalten, während ich unwesentliche Details und wiederholende Ausdrücke entfernt habe. Hier sind einige der Dinge, die gekürzt oder weggelassen wurden:

- \ Die Erwähnung der eigenen Leseerfahrung und anfänglichen Meinung zum Buch.
- \ Einige persönliche Meinungsäußerungen und Bewertungen der Charaktere.
- \ Einige spezifische Beispiele und Details aus der Handlung.
- \ Wiederholungen von Aussagen und Gedanken.
- \ Zusätzliche Vergleiche und Erwähnungen von anderen Büchern.
- \ Einige abschweifende Absätze und Sätze, die nicht direkt zur Kernaussage beitragen.

*In shortening the text, I tried to keep the key messages and main themes, while removing non-essential details and repetitive phrases. Here are some of the things that were shortened or omitted:*

- \ *Mentioning one's own reading experience and initial opinion of the book.*
- \ *Some personal expressions of opinion and evaluations of the characters.*
- \ *Some specific examples and details from the plot.*
- \ *Repetition of statements and thoughts.*
- \ *Additional comparisons and mentions of other books.*
- \ *Some digressive paragraphs and sentences that did not directly contribute to the main message.*

While the omission of repetitions and details may fully meet the expectation of a summary, it is quite revealing that ChatGPT would shorten the text in terms of opinions, evaluations, and personal statements, which form a crucial if not constitutive part of book reviews (Klemm et al., 2012). This could indicate that ChatGPT's shortening or summarizing strategies are not necessarily genre-specific but more generic. However, the fact that the greeting formulae are retained in the condolences and congratulatory texts suggests another direction. It should also be noted here that depending on the prompt – whether a simple shortening of a certain percentage or a one-sentence summary is requested – ChatGPT is likely to use different strategies. However, in the same transcript, the prompt “There's another text coming up in a moment, where you should put the core idea in a sentence of its own again” (original prompt in German: “Gleich kommt noch ein Text, bei dem du wieder die Kernaussage in einen eigenen Satz bringen sollst”) is extended in ChatGPT's initial response by including the aspect of shortening: “Of course, I'm standing by **to shorten** the next text and put the core idea into a sentence of its own.” (emphasis added by the authors; original German: “Natürlich, ich stehe bereit, um den nächsten Text zu kürzen und die Kernaussage in einen eigenen Satz zu fassen.”). Thus, it remains open where the line between condensing and summarizing is to be drawn exactly and the strategies used by the AI seem to be identical for both tasks. In this respect, the last bullet point in example (7) is also interesting in that it seemingly reveals ChatGPT's understanding of what summaries or shortenings generally aim at, namely, delivering the main message. However, this argument is rather tautological insofar as that understanding was equally formulated in the user's prompt.

## 4.4 Generating Texts

As previously indicated, it is important to follow genre conventions when generating a text. If ChatGPT is supposed to generate a text, it means that a user prompts ChatGPT to generate an example of one of the defined genres (“Generiere mir ein Beispiel für [Textsorte].” and “Verfasse ein weiteres Beispiel dieser Textsorte.”; in English, “Generate an example of [genre].” and “Write another example of this genre.”). The responses are expected to be an internally consistent and understandable text that meets the formal and content-related characteristics of the genre in accordance with human genre competence. To achieve this, we will consider ChatGPT’s performance in relation to (in the sense of deviation or conformity) these expectations. Apart from ChatGPT’s obvious struggle to calculate word counts (see above), the most striking observation with regard to the task of generating texts is the lack of internal consistency, which undermines ChatGPT’s ability to avoid contradictions within the same reaction to a prompt. Comparing the characteristics ChatGPT listed as important for a genre<sup>16</sup> in the context of the texts generated, it became apparent that there exists a certain dissonance between the generated texts and the features described as being characteristic of them. In particular, ChatGPT seems to struggle (in the sense of not meeting a competent reader’s genre-competence-based expectations) to generate individualized texts. Considering that ChatGPT defines “individuality” – in the sense of unique personality traits – as characteristic of the genres “condolence letter” and “birthday wishes,”<sup>17</sup> the generated texts turn out quite formulaic, the following example shows:

(8)

**Prompt:** Generiere mir ein Beispiel für einen Geburtstagsglückwunsch.

*Generate an example of a birthday wish.*

**ChatGPT:** Du bist ein wunderbarer Mensch, der stets mit seinem positiven Wesen und seinem warmen Herzen strahlt. [...] Deine positive Ausstrahlung und dein liebevolles Wesen sind inspirierend.

*You are a wonderful person who always shines with their positive nature and warm heart. [...] Your positive attitude and loving nature are always inspiring.*

<sup>16</sup> Original prompt (in German): “Worauf ist zu achten, wenn man eine/n [Textsorte] verfasst?” (in English, “What should you look out for when writing a [genre]?”)

<sup>17</sup> In response to the question, “What should you look out for when writing a birthday wish?” ChatGPT wrote the following (in excerpts): “1. Individuality: [...] focus on their unique personality traits, interests, or relationships. [...]” (Note that the prompt and answer have been translated).



Although an attempt to address personality traits is evident here (e.g., *positive nature*), these turn out to sound platitudinous, becoming very general and hardly individual. It is obvious that ChatGPT does not have the necessary information at its disposal to individualize the message.

In contrast, the generated job advertisements were expectedly structured with a focus on crucial information, providing a clear and immediate overview by structuring the job advertisement around different sections (introduction, job duties, expectations, offer from employer, contact information) using bullet points.<sup>18</sup> In contrast to the congratulatory texts, the (re)use of formulaic expressions and a repetitive structure is not only expected of job advertisements but is downright necessary to enable efficiency and the fast orientation of potentially interested parties. As a first interim conclusion regarding the task of generating texts, it can be stated that ChatGPT performs much better with clearly structured texts, which do not include personal information. This is unsurprising given that ChatGPT simply does not have access to the personal information required for congratulatory texts, whereas it likely has enough data for generic job advertisements.

A final observation concerning the task of generating texts concerns the discursive essay genre (“Erörterung”), a consecutive written form featuring clear structure, arguments that are interlinked, and transitions between paragraphs (cf., e.g., Feilke, 1990; Schicker & Akbulut 2023).<sup>19</sup> The structure of a discursive essay can be linear or dialectical. In the case of a linear arrangement of arguments, they appear in ascending order, moving from the weakest argument at the beginning to the strongest argument at the end. A dialectical arrangement sees pro and contra arguments alternate. In all generated texts (more than one example was asked for), ChatGPT consistently opted for the linear arrangement, failing to consider the weighting of the arguments in the meta-reflection or address other possible arrangements. Rather than creating a text with a focus on discursively interwoven and contrasted arguments, ChatGPT’s essays simply conveyed its information in the most effective and clearly structured manner, much like it did with the job advertisements. ChatGPT obviously does not do justice to the complex structure of a discursive essay. It is possible that this genre, which is typical in the school context, is insufficiently represented in the data on which ChatGPT was trained – in principle, ChatGPT is quite capable of writing argumentative texts.<sup>20</sup>

---

<sup>18</sup> Original prompt (in German): “Generiere mir ein Beispiel für eine Stellenanzeige.” (in English, “Write an example of a job offer.”)

<sup>19</sup> Original prompt (in German): “Verfasse ein Beispiel für eine Erörterung.” (in English, “Write an example of a discursive essay.”)

<sup>20</sup> In their exploratory study, Schicker and Akbulut (2023) show that ChatGPT is able to produce a coherent, elaborate, and largely grammatically correct argumentative text that is rated as good by prospective teachers.

Overall, ChatGPT's performance mostly fulfilled the expectations of a genre-competent reader in terms of generating texts. The tool is quite able to generate cohesive and coherent texts and differentiate between genres by identifying their different characteristics. However, due to their required personalization and emphasis, ChatGPT demonstrates limitations regarding internal consistency and performance for certain specific genres, especially contact texts. Furthermore, ChatGPT's preference for conveying information as effectively as possible when generating a text, as observed in the examples above, may not be suited to all genres.

## 5 Conclusion and Teaching Implications

In conclusion, the analysis of ChatGPT's performance on various genre-related tasks reveals several key findings. ChatGPT was particularly successful at generating and classifying texts (both common primary tasks of LLMs). ChatGPT performed strongly in the context of generating structured and cohesive texts, especially for clearly structured and impersonal genres, such as job advertisements. However, it demonstrated challenges in generating emphatic texts, often resulting in formulaic and non-individualized responses for genres such as condolence letters and birthday wishes. This limitation is due to the lack of personal information available to the model.

With regard to text classification, ChatGPT features a very reliable text classification system, accurately assigning texts to their respective genres based on content and linguistic style. It appropriately justifies its classifications using both text-internal and text-external criteria, and it also performs well at the task of recognizing prototypical examples of a given genre.

ChatGPT performed moderately well in the context of text analysis, where it was strong on the linguistic surface but had more difficulty with thematic coherence and more complex stylistic figures. ChatGPT's coherence analysis was generally reasonable but varied in depth based on text complexity. It tended to provide more detailed analysis for shorter texts and sometimes over-interpreted content. However, its analyses for longer and more complex texts were more concise. ChatGPT generally faces challenges in recognizing and interpreting figures of speech, which require nuanced comprehension.

In contrast, ChatGPT did not perform so well at the task of condensing/summarizing, demonstrating difficulties in accurately condensing texts to a specified length due to inaccurate calculations of text length.<sup>21</sup> However, it performed well in terms of content condensation, ensuring that the main message remained intact. This indicates that its challenges primarily concern the text's surface rather than the preservation of the content.

Overall, ChatGPT demonstrated competence at understanding and generating a variety of genres but challenges in terms of personalization, emphasis, recognizing figures of speech, and accurately shortening texts to a specified length. Its performance varies depending on the typicality of the text and the characteristics of the genre: The more formalized a genre, the better ChatGPT performs (which is unsurprising given that ChatGPT's genre competence is based on the genre's prototypical characteristics).

Notably, these study results represent a mere snapshot and concern the performance of the version of ChatGPT (3.5) available in the middle of 2022. It can be assumed that the revealed deficiencies in performance will decrease in the near future. Nonetheless, the didactic potential is obvious, especially when ChatGPT "reflects" on its own activity, with these reflections capturing the conditions of the task itself, revealing crucial aspects of genre competence, even if, in some cases, only in a rudimentary form.

What didactic conclusions can be drawn from these observations? As recognized by Hutz (2021), learners often find it challenging to write texts when ideas, writing goals, audience, text structures, and linguistic means are unclear or absent. In addition to guidance during the writing process, learners need a clear understanding of the genre they are working in. This could be the starting point for the application of ChatGPT in classrooms.<sup>22</sup> On the one hand, by conducting exercises using ChatGPT, students can reflect on the concept of genre competence in detail by, for example, comparing original source texts with ChatGPT's output. ChatGPT's analytical competence and familiarity with various genres can help students deepen their implicit and explicit knowledge of genres via experimental dialogues with ChatGPT.

---

<sup>21</sup> This clearly shows that ChatGPT follows different mechanisms than computer linguistic methods for machine tagging that have been working very reliably for quite a long time.

<sup>22</sup> See also Schicker & Akbulut (2023, pp. 192–193), who also outline some ideas about the didactic potential of ChatGPT for developing genre competence.

On the other hand, ChatGPT can be used for the writing process itself. Ideally, writing involves an interactive approach that allows for the exchange of ideas about content and linguistic style, including the structure and logic of argumentation and the appropriateness of the linguistic presentation. In this way, a text is created step by step with several writing and revision phases. In class, exchanges about writing processes and strategies, as well as about written work, often cannot take place to a desirable extent due to limited resources and the fact that students tend to work in isolation.

Meanwhile, collaborative or cooperative writing is useful not only for the acquisition of genre competence and further writing competence but also as a fundamental skill. Herein lies the value of text-generating AI tools such as ChatGPT. According to the data, ChatGPT's genre competence is not (yet) fully developed, but it is going in the right direction to eventually adequately facilitate exchanges about writing and written work and initiate discussions about what a particularly typical example of a genre looks like. Working with ChatGPT links language (the surface level), content (the deep structure), and contexts (the extra-linguistic situational embedding) when engaging with genres. Therefore, it has the potential to enable learners to understand, review, and systematically write appropriate texts. In this way, ChatGPT can contribute to the development of genre competence.

## References

- Adamzik, K. (2016). *Textlinguistik. Grundlagen, Kontroversen, Perspektiven* (2nd ed.). De Gruyter.
- Ahn, H. (2012). Teaching writing skills based on a genre approach to L2 primary school students: An action research. *English Language Teaching*, 5(2), 2–16. <http://dx.doi.org/10.5539/elt.v5n2p2>
- Augst, G., Disselhoff, K., Henrich, A., Pohl, T., & Völzing, P.-L. (2007). *Text – Sorten – Kompetenz: Eine echte Longitudinalstudie zur Entwicklung der Textkompetenz im Grundschulalter*. Peter Lang.
- Bachmann, T., & Feilke, H. (eds.). (2014). *Werkzeuge des Schreibens. Beiträge zu einer Didaktik der Textprozeduren*. Fillibach bei Klett.
- Bachmann-Stein, A. (2021). Die Textsorte konventionelles Kondolenzschreiben. In C. Braun (ed.), *Sprache des Sterbens – Sprache des Todes. Linguistische und interdisziplinäre Perspektivierungen eines zentralen Aspekts menschlichen Daseins* (pp. 15–40). De Gruyter. <https://doi.org/10.1515/9783110694734-002>
- Björk, L. (2003). Text types, textual consciousness and academic writing ability. In L. Björk, G. Bräuer, L. Rienecker, & P.S. Jörgensen (Eds.), *Teaching academic writing in European higher education* (pp. 29–40). Springer. [https://doi.org/10.1007/0-306-48195-2\\_3](https://doi.org/10.1007/0-306-48195-2_3)
- Brinker, K. (2005). *Linguistische Textanalyse: Eine Einführung in Grundbegriffe und Methoden* (6th ed.). Erich Schmidt Verlag. (Original work published 1985)
- Brinker, K., Cölfen, H., & Pappert, S. (2018). *Linguistische Textanalyse: eine Einführung in Grundbegriffe und Methoden* (9th ed.). Erich Schmidt Verlag.
- Brock, A. & Schildhauer, P. (2017). Communication form: A concept revisited. In A. Brock & P. Schildhauer (eds.), *Communication Forms and Communicative Practices* (pp. 13–43). Peter Lang. <https://doi.org/10.3726/978-3-653-06384-4>
- Brommer, S. (2018a). *Sprachliche Muster in wissenschaftlichen Texten: Eine induktive korpuslinguistische Analyse*. De Gruyter.
- Brommer, S. (2018b). Textsortenspezifische sprachliche Variation ermitteln. Muster und Musterhaftigkeit aus korpuslinguistischer, textlinguistischer sowie stilistischer Perspektive. In K. Adamzik & M. Maselko (eds.), *Variationslinguistik trifft Textlinguistik* (pp. 61–81). Narr Francke Attempto Verlag.

- Brommer, S. (2019a). Die Musterhaftigkeit eines Textes als Grundlage seiner Beurteilung – Korpuslinguistik im Dienste der Schreibdidaktik. In A. Hirsch-Weber, C. Loesch, & S. Scherer (eds.), *Forschung für die Schreibdidaktik: Voraussetzung oder institutioneller Irrweg?* (pp. 145–165). Beltz Juventa.
- Brommer, S. (2019b): Empirisch fundierte Sprachkritik – ein Beitrag zur Operationalisierung der vagen Kategorie, Angemessenheit. *Aptum. Zeitschrift für Sprachkritik und Sprachkultur*, 15(2), 123–133.
- Bubenhofer, N. (2009). Sprachgebrauchsmuster: *Korpuslinguistik als Methode der Diskurs- und Kulturanalyse*. De Gruyter. <https://doi.org/10.1515/9783110215854>
- Bubenhofer, Noah (2022). Wie wir in Zukunft wissenschaftliche Texte schreiben (könnten) – Teil 2. <https://www.bubenhofer.com/sprechtakel/2022/12/18/wie-wir-in-zukunft-wissenschaftliche-texte-schreiben-koennten-teil-2/>
- Esterl, U., & Krieg-Holz, U. (eds.). (2018). *Textmuster und Textsorten*. Themenheft Informationen zur Deutschdidaktik (ide). Zeitschrift für den Deutschunterricht in Wissenschaft und Schule. StudienVerlag.
- Fandrych, C., & Thurmair, M. (2011). *Textsorten im Deutschen. Linguistische Analysen aus sprachdidaktischer Sicht*. Stauffenburg Verlag.
- Feilke, H. (1990): Erörterung der Erörterung. Freies Schreiben und Musteranalyse. *Praxis Deutsch*, 99, 52–56.
- Feilke, H. (2012). Was sind Textroutinen? Zur Theorie und Methodik des Forschungsfeldes. In H. Feilke, & K. Lehnen (eds.). *Schreib- und Textroutinen. Theorie, Erwerb und didaktisch-mediale Modellierung* (pp. 1–31). Peter Lang. <https://doi.org/10.3726/978-3-653-01844-8>
- Fischer, C. (2009). *Texte, Gattungen, Textsorten und ihre Verwendung in Lesebüchern* [Dissertation]. Justus-Liebig-Universität Gießen.
- Fix, U. (2008). Ansprüche an einen guten (?) Text. *Aptum. Zeitschrift für Sprachkritik und Sprachkultur*, 4(1), 3–22. [https://doi.org/10.46771/9783967691283\\_1](https://doi.org/10.46771/9783967691283_1)
- Fix, U. (2019). Text und Textlinguistik. In N. Janich (ed.), *Textlinguistik. 15 Einführungen und eine Diskussion* (2nd ed., pp. 17–34). Gunter Narr Verlag.
- Freudenberg-Findeisen, R. (ed.). (2016). *Auf dem Weg zu einer Textsortendidaktik. Linguistische Analysen und text(sorten)-didaktische Bausteine nicht nur für den fremdsprachlichen Deutschunterricht*. Georg Olms.
- Göpferich, S. (2015). *Text competence and academic multiliteracy. From text linguistics to literacy development*. Narr Verlag.

- Heinemann, W. (2000). Textsorte – Textmuster – Texttyp. In K. Brinker, G. Antos, W. Heinemann, & S. F. Sager (eds.), *Text- und Gesprächslinguistik*, 1. Halbband (pp. 507–523). De Gruyter Mouton.
- Hutz, M. (2021). Schreiben mit dem Genre-Ansatz fördern. Textsortenkompetenz erwerben, Schreibschwierigkeiten überwinden. *Der Fremdsprachliche Unterricht*. Englisch 170, 2–8.
- Hyland, K. (2007). Genre pedagogy: Language, literacy and L2 writing instruction. *Journal of Second Language Writing*, 16(3), 148–164.
- Kaiser, D. (2010). Wissenschaftliche Textsortenkompetenz für deutsche und internationale Studierende. In H. Brandl, S. Duxa, G. Leder, & R. Riemer (eds.), *Ansätze zur Förderung akademischer Schreibkompetenz an der Hochschule* (pp. 11–26). Universitätsverlag Göttingen.
- Klemm, Albrecht/Rahn, Stefan/Riedner, Renate (2012). Die Rezension als studentische Textart zur Einübung von zentralen wissenschaftssprachlichen Handlungen. *Informationen Deutsch als Fremdsprache* 39 (4), 405–435. <https://doi.org/10.1515/infodaf-2012-0404>.
- Krieg-Holz, U. & Bülow, L. (2016). *Linguistische Stil- und Textanalyse: Eine Einführung*. Narr Francke Attempto Verlag.
- Lehnen, K. & K. Schindler (2010). Berufliches Schreiben als Lernmedium und -gegenstand. Überlegungen zu einer berufsbezogenen Schreibdidaktik in der Hochschullehre. In T. Pohl & T. Steinhoff (eds.): *Textformen als Lernformen* (pp. 233–256). Gilles and Francke.
- Locher, Franziska Maria/Philipp, Maik (2023). Measuring reading behavior in large-scale assessments and surveys. *Frontiers in Psychology* 13, <https://doi.org/10.3389/fpsyg.2022.1044290>.
- Lomborg, S. (2014). *Social media, social genres: Making sense of the ordinary*. Routledge.
- Nünning, A. (2008). Textsortenkompetenzen. In V. Nünning (ed.), *Schlüsselkompetenzen: Qualifikationen für Studium und Beruf* (pp. 91–104). J.B. Metzler. <https://doi.org/10.1007/978-3-476-05226-1>
- Philipp, M. (2015). *Lesestrategien. Bedeutung, Formen und Vermittlung*. Weinheim: Beltz
- Pitz-Klauser, P. (2019). *Analysieren, interpretieren, argumentieren: Grundlagen der Textarbeit fürs Studium*. Narr Francke Attempto Verlag.
- Renkema, J. & Schubert, C. (2018). *Introduction to Discourse Studies*. John Benjamins.
- Rezat, S. & Feilke, H. (2018). Textsorten im Deutschunterricht. Was sollten LehrerInnen und SchülerInnen können und wissen? *Informationen zur deutschdidaktik* (ide), 2, 24–38.



- Sandig, B. (2000). Text als prototypisches Konzept. In M. Mangasser-Wahl (ed.), *Prototypentheorie in der Linguistik. Anwendungsbeispiele – Methodenreflexion – Perspektiven* (pp. 93–112). Stauffenburg.
- Schicker, S. & Akbulut, M. (2023). ChatGPT – maschinelle und menschliche Textsortenkompetenz. In S. Schicker & L. Saletović (eds.) (2023). *Sprachliche Handlungsmuster und Text(sorten)kompetenz. Ein Sammelband im Rahmen der IDT 2022* (pp. 169–196). Library Publishing University Graz. <https://doi.org/10.25364/9783903374263>.
- Schicker, S. & Saletović, L. M. (eds.) (2023). *Sprachliche Handlungsmuster und Text(sorten)kompetenz. Ein Sammelband im Rahmen der IDT 2022*. Library Publishing University Graz. <https://doi.org/10.25364/9783903374263>.
- Schneider, J. G., & Zweig, K. A. (2023). Grade prediction is not grading: On the Limits of the e-rater. In R. Groß & R. Jordan (eds.), *KI-Realitäten. Modelle, Praktiken und Topologien maschinellen Lernens* (pp. 93–112). transcript Verlag. <https://doi.org/10.14361/9783839466605-005>
- Searle, J. R. (1976). A classification of illocutionary acts. *Language in Society*, 5(1), 1–23. <https://doi.org/10.1017/s0047404500006837>
- Siepmann, D. (2017a). Textsortenwissen. In S. Schierholz & L. Giacomini (eds.), *Wörterbücher zur Sprach- und Kommunikationswissenschaft (WSK) Online*. De Gruyter. [https://www.degruyter.com/database/WSK/entry/wsk\\_idf9f4f6aa-4af9-4dc5-8ee7-418ee3aa250a/html](https://www.degruyter.com/database/WSK/entry/wsk_idf9f4f6aa-4af9-4dc5-8ee7-418ee3aa250a/html).
- Siepmann, D. (2017b). Textsortenkompetenz. In S. Schierholz & L. Giacomini (eds.), *Wörterbücher zur Sprach- und Kommunikationswissenschaft (WSK) Online*. De Gruyter. [https://www.degruyter.com/database/WSK/entry/wsk\\_idb483c16e-ad58-45ca-b550-e107afbb4d9a/html](https://www.degruyter.com/database/WSK/entry/wsk_idb483c16e-ad58-45ca-b550-e107afbb4d9a/html)
- Steinhoff, T. (2007). Wissenschaftliche Textkompetenz: Sprachgebrauch und Schreibentwicklung in wissenschaftlichen Texten von Studenten und Experten. Max Niemeyer. <https://doi.org/10.1515/9783110973389>
- Stöckl, H. (2015). From text linguistics to multimodality. Mapping concepts and methods across domains. In J. Wildfeuer (ed.), *Building bridges for multimodal research. International perspectives on theories and practices of multimodal analysis* (pp. 51–75). Lang.
- Susteck, S. & Perder, C.) (2023). Schreiben durch Künstliche Intelligenz. ChatGPT und automatisierte Lyrikanalysen. *Medien im Deutschunterricht*, 5(2). <https://journals.ub.uni-koeln.de/index.php/midu/article/view/1970/2040>
- Swales, J. & Feak, C. (2012). *Academic writing for graduate students: Essential tasks and skills* (3rd ed.). University of Michigan Press.

- Trosborg, A. (1997). Text typology: Register, genre and text type. In A. Trosborg (ed.), *Text Typology and Translation* (pp. 3–23). John Benjamins.
- Van Dijk, T. A. (1972). Foundations for typologies of texts. *Semiotica*, 6(4), <https://doi.org/10.1515/semi.1972.6.4.297>.
- Wendt, C., Burhfeind, I., Frick, K., & Neumann, A. (2023). Mit KI im Deutschunterricht schreiben – Impulse für Lehrer\*innen für den Unterricht in der Zukunft. *k:ON – Kölner Online Journal Für Lehrer\*Innenbildung*. *k:ON - Kölner Online Journal* (7), 321–340.
- Williamson, P. A. (2021). *Academic writing skills*. The University of Queensland. <https://doi.org/10.14264/54e8047>
- Zhao, W. X., Zhou, K., Li, J., Tang, T., Wang, X., Hou, Y., Min, Y., Zhang, B., Zhang, J., Dong, Z., Du, Y., Chen, Y., Chen, Y., Chen, Z., Jiang, J., Ren, R., Li, Y., Tang, X., Liu, Z.,... Wen, J. (2023). A survey of large language models. *arXiv* (Cornell University). <https://doi.org/10.48550/arxiv.2303.18223>

## Appendix

### Student Assignment (original in German)

Wir sind in die Schreibexperimente-Phase gestartet. In dieser Zeit sollen Sie ausprobieren, wie gut ChatGPT Folgendes kann:

- \ Texte klassifizieren
- \ Texte analysieren
- \ Texte kondensieren/zusammenfassen
- \ Texte generieren

Damit sich die Experimente besser vergleichen lassen, sind im Folgenden ein paar **Vorgaben zum Vorgehen** formuliert. Bitte beachten Sie diese bei der Durchführung Ihrer Experimente.

**Grundsätzliches:**

- \ Wichtig: Testen Sie die vier oben genannten Funktionen (1. klassifizieren, 2. analysieren, 3. kondensieren/zusammenfassen, 4. generieren) jeweils für sich separat, d. h. starten Sie jeweils einen neuen Dialog.
- \ Wählen Sie vorab drei bis fünf Originaltexte (= vom Menschen verfasste Texte) Ihrer gewählten Textsorte als Ausgangstexte aus.

**Klassifizieren testen**

- \ Starten Sie Ihren Dialog zum Klassifizieren mit folgendem Prompt: “Ich gebe Dir gleich einen Text. Um welche Textsorte handelt es sich dabei und wie kommst Du zu dieser Einschätzung?”
- \ Testen Sie diesen Prompt mit weiteren Originaltexten.
- \ Fragen Sie ChatGPT nach der Prototypizität der jeweiligen Texte und Unterschieden und Gemeinsamkeiten.

**Analysieren testen**

- \ Starten Sie Ihren Dialog zum Analysieren mit folgendem Prompt: “Ich gebe Dir gleich einen Text. Analysiere den Text hinsichtlich seiner Kohärenz.”
- \ Fahren Sie mit folgendem Prompt fort: “Analysiere den Text auf sprachlich-stilistischer Ebene.”
- \ Testen Sie diese beiden Prompts mit weiteren Originaltexten.
- \ Formulieren Sie Prompts nach Belieben, um die Analysefähigkeit von ChatGPT zu testen.

**Kondensieren/zusammenfassen testen**

- \ Starten Sie Ihren Dialog zum Kondensieren/Zusammenfassen mit folgendem Prompt: “Ich gebe Dir gleich einen Text. Bitte kürze ihn um ca. 30 Prozent auf zwei Drittel seiner Länge.”
- \ Fahren Sie mit folgendem Prompt fort: “Fasse die Hauptaussage des ursprünglichen Textes in eigenen Worten in einem Satz zusammen.”
- \ Testen Sie diese beiden Prompts mit weiteren Originaltexten.
- \ Formulieren Sie weitere Prompts, in denen Sie ChatGPT bitten, verschiedene Aspekte zu kürzen oder zusammenzufassen.

**Generieren testen:**

- \ Starten Sie Ihren Dialog zum Generieren mit folgendem Prompt: “Generiere mir ein Beispiel für eine Danksagung/Anzeige/Rezension [Ihre Textsorte].”

- \ Fahren Sie mit folgendem Prompt fort: “Verfasse ein weiteres Beispiel dieser Textsorte.”
- \ Fragen Sie als nächstes: “Worauf ist zu achten, wenn man eine [Ihre Textsorte] verfasst?”
- \ Führen Sie den Dialog nach Belieben fort und überlegen Sie, welche Prompts Ihnen helfen herauszufinden/zu prüfen, wie gut ChatGPT beim Generieren Ihrer Textsorte ist.

---

**Date received:** November 2023

**Date accepted:** April 2024