



IRAQI
Academic Scientific Journals



العراقية
المجلات الأكاديمية العلمية



ISSN: 2663-9033 (Online) | ISSN: 2616-6224 (Print)

Journal of Language Studies

Contents available at: <https://jls.tu.edu.iq/index.php/JLS>

A Digital Stylistic Exploration of AI-Generated Text: Prompt Engineering for Journalistic and Legal Genres in ChatGPT-4

Asst. Prof. Dr. Saza Ahmed Fakhry Abdulla*

Department of English, College of Languages, University of Sulaimani, Sulaimaniyah, Iraq
saza.fakhry@univsul.edu.iq

&

Bwar Mustafa Ali

Department of English, College of Languages, University of Sulaimani, Sulaimaniyah, Iraq
bwar.04000401@univsul.edu.iq

Received: 10/12/2025, Accepted: 25/12/2025, Online Published: 30/06/2026

Abstract

In this study, a formal and textual stylistic approach was employed to investigate specific prompts in journalistic and legal genres that were provided to Generative Pre-trained Transformer-4 (ChatGPT-4) for generating prompts. ChatGPT-4 generates three prompts: minimal, tone-specified, and constraint-based. Each one of these prompts includes a prompt, an output, and a stylistic analysis. The purpose is to determine how prompt design

* **Corresponding Author:** Saza Ahmed Fakhry, Email: saza.fakhry@univsul.edu.iq.

Affiliation: University of Sulaimani – Iraq.

©This is an open access article under the CC by licenses <http://creativecommons.org/licenses/by/4.0>



influences stylistic variation, evaluate the AI's capacity for self-analysis, and highlight the importance of prompt engineering. Leech and Short's (2007) stylistic checklist of lexical, grammatical categories, and cohesion has been used as a theoretical framework for analysing the AI outputs. In the process of data analysis, a qualitative research method was used to evaluate ChatGPT-4's ability to generate content within the two selected genres to perform a stylistic analysis, and evaluate its outputs against recognised standards of the stylists' toolkit (Leech and Short 2007, Leech et al 1989). In the analysis, the following linguistic features are used: lexical features (nouns: abstract and concrete, verbs, adjectives), grammatical features: sentence (complexity, length, elements, types, structure), tense, and voice; Textual features: punctuation, specialised vocabulary, emotive lexis, word choice, and tone. The findings show that the legal genre had longer and more syntactically complex sentences, more lexical abstraction, a higher frequency of present-tense verbs, specialised vocabulary, and formal word choice. The journalistic genre, on the other hand, is characterised by a greater use of past-tense verbs, shorter syntactic structures, and a reportorial tone. Although declarative sentences and an active voice were used in both genres, legal prompts were distinct in that they contain a greater variety of authoritative tones and foregrounding punctuation. While AI is capable of identifying surface-level features, prioritising tone and genre over grammatical detail, it was less accurate and inadequate. It disregards detailed linguistic categorisations by methodically analysing lexical choice, syntactic complexity, and textual cohesion; human analysis demonstrates superior accuracy. Ultimately, ChatGPT-4 often replicates the style of stylistic critique, whereas stylists validate their conclusions through explicit linguistic evidence, particularly in abstract domains where linguistic validity predominates.

Keywords: ChatGPT-4, digital stylistics, journalism, law, prompt engineering

الأسلوبية الحاسوبية للنصوص المولدة بالذكاء الاصطناعي: هندسة المطالبات للنصوص الصحفية والقانونية

بواسطة المحوّل التوليدي المُدرَّب مُسبقاً - الإصدار الرابع

أ. م. د. سازه احمد فخري عبد الله
قسم اللغة الإنكليزية، كلية اللغات، جامعة السليمانية

و

بوار مصطفى علي
قسم اللغة الإنكليزية، كلية اللغات، جامعة السليمانية

المستخلص

في هذه الدراسة، استُخدم منهج شكلي ونصي للتحليل الأسلوبي في البحث على المطالبات المحددة من قبل المحوّل التوليدي المُدرَّب مُسبقاً - الإصدار الرابع (جات جي بي تي 4) في النصوص الصحفية والقانونية. جات جي بي تي 4 صمم ثلاثة أنواع من المطالبات منها: الأدنى، ومحددة النبرة، والقائمة على القيود. يتضمن كل منها الأمر، المخرجات، والتحليل الأسلوبي. تهدف هذه الدراسة إلى تحديد كيفية تأثير تصميم المطالبات على التباين الأسلوبي، وتقييم قدرة الذكاء الاصطناعي على التحليل الذاتي، وتبسيط الضوء على أهمية هندسة الأوامر. اعتمدت الدراسة على الإطار النظري لـ ليتش وشورت (2007) للتحليل الأسلوبي والتي تشمل الفئات المعجمية، والنحوية، وعناصر التماسك النصي لتحليل مخرجات الذكاء الاصطناعي. عند عملية تحليل البيانات، استُخدم منهج البحث النوعي لتقييم قدرة نموذج جات جي بي تي على توليد المحتوى ضمن النصيين المختارين، وإجراء تحليل أسلوبي لهما، بالإضافة إلى تقييم مخرجاته في ضوء المعايير المعتمدة ضمن مجموعة أدوات وضعت من قبل علماء الأسلوبية منهم، ليتش وشورت (2007) وليتش وآخرون (1989). اعتمد التحليل على السمات اللغوية التالية: السمات المعجمية (الأسماء: المجردة والمادية، الأفعال، الصفات)؛ والسمات النحوية: الجملة (من حيث التعقيد، والطول، والعناصر، والأنواع، البنية)، والزمن، والصيغة) المبني للمعلوم والمجهول)؛ والسمات النصية: علامات التنقيط، والمفردات المتخصصة، والكلمات العاطفية، واختيار الألفاظ، والنبرة. وأظهرت النتائج أن النص القانوني تميز بجملة أطول وأكثر تعقيداً من الناحية التركيبية، مع نزوع أكبر نحو التجريد المعجمي، وكثرة في استخدام الأفعال المضارعة، ومفردات متخصصة، واختيار رسمي للألفاظ. وفي المقابل، النص الصحفي إلى استخدام أفعال الماضي، والجملة الأقصر والأكثر بساطة، مع تبني نبرة تقريرية. ورغم استخدام الجمل الخبرية وصيغة المبني للمعلوم في كلا النصيين، إلا أن مخرجات الأوامر القانونية كانت متميزة باحتوائها على تنوع أكبر في النبرات السلطوية واستخدام علامات التنقيط بشكل انحرافي لأغراض الإبراز الأسلوبي. مع ذلك، ظل التحليل الأسلوبي الذاتي الذي ولدته جات جي بي تي 4 سطحيًا وغير كافٍ، حيث غلبت النبرة ونوع النص على التفاصيل النحوية، وكانت أقل دقة من تحليل المحترفين والمتخصصين في علم الأسلوبية الذين يقدمون تصنيفاً أدق وأكثر تفصيلاً للسمات الأسلوبية مثل التراكيب النحوية، وأنواع الأفعال، وتجريد الأسماء والتماسك النصي. وفي الختام، يكتفي جات

جي بي تي 4 غالباً بتقليد اسلوب النقد الأسلوبي، بينما يوثق الأسلوبيون استنتاجاتهم عبر أدلة لغوية صريحة، لا سيما في المجالات المجردة التي تفرض فيها الصلاحية اللغوية سيادتها.
الكلمات المفتاحية: المحوّل التوليدي المُدرَّب مُسبقاً - الإصدار الرابع 4، الأسلوب الرقمي، الصحافة، القانون، الهندسة الفورية

1. Introduction

The emergence of advanced large language models (LLMs) such as ChatGPT-4 marks a turning point in digital communication and professional writing, enabling coherent texts that imitate journalistic reporting and legal exposition. However, a critical underexplored challenge lies in LLMs' capacity to autonomously generate effective prompts for stylistic adaptation and perform expert-level stylistic analysis of their own outputs, essential for reliable professional use yet absent from prior research. Previous research has effectively documented the surface-level stylistic imitation of LLMs, demonstrating their ability to repeat the formal conventions of genres such as news articles and legal documents (Dale, 2021) and even the specific stylistic fingerprints of individual authors (Ippolito et al., 2022). This limitation highlights that LLM outputs, while useful for surface-level stylistic tasks, require human oversight for analyses demanding fine-grained grammatical and semantic evaluation.

Evlin et al. (2019) refer to foundational BERT work, which established bidirectional context understanding as key for tasks such as masked language modeling. Tikhonov et al. (2023) explore prompt engineering techniques to optimise LLM outputs, building on earlier zero-shot methods. Their analysis shows how refined prompts enhance coherence and task performance, following instructions in a way comparable to humans. Ashwell (2022) studied AI-human stylometric comparisons in creative fields, showing AI's ability to generate text, but did not consider AI's self-prompting or self-analysis in professional areas such as journalism and law, leaving an important research gap.

This study addresses that gap through Digital Stylistics by investigating ChatGPT-4's dual role in generating targeted prompts: minimal, tone-specified, and constraint-based, and conducting stylistic analysis. ChatGPT-4's outputs are rigorously compared with analyses conducted by professional human stylisticians. The comparison evaluates differences in depth, accuracy, and detail, with particular attention to how ChatGPT-4 adapts lexical choices, grammatical categories, and textual features. Building on this framework, the study focuses on the following two research questions:

1. How can ChatGPT-4's outputs in journalistic and legal genres be practically and efficiently analysed from a stylistic perspective?
2. To what extent does the stylistic analysis generated by ChatGPT-4, under optimised prompting, meet the standards of expert human stylisticians' analyses in terms of accuracy, interpretive depth, and comprehensiveness?

The originality of this study lies in its examination of ChatGPT-4's dual capacity to generate targeted prompts and to conduct stylistic analysis within journalistic and legal genres that remain underrepresented in AI stylistics research. It further distinguishes itself through a rigorous comparison of AI-generated analyses with those produced by professional human stylisticians, quantitatively assessing differences in depth, accuracy, and level of detail.

2. Theoretical Framework

The theoretical framework of this study is based on two main parts: (a) LLMs' stylistic generation influenced by prompt design, and (b) stylistics as a framework for analysing and comparing AI- and human-produced texts.

2.1 Prompt Design and Stylistic Variation in Language Models

ChatGPT-4 is a large language model that generates text by predicting word sequences based on patterns learned from its training data (Brown et al., 2020). Its linguistic output is therefore highly dependent on the phrasing and specificity of the provided prompt. Studies on prompt engineering demonstrate that the structure and constraints of a prompt determine the model's tone, lexical range, and grammatical complexity (Reynolds & McDonell, 2021; Jurafsky & Martin, 2023). This relationship explains why the present study employs three prompt types: minimal, tone-specified, and constraint-based to examine how increasing levels of instruction affect stylistic accuracy and genre consistency.

Previous research indicates that minimal prompts usually produce general or less refined stylistic features, while tone-specific or constraint-based prompts generate outputs that are more controlled and appropriate to the target genre (Das, Sanyal, & Chakraborty, 2024). Building on this, prompt variation is used in this study as an experimental tool to explore ChatGPT's stylistic flexibility. Three types of prompts are examined: minimal, which provide no stylistic guidance; tone-specified, which indicate a desired communicative tone

(e.g., critical, hopeful, neutral); and constraint-based, which set explicit rules for content, format, genre, or structure (e.g., courtroom monologue, parable).

2.2 Stylistics and Digital Stylistics

Stylistics can be defined as the "linguistic study of style, is rarely undertaken for its own sake, simply as an exercise in describing what use is made of language" (Leech and Short, 2007, p.11). focusing on how linguistic choices create communicative effects. Bradford states that stylistics offers a thorough analysis "of the text in hand and as a result, deepens the process of interpretation and evokes critical thinking" (1997, p. xii). According to Short, stylistics is 'a method of linking linguistic form, via reader inference, to interpretation in a detailed way and thereby providing as much explicit evidence as possible for and against particular interpretations of texts' (1996, p. 27).

The basic objective of stylistics lies in the explanation of 'how readers get from the words of a text (a) an understanding of it and (b) a felt response to it' (Leech and Short, 2007: 289, cited in Abdulla, 2016). **Digital stylistics** is an emerging multidisciplinary field that combines traditional stylistic analysis (the study of linguistic style in literary and non-literary texts) with digital methods and computational tools to analyse textual patterns, linguistic features, and stylistic variations in digital and digitised texts. It is defined as a field that uses computational and statistical methods to analyse literary style and language use, focusing on relationships between stylistic function and linguistic choices (Hesselbach et al., 2024).

Leech and Short (2007: 61) classify the stylistic checklist into four categories: **Lexical Categories**, which can be classified into a 'General' category to explain whether the vocabulary is 'simple or complex, descriptive or evaluative, specific or general, formal, informal' or whether the text shows any 'rare use of specialised vocabulary'. Nouns are classified into concrete and abstract. Concrete nouns are defined as those that are detected by the senses of perception and could be 'observable and measurable'. In contrast, abstract nouns refer to those nouns that are 'non-observable' and 'non-measurable' (Quirk, Greenbaum, Leech & Svartvik, 1985: 247). Some concrete nouns refer to people, things, animals, and places, all of which are used as countable nouns. Other concrete nouns refer to substances, liquids, and gases, which are uncountable nouns (Leech, 1989: 90-91).

Abstract nouns are those nouns that have ‘no physical form’ and cannot be seen or touched. They can stand for ‘general feelings, ideas and concepts’, qualities, states, events and actions, and a number of them are recognised by ‘their endings’ such as ‘ion’, ‘al’, ‘sion’, ‘tion’ (Leech 1989: 9-11). They classify adjectives as (physical, psychological, visual, auditory, describing colour, referential, emotive, evaluative, etc.). They also classify verbs as states, actions or events, movement, physical acts, Speech Acts, psychological states or activities, perceptions, transitive or intransitive, factive or non-factive. According to Leech (1989: 20), adverbs are classified according to meaning: manner, place, time, degree, frequency, linking, comment or attitude, adding and limiting, viewpoint, and length of time.

Grammatical categories involve six types: sentence types include declarative, negative, exclamatory, interrogative, and imperative. Sentence complexity involves a simple or complex structure, the average sentence length (in number of words), the ratio of dependent to independent clauses. Clause types such as relative clauses, adverbial clauses, and different types of nominal clauses (*that*-clauses, *wh*-clauses, etc.). Clause elements such as objects, complements, and adverbials of a transitive or intransitive verb. They also refer to the uses of tense: the simple past tense, occurrences and functions of the present tense; modal auxiliaries such as *can*, *shall*, *must*, *would*, etc. The third category is cohesion, which involves the text containing logical or other links between sentences that tend to rely on implicit connections of meaning, and the avoidance of repetition. These selected stylistic features are consistent with widely recognised stylistic methodologies employed to analyse levels of formality and domain-specific language use in professional genres (Leech & Short, 2007; Hoover, Culpeper, & O’Halloran, 2023).

Halliday & Hasan (1989) define **specialised vocabulary** as lexical items restricted to particular registers or discourse communities, whose meanings are shaped by professional, academic, or institutional contexts. It is also referred to as domain-specific or technical vocabulary. Simpson (2004) sees **word choice** as part of the lexical level of stylistic analysis, shaping interpretation and discourse function. Simpson situates **emotive lexis** within evaluative language that shapes reader alignment and interpretive stance. Short defines **tone** as a stylistic effect emerging from “patterns of linguistic choices which signal attitude” (1996). This study aims to examine ChatGPT-4’s self-analysis of its own outputs and contrasts it with stylisticians’ analytical examination. This comparison highlights the

limits of AI interpretive ability and underlines the need for human expertise in detailed stylistic validation.

3. Methodology

In this work, a qualitative research method is used to analyse ChatGPT-4 generated text prompts of two different genres. The **data** has been **collected** from ChatGPT-4 to create journalistic and legal prompts.

Rationale for data selection

The **data** have been **selected** from these two genres because they are interconnected institutional fields that shape public knowledge through authoritative language. Both rely on shared stylistic features such as objectivity, structured argumentation, and controlled lexical choices, while serving different communicative purposes. Their inclusion allows an assessment of whether ChatGPT-4 can generate genre-appropriate legal and journalistic texts. For this reason, choosing them is not random but based on their shared responsibility to handle information carefully and accurately.

Research Procedures and Steps:

During the process, the following steps were undertaken:

1. For each genre, three prompts were designed.
2. These prompts were provided to ChatGPT-4 to generate responses, as outlined below:
 - a. For the journalistic and legal genres, write three example prompts. Then, generate an output for each prompt within each discipline.
 - b. Provide a thorough stylistic analysis for each of the three generated outputs per genre.
3. All responses were saved and categorised according to discipline, prompt variation, and AI-generated output type.
4. For each genre, ChatGPT-4 generated three types of responses: minimal, tone specified, and genre-constraint. Each AI-generated response includes the prompt, the output, and a stylistic analysis.
5. The outputs and the stylistic analysis in the minimal, tone-specified, and genre constraint prompts were analysed stylistically by the researchers first and compared with the stylistic analysis provided by ChatGPT-4, followed by some findings.

6. The theoretical framework of Leech and Short (2007) and Leech et al (1983) is adopted.
7. The researchers performed a manual stylistic analysis of AI-generated textual outputs and analysed the following selected stylistic features:
 - Lexical:** nouns (concrete and abstract), verbs, adjectives, and adverbs.
 - Grammatical:** sentence (complexity, length, elements, types, structure), and tense.
 - Textual:** punctuation, specialised vocabulary, emotive lexis, word choice, and tone.
8. Examples from the data are presented within the analysis.

4. Analysis

The analysis focuses exclusively on the outputs generated by ChatGPT in minimal, tone-specified, and genre-constraint prompts across two selected professional genres.

Additionally, ChatGPT's stylistic analysis will be re-analysed and reassessed by the researchers, who are specialised in stylistics.

In the **Journalistic Prompts**, ChatGPT 4's generated minimal prompt is "*Report on a train accident.*" Based on the prompt that ChatGPT 4 generated, the output was: "*A freight train derailed Tuesday evening in northern Ohio. No fatalities were reported. Investigations are ongoing.*". Regarding **lexical features**, the majority of nouns are abstract, such as *Tuesday evening*, *fatalities*, and *Investigations*, and a *freight train* is a concrete noun. The verbs are *derailed*, *reported*, and *are*. These verbs are dynamic, stative, and copular verbs, respectively. In relation to **grammatical features**, the output response consists of three short and simple sentences. The length of the first sentence contains nine words and follows an SVA structure, where *A freight train* functions as the subject, *derailed* as the verb, and *Tuesday evening in northern Ohio* as the adverbial. The second sentence contains four words, which consist of the subject *No fatalities*, with a negative particle and the passive verb phrase 'were reported. The third sentence contains three words and includes the subject *Investigation*, the present tense verb *are*, and the complement *ongoing*. All three sentences are declarative and exhibit structural simplicity. There is subject-verb agreement in both number and person in the clauses *No fatalities were* and *Investigations are*. The plural suffix -s is correctly used in the nouns *fatalities* and *investigations*. The tense used is the simple past. Regarding **textual features**, the vocabulary is characterised

by specialised lexis, particularly related to journalistic genres such as *as reported* and *investigations*. No emotive lexis is shown. The word choices are simple, clear, and concise, and less formal. The tone is optimistic and supportive. No cohesive ties are used since the three sentences could be connected by using the coordinator *and* in the second sentence and the subordinator in the third one.

ChatGPT-4 generated the following **stylistic analysis**: “*The writing adheres to the inverted pyramid style: key facts front-loaded, supporting details follow. Lexis is factual, with temporal and geographical specificity. The passive construction 'were reported' implies journalistic objectivity. Sentence lengths are short, supporting skimmability. No emotive or evaluative language is present*”. ChatGPT-4’s analysis provides no detailed examination of lexical features, such as nouns, verbs, etc., word choices, sentence length, types, and complexity, grammatical structures such as tenses, sentence elements, word choice, and textual features such as specialised vocabulary, tone, and cohesive ties as analysed above.

In a **tone-specified prompt**. ChatGPT 4 generated “*Report on government corruption in a restrained but skeptical tone. Based on this prompt, the output was: “The ministry’s procurement process, while technically legal, raises questions of transparency and ethical oversight.”*”

With respect to **lexical features**, the majority of nouns are abstract, such as *procurement process, questions, transparency, and oversight*, with one concrete noun, *ministry*. The verb *raise* is stative, and the two adjectives, *legal* and *ethical*, are evaluative. Regarding **grammatical features**, the output response comprises a complex sentence containing one independent clause: *the ministry’s procurement process raises questions of transparency and ethical oversight*, and a dependent clause, *while technically legal*. The average sentence length is 14 words. The sentence has a subject, verb, complement, and adverbial order, where *the ministry’s procurement process* functions as the subject, *raises* as the verb, and *questions* as the object, and *of transparency and ethical oversight* as the adverbial. The sentence is declarative in type and structurally complex. The sentence is in the present tense and active in voice. In the use of **textual features**, the vocabulary is characterised by specialised governmental lexis, particularly related to journalistic genres: *procurement process, technically, and transparency*. These words are frequently used in the news. Word

choices exhibit formality, namely *procurement*, and *lacks transparency*, which are favoured by Chat GPT 4, instead of "*buying things*" and *seem shady*, respectively. No overt emotive lexis is shown. The tone is authoritative and neutral. The sentence is cohesively tied since the coordinator *but* and the subordinator *while* are used.

ChatGPT-4 generated the following **stylistic analysis**: *This sentence relies on qualification and subtle critique. The concessive clause ("while technically legal") uses hedging to suggest concern without outright accusation. The noun phrase "questions of transparency" reflects bureaucratic diction, maintaining professionalism while hinting at controversy. This is typical of editorial reporting or watchdog journalism.* ChatGPT-4 provides no detailed analysis, and drills down into the formal stylistic features, such as lexical features: nouns, verbs, adjectives, etc., sentence types, sentence complexity, grammatical features such as tenses, sentence elements, and textual, including word choices, specialised vocabulary, and tone as shown above.

In a **constraint-based prompt**, ChatGPT 4 generated: "*In four sentences, write an objective news brief on a climate protest.*" Based on this prompt, the output was: "*Activists gathered outside the City Hall at noon. They held banners and gave speeches. Police maintained a perimeter. No arrests were made.*"

With respect to **lexical features**, nouns such as *activists*, *city hall*, *banners*, *arrests*, *police*, and *perimeter* are concrete. One abstract: *speeches*. The verbs mostly refer to dynamic verbs, such as *gathered* and *gave*, and stative verbs, such as *held* and *maintained*. Adverbs of place, such as *outside the City Hall*, and time, such as *at noon*, are noticeably used. Regarding **grammatical features**, the output response contains four sentences: three simple sentences and one compound. The first sentence is simple, and follows a subject, verb, adverbial order, where *Activist* functions as the subject, *gathered* as the verb, and *outside the City Hall at noon* as the adverbial. The second sentence is compound, and has a subject *they*, the two verbs *held* and *gave* are joined by a coordinator, *and* to connect the two nouns *banners* and *speeches*, respectively, where they are the objects of the sentence. The third sentence is simple and has a subject, *police*, a verb, *maintained*, and an object, *a perimeter*. The last sentence is simple and has the subject *No arrests*, and the verb *be* and the past participle of *make* to form a passive construction *were made*. The length of the first sentence is 8, the second is 6, and the third and the last sentence are 4 words. The sentence

type is declarative and simple in structure. Regarding **textual features**, the word choices are simple, clear, uncomplicated, and less formal. The text lacks cohesive devices, as the first and second sentences along with the third and fourth could be joined by the coordinating conjunction 'and' to form compound sentences. The language is objective, devoid of emotional expressions, and avoids the use of descriptive adjectives. The tone is neutral, detached, and reportorial. The four sentences are not cohesively tied due to the lack of connectors to connect the four simple sentences.

ChatGPT-4 generated the following **stylistic analysis** “*The syntax is repetitive, mirroring a subject–verb–object format for all four sentences. The repetition enhances neutrality and clarity. The vocabulary is deliberately bland to avoid bias, focusing strictly on observable actions. The tone is dry and matter-of-fact, suited to newswire or public service bulletins. There is no judgment or narrative voice.*” ChatGPT-4’s analysis does not thoroughly address lexical features: nouns, verbs, adjectives, grammatical features: sentence complexity, length, elements, and textual features such as emotive language, tone, as illustrated earlier.

In the **Legal Prompts**, ChatGPT 4’s generated minimal prompt is “*Explain tort law.*” Based on this prompt, the output was: “*Tort law involves civil wrongs that cause harm to individuals. It includes negligence, intentional harm, and strict liability.*”

In the use of **lexical features**, the output employs predominantly abstract nouns *tort, law, wrongs, harm* (repeated), *negligence*, and *liability*, with a single concrete noun: *individuals*. The **grammatical features** of the output response comprise two clauses: The first clause is structurally complex, containing an independent clause *Tort law involves civil wrongs* accompanied by a dependent subordinate *that*-clause *that cause harm to individuals*. The second clause is compound *It includes negligence, intentional harm, and strict liability* combined with a coordinator *and* to join items: *negligence, intentional harm, and strict liability*. The average sentence length is 10 words in the first sentence and 8 words in the second. The first sentence follows an SVOA (Subject-Verb-Object-Adverbial) structure, where *Tort law* functions as the subject, *involves* as the verb, and *civil wrongs* as the object, *that cause harm to individuals* as the adverbial. The second sentence exhibits an SVC (Subject-Verb-Complement) structure, where the pronoun *It* acts as a subject, *includes* as the verb, and the noun phrase “*negligence, intentional harm, and strict liability*”

functions as the complement. The first sentence's SVOA pattern is simple in terms of structural complexity and maintains a consistent declarative mood. All verbs are stative verbs and are in the present tense: *involves*, *cause*, *includes* refer to facts, while the evaluative adjectives *civil*, *intentional*, and *strict* modify the following nouns for categorising wrongs, *harm*, and *liability*. Regarding **textual features**, the response includes specialised legal lexis, including technical terms *Tort law*, and *Civil wrongs*. The pronoun *it* functions anaphorically, referencing *Tort Law*. Word choices are formal, direct, and unambiguous. There is a lack of emotive language, figurative language. The tone is neutral, serious, and informative.

ChatGPT-4 generated the following **stylistic analysis**: *This text adopts a didactic and impersonal tone suitable for introductory legal explanation. The short sentence length and declarative structure promote readability and clarity. Lexical choices such as “civil wrongs” and “negligence” are domain-specific, reflecting legal terminology without rhetorical elaboration. The passive grammatical structure distances the writer’s voice, supporting the authoritative tone. The absence of modality or metaphor preserves neutrality and objectivity, aligning with the stylistic conventions of legal education materials.* ChatGPT-4’s analysis is limited to a general interpretation of tone and style provides no detailed explanation regarding the stylistic features, of lexical features such as nouns, verbs, etc., sentence types, sentence complexity, grammatical structures such as tenses, sentence elements, as well as textual features: word choices, specialised vocabulary, and tone as shown above.

In a **tone-specified prompt**, ChatGPT 4 generated “*Explain contract breaches in a stern and authoritative tone.*” Based on this prompt, the output was: “A contract is not a suggestion—it is a *binding legal obligation. A breach is a violation of trust and law, subject to restitution or remedy.*”

Regarding **lexical features**, there is an apparent use of abstract nouns such as *contract*, *suggestion*, *obligation*, *breach*, *violation*, *trust*, *law*, *restitution*, and *remedy*. In connection with **grammatical features**, ChatGPT-4’s output consists of two independent clauses, each of which expresses a complete thought. The first sentence includes “*A contract is not a suggestion—it is a binding legal obligation.*”. These sentences consist of three clauses, connected by commas and conjunctions, to maintain the formality of a legal document.

There is also a compound sentence (two independent clauses joined by an em dash: the first clause is *"A contract is not a suggestion,"* and the second clause is *it is a binding legal obligation*. The average sentence length is 12 words in sentence one and 14 in sentence two. The two sentences are declarative in mood and complex in their syntactic arrangement. The sentence *"A breach is a violation of trust and law, subject to restitution or remedy"* is a compound sentence. The main clause is *"A breach is a violation of trust and law,"* is simple and consists of a SVC pattern where *A contract* is the subject, the verb is a copula *is* in the negative form, and *a suggestion* is a complement. The second sentence comes immediately after the dash, starting with a subject *it*, and another linking verb *is*, and *a binding legal obligation* as a complement. The third sentence also includes a SVC pattern where *A breach* is a subject, *is* is a linking verb *a violation of trust and law, subject to restitution or remedy* is a complement. The phrase *subject to restitution or remedy* is a modifying phrase, but not a full clause. The tense is simple present because the AI response deals with facts. There is also a parallel structure in *"A breach is a violation of trust and law"*. The two items have the same syntactic structure, *"trust and law"*.

Regarding **textual features**, the selected nouns are specialised vocabulary and are accurate for a legal genre. There is a logical cohesion of the cause-and-effect relationship between the following pairs: *"contract" → "breach" → "violation" → "restitution/ remedy"*. The output is not cohesively tied since no subordinating conjunctions, such as *however, so, thus*, etc., are used, but rather the clauses are connected with a dash and a comma, besides the coordinating conjunction *and* is used once. The word choices are formal and direct. No emotive lexis is shown. The tone is highly confident, persuasive, and objective.

ChatGPT-4 generated the following **stylistic analysis**: *The tone is intensified through emphatic declaratives and formal legal diction. The em-dash in the first sentence acts as a rhetorical pause, adding dramatic force to the contrast between "suggestion" and "binding legal obligation."* The word choices *"violation," "trust,"* and *"restitution"* carry moral and legal weight. Syntactically, the construction of both sentences uses balanced phrasing and parallel structure, which reinforces the rigid and definitive tone demanded by the prompt. ChatGPT-4's **stylistic analysis** remains somewhat general since it focuses on the tone, which is strengthened through clear uses of declarative sentences and formal

vocabulary. The punctuation adds emphasis to key ideas, while the balanced structure contributes to a firm and authoritative tone. ChatGPT 4 has not accurately examined the formal features, such as lexical features: nouns, verbs, etc., the grammatical features, such as sentence (complexity, length, elements, types), syntactic structure, tenses, voice, and the textual features: word choices (formality), specialised vocabulary, tone, as shown above.

In a **constraint-based prompt**, ChatGPT-4 generated "*Describe the Miranda rights as a courtroom monologue.*" Based on this prompt, the output was: "*Your Honor, before my client uttered a word, the officer should have informed him—clearly—that he had the right to remain silent. That right, fundamental and constitutional, was denied.*"

In the selection of **lexical features**, the nouns are strikingly impressive: *honor, client, office*, the verbs are in the past simple tense: *uttered, informed, had, denied*, and the adjectives are mostly evaluative, *silent, fundamental, and constitutional*. The adverb of manner *clearly* is located between the two dashes. Regarding **grammatical features**, the output consists two sentences: one is complex and another is simple. One main clause: *the officer should have informed him—clearly*. This sentence can stand alone and expresses a complete thought, and another clause: *before my client uttered a word, and that he had the right to remain silent*. The former begins with the subordinating conjunction "*before*" and cannot stand alone, functioning as an adverbial clause of time, while the latter consists of a *that clause* that functions as an adverbial clause. The first sentence has SAOVOA structure, where *Your Honor* is the subject *before my client uttered a word* is an adverbial clause, *the officer* is the object, *should have informed* is the verb in the passive, *him is the object*, and *clearly that he had the right to remain silent* is the adverbial. The average sentence length in sentence 1 is 23 words, and in sentence 2 is 7 words. The sentence is declarative in type and complex in its grammatical construction. In the use of **textual features**, the output contains two separate sentences. The first sentence consists of two clauses connected by commas, dashes, and colons, containing specialised legal vocabulary. The temporal phrase "*before my client uttered a word*" shows that the violation happened first, highlighting how early the mistake occurred in the process. The overuse of subordinators is a noticeable foregrounding stylistic feature for revealing the complexity and formality of the clauses. The use of passive construction with the modal auxiliary expressing necessity is strikingly noticeable: *should have informed*. The use of a dash

makes the text appear less cohesive, which is a noticeable foregrounding feature. There is an anaphoric reference to the pronouns *him/he*, which are cohesively tied. There is an anaphoric reference to the pronouns "him"/"he", which are cohesively tied. Word choices are formal, and the tone is respectful, furious, and demanding.

ChatGPT-4 generates the following **stylistic analysis**: *This excerpt mimics courtroom rhetoric with its formal address ("Your Honor"), first-person perspective, and deliberate appeal to authority ("fundamental and constitutional"). The use of the modal "should have" denotes legal obligation, while the parenthetical interruption "—clearly—" simulates spoken delivery and adds emotional emphasis. The anaphoric repetition of "that right" foregrounds the legal stakes. The prose is persuasive.* As shown above, ChatGPT4's analysis has not accurately examined the formal features such as lexical features: nouns, verbs, etc., the grammatical features: sentence (complexity, length, elements, types), tenses, voice, and the textual features such as word choices (formality), specialised vocabulary, tone, the foregrounding features in the writing mechanics (punctuations).

5. Discussion

The analysis indicates that ChatGPT-4's outputs in the two selected genres can be efficiently examined for the following stylistic features:

Lexical features: across the three prompts, the number of abstract nouns was higher in the legal genre (24) than in the journalistic genre (10), while the number of concrete nouns was higher in the journalistic genre (7) than in the legal genre (3). The number of verbs and adjectives was 10 and 5 in the legal genre, and 6 and 2 in the journalistic genre, respectively. The number of adverbs was 3 in the journalistic genre and 1 in the legal genre. The selection of verbs differs: 8 lexical verbs in journalistic prompts and 4 in legal prompts; 3 copular verbs were observed.

Grammatical features: the average sentence length is higher in legal prompts (74) than in journalistic prompts (44). Sentence complexity in the legal genre includes three simple, one complex, and one compound sentence, while in the journalistic genre it includes two simple, one complex, and one compound sentence. The sentence structures in the legal genre were as follows: S1=SVO, S2=, SVS3=SVC S= SVCAVOA, S1= SVA, S2=SVOVO, S3= SVO, and S4= SV. In the journalistic genre, they were S1=SVOA, S2= SVC, S1=SVC, S2=SVC, and S1=SAOVOA; S2=SVC. Sentence types were declarative

in both genres. The predominant sentence structure was simple in the journalistic genre and complex in the legal genre. The use of the active voice was higher than the passive voice in the journalistic genre, whereas in the legal genre, the use of active and passive voice was equal. Tense usage varied: the past tense predominated in the journalistic genre, while the present tense was more frequent in the legal genre.

Textual features: In the journalistic genre, no foregrounding punctuation marks, such as dashes or colons were observed, whereas in the legal genre, such punctuation was present. Specialised vocabulary was more predominant in the legal prompts than in the journalistic ones. Word choice was also more formal in the legal prompts compared to the journalistic prompts. No emotive lexis was identified in either genre. The tone in the journalistic genre was optimistic, reassuring, authoritative, neutral, detached, and reportorial, while in the legal genre, it was neutral, serious, informative, confident, persuasive, objective, respectful, and demanding. The tables below summarise the analysis of the two selected genres:

Table 1: Journalistic vs Legal Prompts for Lexical Features

Lexical Features	Journalistic Prompts			Legal Prompts		
	Minimal	Tone-Specified	Constraint-based	Minimal	Tone-specified	Constraint-based
Abstract nouns	4	5	1	6	9	9
Concrete nouns	1	1	5	1	1	1
Verbs	1	1	4	3	3	4
Adjectives	/	2	/	3	1	1
Adverbs	/	1	2	/	/	1

Table 2: Journalistic vs Legal Prompts for Grammatical Features

Grammatical Features	Journalistic Prompts			Legal Prompts		
	Minimal	Tone-Specified	Constraintbased	Minimal	Tone-specified	Constraint-based

Complexity	3 simple Sentences	1 complex sentence	3 simple sentences & 1 compound	2 sentences: 1 complex & 1 compound)	2 sentences: 2 clauses	2 sentences: 1 complex & 1 simple sentence
Length	S1=9 S2=4 S3=3	S= 14	S1=8; S2=6 S3=4 S4=4	S1=10 S2=8	S1=12 S2=14	S1=23 S2=7
Elements	S1=SVO S2=SV S3=SVC	S=SVCA VOA	S1= SVA S2=SVOVO S3= SVO S4= SV	S1=SVOA S2= SVC	S1=SVC S2=SVC	S1=SAOVOA S2=SVC
Type	Declarative					
Structure	Simple	Complex	Simple	simple	complex	complex
Tense	past, past, present	Past	Past	past	present, present	present, Past
Voice	active, passive, active	Active	active, active, active, passive	passive	active	passive, active

Table 3: Journalistic vs Legal Prompts for Textual Features

Textual Features		Journalistic Prompts			Legal Prompts		
		Minimal	Tone-Specified	Constraint-based	Minimal	Tone-specified	Constraint-based
Punctuation	/	/	/	/	dash, comma	Comma, dash, colon	
Specialised vocabulary	2	4	2	11	9	9	
Emotive Lexis	/	/	/	/	/	/	
Word choice	Less formal	Formal	Less formal	Formal	Formal	formal	

Tone	Optimistic, reassuring	Authoritative, neutral	Detached, reportorial	Neutral, serious, informative	Confident, persuasive, objective	Respectful, serious, demanding
------	------------------------	------------------------	-----------------------	-------------------------------	----------------------------------	--------------------------------

In addition, the analysis indicates a marked difference in analytic depth, linguistic accuracy, and descriptive precision between ChatGPT-4’s self-analyses and those performed by professional stylisticians. ChatGPT-4’s analyses focus on surface-level stylistic features, such as tone and formality, while omitting detailed lexical classification, grammatical precision, and functional explanation. For instance, while ChatGPT-4 can identify a text's general tone and some genre-specific features, its analysis remains shallow, focusing predominantly on themes and factual content rather than on deeper elements such as word choice, sentence structure, sentence complexity, and textual organisation. In practice, AI analysis tends to focus on thematic strictness, whereas professional human stylisticians examine and evaluate specific linguistic features. These include sentence structure (e.g., length, complexity, and type), formality, word choice, and tone, as well as the identification of cohesive devices. This distinction highlights a key divergence: AI often imitates authoritative analysis, while stylisticians validate their conclusions through apparent stylistic evidence. Professional stylisticians consistently provide precise linguistic categorisations, such as distinguishing between stative and dynamic verbs, between abstract and concrete nouns, and between past and present, emotive lexis, specialised vocabulary, formality, cohesive devices, and punctuation in the two genres that ChatGPT4 failed to recognise.

6. Findings

To present the findings, it is important to restate the research questions addressed in this study. The first research question was: *How can ChatGPT-4’s outputs in journalistic and legal genres be practically and efficiently analysed from a stylistic perspective?* The findings are summarised as follows:

A. lexical features

1. Legal texts contained a higher number of abstract nouns, while journalistic texts favored concrete nouns.
2. The legal genre showed greater use of verbs and adjectives, whereas the journalistic genre used more adverbs.

3. Verb selection differed across genres, with journalistic texts favoring lexical verbs and legal texts showing more restricted verb usage.

B. Grammatical Features

1. Sentence length and complexity were higher in the legal genre than in the journalistic genre.

2. Legal texts predominantly used complex sentence structures, while journalistic texts favored simple sentences.

3. Both genres relied mainly on declarative sentences.

4. The journalistic genre favored active voice, while the legal genre balanced active and passive constructions.

Tense usage differed, with the past tense common in journalistic texts and the present tense dominant in legal texts.

C. Textual Features

1. Foregrounding punctuation and specialised vocabulary were more prominent in legal texts.

2. Legal texts displayed a higher level of formality in word choice.

3. Emotive lexis was absent in both genres.

4. Journalistic texts showed a range of tones (e.g., reportorial, neutral, reassuring), while legal texts maintained a more formal and authoritative tone.

The second research question was: *To what extent does the stylistic analysis generated by ChatGPT-4, under optimised prompting, meet the standard of expert human analysis in terms of accuracy, interpretive depth, and comprehensiveness?*

The comparison between ChatGPT-4's self-analysis and professional stylisticians' analysis reveals clear differences in analytical quality, as shown below:

1. ChatGPT-4's stylistic analysis remains largely surface-level, focusing on tone, formality, and general genre characteristics. The AI often omits detailed lexical classification, grammatical precision, and functional explanation of stylistic choices.

2. Professional stylisticians provide systematic linguistic categorisation, including distinctions between: abstract vs. concrete nouns, stative vs. dynamic verbs, tense usage, specialised vs. emotive lexis, punctuation, and cohesive devices

While ChatGPT-4 can imitate an analytical tone, human stylisticians support their interpretations with explicit linguistic evidence. The gap between AI and human analysis

becomes most evident in complex texts, where ChatGPT-4 can match expected tone but fails to accurately assess linguistic validity.

The comparative analysis of ChatGPT-4's self-analysis and human stylists' assessments reveals that, while the AI shows a strong sense of genre fidelity, it still has a limited capacity for detailed stylistic interpretation. ChatGPT-4 can capture tone, general structure, and purpose, but it struggles to analyse fine-grained grammatical and textual features. In contrast, professional stylists provide richer interpretations, mapping sentence structures, clause functions, and lexical choices that truly define a genre's identity. Although the AI can reproduce the patterns of specific genres with high reliability, this is achieved by recognising patterns in large amounts of text rather than through genuine understanding. It imitates common features without grasping their rhetorical purpose, highlighting the gap between mimicry and mastery. These findings emphasise that human expertise remains crucial to ensure AI-generated texts meet not only stylistic conventions but also the broader communicative and rhetorical standards expected in professional writing

7. Conclusion

It is concluded that stylistic analysis remains an effective framework for examining AI-generated texts across genres. Although ChatGPT-4 can convincingly reproduce the surface features of journalistic and legal writing, its performance remains largely imitative rather than interpretive. While the model is capable of identifying general stylistic patterns, it lacks the analytical judgment, linguistic precision, and theoretical grounding required for comprehensive stylistic evaluation. In contrast, professional stylists demonstrate the expertise necessary to explain how and why specific stylistic choices operate within a genre. Consequently, AI can approximate the appearance of genre-specific writing but cannot fully account for its functional and rhetorical practicalities. Human expertise, therefore, continues to play a decisive role in ensuring analytical accuracy and rhetorical integrity in professional discourse. This study is **limited** to two genres, a specific dataset, and a single AI model, which may narrow the scope of the findings. **Future research** could explore additional genres, larger datasets, and different AI systems to increase our understanding of how AI produces and structures stylistic features.

References

- Abdulla, S. A. F. (2016). An Eclectic Model for the Stylistic Exploration of Mind Style in Fiction (Doctoral dissertation, University of Huddersfield).
- Ashwell, J. (2022). Stylometric comparisons of human versus AI-generated creative writing. [Publisher/URL from prior search, e.g., cora.ucc.ie repository].
- Bradford, R. (1997), *Stylistics*. London and New York: Routledge.
- Brown, T. B., Mann, B., Ryder, N., Subbiah, M., Kaplan, J., Dhariwal, P., Amodei, D. (2020). Language models are few-shot learners. *Advances in Neural Information Processing Systems*, 33, 1877–1901.
- Dale, R. (2021). GPT-3: What’s it good for? *Natural Language Engineering*, *27*(1), 113–118. <https://doi.org/10.1017/S1351324920000601>.
- Das, A., Sanyal, D., & Chakraborty, T. (2024). Stylo AI: Distinguishing AI-generated content with stylometric analysis. arXiv. <https://arxiv.org/abs/2405.10129>.
- Devlin, J., Chang, M.-W., Lee, K., & Toutanova, K. (2019). BERT: Pre-training of deep bidirectional transformers for language understanding. arXiv preprint arXiv:1810.04805. <https://arxiv.org/abs/1810.04805>.
- Halliday, M. A. K. & Hasan, R. (1989). *Language, Context, and Text*. Oxford.
- Hesselbach, R., Calvo Tello, J., Henny-Krahmer, U., Schöch, C., & Schlör, D. (Eds.). (2024). *Digital stylistics in Romance studies and beyond*. Heidelberg University Publishing
- Hoover, D. L., Culpeper, J., & O’Halloran, K. (2023). *Digital stylistics: Computational approaches to style in literature*. Routledge.
- Ippolito, D., Duckworth, D., Callison-Burch, C., & Eck, D. (2022). Automatic detection of generated text is easiest when humans are fooled. In *Proceedings of the 60th Annual Meeting of the Association for Computational Linguistics* (Vol. 1, pp. 427–443). Association for Computational Linguistics. <https://doi.org/10.18653/v1/2022.acl-long.32>
- Jeffries, L., & McIntyre, D. (2010). *Stylistics*. Cambridge University Press.

- Jurafsky, D., & Martin, J. H. (2023). *Speech and language processing* (3rd ed. draft). Stanford University. <https://web.stanford.edu>.
- Leech, G. N., Cruickshank, B., & Ivanič, R. (1989). *An A-Z of English Grammar and Usage*. Edward Arnold: London.
- Leech, G., & Short, M. (2007). *Style in fiction: A linguistic introduction to English fictional prose* (2nd ed.). Pearson Education.
- Quirk, R., Greenbaum, S., Leech, G., & Svartvik, J. (1985). *A Comprehensive Grammar of the English Language*. London: Longman.
- Reynolds, L., & McDonell, K. (2021). Prompt programming for large language models: Beyond the few-shot paradigm. *arXiv preprint arXiv:2102.07350* . <https://arxiv.org/abs/2102.07350>.
- Short, M. (1996). *Exploring the language of poems, plays and prose*. Routledge.
- Simpson, P. (2004). *Stylistics: A Resource Book for Students*. Routledge.
- Tikhonov, A., & others (2023). [Specific title on prompt engineering; inferred from context as exploring zero-shot methods—exact paper may vary, e.g., related to arXiv works on LLM prompting].