

Text-Linguistic Analysis in Forensic Authorship Attribution

*Eilika Fobbe**

Abstract

Authorship analysis of anonymous texts is one of the more frequently required tasks in a forensic setting. Its main purpose is either to gain information on the author's linguistic and social background to support an investigation or to match an anonymous text to a suspect's previous writings. One approach to authorship attribution is pragmatic stylistic analysis, which is grounded in text-linguistic research, cf. Brinker (2002), Sandig (2006), Püschel (2009), and Brinker et al. (2018), and holds a broad and holistic view on style. In its analyses, it focuses on the functional and pragmatic aspects of style as part of a communicative strategy. A central element, especially in Brinker's approach, is the thematic text pattern. How individuals argue, how they arrange textual patterns and how they express their demands most certainly reveal aspects of their individuality (Brinker, 2002; Brinker et al., 2018), suggesting that these cannot be easily suppressed or disguised. The paper applies Brinker's approach to three very short anonymous extortion letters illustrating how text-linguistic analysis can contribute to authorship attribution. Firstly, the text structure of the letters will be analyzed and secondly, the relations between text pattern and stylistic features will be examined in detail, and the relevance of the text patterns will be discussed with reference to the author's idiolect. The paper aims to point out the possibilities of a more in-depth textual analysis beyond the analysis of surface structures, especially in cases of very small data sets.

Keywords

text-linguistics, genre, stylistics, thematic text pattern, authorship analysis, extortion letter

**Eilika Fobbe*: Bundeskriminalamt Wiesbaden (Germany), eilika.fobbe@bka.bund.de. The author would like to thank the anonymous reviewer and the editors for their critical comments and supportive suggestions on a prior version of this article.

1. Introduction

In the context of language and law, the field of forensic linguistics deals with questions of language and the effects that may arise from a particular use of language. Since the circumstances under which a specific language use leads to legal consequences differ, so do the definitions of *forensic linguistics*. In a broader sense, the term encompasses legal language, courtroom interaction, language analysis of asylum seekers, police interrogations as well as the analysis of forensic texts (Coulthard & Johnson, 2010: 6), covering all kinds of communication. In a narrower sense, *forensic linguistics* is limited to the analysis of language when used in evidence and its linguistic evaluation posing a problem for legal judgment (Bierwisch, 1992: 56–57). Also, *forensic linguistics* is used synonymously to the analysis of forensic texts and the attribution of authorship (Ehrhardt, 2018: 167).

The term *authorship attribution* is often used to cover both *authorship identification* and *authorship verification*. Where the linguist compares a given text to a text of known authorship to determine if the same author has written both, this is referred to as *authorship verification*. The method's aim is not to identify an author but to validate the match, thus providing the basis for further conclusions about the likely author.¹ The term *authorship profiling* specifies the linguist's analysis of one or more anonymous texts to facilitate the operative investigation by specifying the social background of the author, primarily based on the evaluation of the writing skills displayed. In any of these cases, the methods usually applied are error analysis and the analysis of style.

Error analysis examines the deviations in a text to a prescriptive, often codified, norm. Errors can be described and assessed both in terms of their occurrences and their typicality. Stylistic analysis is based on defining language rather by variation than by homogeneity. The analysis of stylistic features thus works within the frame of style as a deviation from a norm and as variation within a norm (McMenamin, 2012: 488). Deviation, as well as variation, is produced either by choice or by habit whereby habits are considered generally originating from one's deliberate choices that have become habitual over time. Style is not restricted to classical stylistic figures, or the vocabulary of texts, although both represent relevant aspects of style. Many believe that style refers to the linguistic design of a text in its entirety. Style is therefore understood to be a characteristic of all texts where the writer produces style to evoke an effect in the reader, and the reader perceives style by checking it against the stylistic expectations he or she has towards the text (Spillner, 2009: 1762–1763).

Many studies in the field of authorship attribution have dealt with some controversial issues relating to appropriate methods and standards for admissibility in court.² One

¹ Therefore, the term *authorship identification*, which is also used for attribution problems by computational linguists (Stamatatos, 2009: 539), is misleading in the forensic setting.

² For U.S. American courts, cf. Shuy (2006), Chaski (2012). Because the legal system in Germany operates differently, the requirements for admissibility in German courts differ from those in the United States, cf. Kniffka (2007).

major topic of discussion is the application of computer-based or computer-driven techniques to the analysis of forensic texts. It has been pointed out on several occasions that many studies of this kind do not serve the needs of forensic work, also for methodological reasons (Chaski, 2012: 489–490; Grant, 2007: 3). A typical argument is that the sparseness of data available sets limits on the applicability of statistical analysis, but in times of social media, forensic text analysis also faces an increasing amount of data hardly manageable without computational support. Recent improvements of machine learning based applications have now offered new capabilities even to typical forensic constellations, such as cross-genre comparison, working with corpora of non-literary short texts, as well as unbalanced corpora. The argument of small data quantity has therefore become less essential but still remains.

Limitations of another kind arise from the fact that computational research in this field scarcely addresses theoretical issues of forensic authorship analysis. Authorship attribution inside and outside the field has always worked on the assumption that, in principle, one can discriminate texts by analyzing their language and that these language differences would point to different authors eventually. Also, an individual's language use can be determined and described by identifiable features. This central belief has led to many efforts to find the best features (cf. the survey of Stamatatos, 2009), which ideally “remain relatively stable over topic shifts and genre variation” (Stamatatos, 2018: 462) and have a differentiating value for future analysis. It was also hypothesized that with the use of as many features as possible, their discriminative value would always increase. Using machine learning systems has now made it possible to analyze several features simultaneously, rendering the frequencies of their combined occurrences. However, it also was proved that increasing the number of features does not necessarily improve the number of correct attributions – on the contrary, the application of a subset of selected features yields much better results (cf. Iqbal et al., 2008: 44).

It is worth noticing that these studies refer almost exclusively to surface structure. Surface structure features are easily identifiable by algorithms, and the more a feature occurs in a text, the better the results of the algorithmic analysis. Understandably, many of these studies use either a structuralistic concept of style or an additive one where the sum of its linguistic elements defines the style of a text. However, not only does it remain unclear how a linguistic element acquires its stylistic value in a linguistic sense,³ but it also reveals a lack of differentiation between language and style, which bears the problem that nothing can be stated about the style of a text when referring to the findings of a computer-driven or computer-based analysis. From a linguistic perspective, it is the rather unreflected view on a frequency-based definition of style accompanied by the neglect of findings in current linguistic research, which raises some concerns (cf. Fobbe, 2021).

³ From a statistical point of view a feature's stylistic value is defined by its statistically significant occurrence.

If there is not enough data available, statistical evaluation is not possible and an approach not dependent on statistical methods has to be taken. How sound the respective results are, fairly depends on how well the underlying stylistic concept or theory meets the aim of the stylistic analysis. Supposing, that their characteristics can differentiate texts and do point to their authors, the underlying definition of style should cover the author's production of style, as well as the perception of style by the reader. It should also explain how the language of the texts reflects both.

The problem with automated systems is not their method of feature extraction or evaluation but the underlying theory of style, which still cannot provide a sufficient explanation of how a feature's frequent occurrence in itself points to an individual author. Any analysis of style refers – explicitly or not – to a theory of language or at least to a theory of grammar; if there is no such theory, then the analysis is not a stylistic one (Toolan, 2014: 20), a problem only some of the computer-based studies are aware of.

The present article addresses the issue of style in the context of discourse analysis, especially text-linguistics, following an approach to textual analysis whose focus lies on the communicative function of a text. By integrating pragmatics into the analysis of style and expanding it to the textual level, the text-linguistic concept represents a method that offers a comprehensive approach to the variety of stylistic manifestations in texts.

The approach also addresses the limitations mentioned above set by the amount of data available — although it is usually applied to shorter-length texts, it may be applicable even to larger texts if the identification of patterns described below can be operationalized, and so turn out a useful tool for forensic linguistic purposes.

2. The Text-Linguistic Approach to Texts

As mentioned above, the analysis of the distribution of surface structure features is only one possible way of approaching authorship attribution. A text-linguistic approach would add to the analysis of surface structure elements both their pragmatic functions and their role within the composition of the text. According to text-linguistics, texts are described as “holistic entities” (Kniffka, 2007: 232) and analyzed not only by their vocabulary, grammar and rhetorical figures but also by topic, textual patterns and medium. Due to its close relation to pragmatic stylistics text-linguistics refers to a more holistic concept of style, which is by definition an integral part of any text (Sandig, 2006: 3–4). As a consequence, when speaking of stylistic analysis as part of a text-linguistic oriented analysis, it is considered to be the examination of all linguistic aspects of a text.⁴

⁴ The scope of *text-linguistics* is different from that of *discourse* and *discourse analysis*. While *discourse* may refer to both spoken and written texts, *text-linguistics* focuses on the written text and “characterizes ‘text’ as a medium dependent notion” (Esser, 2009: 7).

The text-linguistic approach, represented by the work of Klaus Brinker and applied here, emphasizes the importance of pragmatics as the most comprehensive aspect of text-linguistics (Brinker et al., 2018: 17). His method reflects the turn of text-linguistics towards an integrative approach, which has abolished the complete assimilation of the text constitution in speech act categories and transferred it into a model that distinguishes between levels of description (Feilke, 2000: 67). While using Brinker's approach to speech act theory provides the theoretical framework for defining a text as text, he regards speech act types as to refer to the different levels of description, (i.e. the illocutionary act to the communicative-pragmatic level, the propositional act to the level of the topic or theme and the locutionary act to grammar). At the same time, the analysis carefully differentiates between the levels without treating them as isolated units. To the contrary, it is the coherence of these levels' constituents, i.e. between sentences, thematic patterns and pragmatic acts, which is the focus of textual analysis. According to Brinker, the linguistic means only have a serving role for the other levels and must be seen and interpreted in terms of their indicative function for the communicative goal and content of the text (Brinker, 2000a: 164). Partly drawing on van Dijk's concept of macrostructures, Brinker places the topic or theme of the text at the centre of his approach, arguing that it is by the topic that presuppositions, substitutions, and implicit recurrences 'work' because the topic lays the foundation for a specific perspective that makes coherent presuppositional connections and expectations possible in the first place (Feilke, 2000: 69).

As a consequence to the important role that pragmatics plays for text-linguistics, a text is primarily considered to be serving the speaker's communicative intentions, and its realization is interpreted as a complex linguistic action (Brinker, 2002: 42). It is also believed that the speaker composes a text sample exactly the way he or she believes to be the most promising to achieve his communicational goal. This assumption refers not only to the language employed but also to the choice of thematic text patterns in structure and sequencing (Püschel, 2009: 52). An integral part of any stylistic analysis of a text is therefore to identify tacit assumptions, presumptions and implications that accompany the reading of the text. By making them explicit, the author's communicative decisions are described, and it becomes clear where they differ from the reader's expectations.⁵

Any text consists of several text patterns realized in different fashions, where the text function merges with specific textual characteristics. Text-types and their text patterns are conventionalized to different degrees, and it is "not that the writer does not have the choice of making his or her own meanings, but the constitutive conventions fundamentally restrict the set of elements available for combination in specific texts" (Carter & Nash, 1990: 10). Thus the choices and restrictions that are subject to the constitutive conventions are socially primed and based on social decisions. According to Fix (2004: 42)

⁵ This description of a pragmatic (text)-stylistic analysis shows parallels to what is described as a relevance-theoretic approach (cf. Clark, 2018: 159–160).

writers differ in using linguistic means because the type of social relationship they want to establish between them and their readers varies.

2.1. Individual Style and Text-Type

In the text-linguistic context, the concept of individual style (or idiolect) comprehensively relates to the whole design of a text and its text patterns. Any text sample with its patterns may fall into a text-type but is created by an individual for an individual purpose. An author adapting a text sample to an individual situation may develop stylistic traits that do not have to be part of the original text-type, and its patterns could be modified to different degrees. Then again, text-types are formally defined by obligatory components both requiring a specific register and showing restrictions with reference to the social role of the persons involved, the topic and the communicative situation. A person who would want to achieve his/her communicative goal would possibly try to meet the text-type based requirements linguistically. The more prescriptive the norm of a text-type, the more consistent its adaptation by the individual on his/her text sample. As a result, the linguistic choices of the individual do not bear ipso facto an individualizing effect on style. So even if one subscribes to the existence of individual style, it does not necessarily mean that it has to be always detectable or would always develop an individualizing function.

2.2. The Text-Type *Extortion Letter*

While early research in text-linguistics was largely interested in the more general aspects of text-types⁶ and their social and linguistic constraints (Adamzik, 2018: 55), research in malicious (forensic) texts focused on both the discrimination between conforming text-type based stylistic features and idiolectal features and the thorough description of thematic text patterns.

The reasoning is clear: The better we know the different components of a text type together with their potential linguistic variations, the easier it is to identify them and to evaluate their potential as idiolectal variations.⁷ These insights could then aid the investigative process as only idiolectal features are regarded as potential cues to the author (Dern, 2009: 36–37).

⁶ The term refers to the written form of a text and to its topic. A *text type* is primarily defined by its communicative function. It can be applied to describe all forms of written communication (including the growing part of digital communication) sharing the extortionate function (Brinker et al., 2018: 142).

⁷ A similar approach is presented by Wright in his study of email greetings and farewells. Wright points out that “a genre-focused approach has strong potential for authorship research” (2013: 72).

Research has also suggested that the text-type *extortion letter* is only minimally shaped by norms and therefore, more open to individual design (Fobbe, 2011: 75). So, for the most part, the author of such a text acts independently, drawing on mere assumptions of how “extortionists” are supposed to communicate and on more or less suitable templates like the formal (business) letter. A research study by Dern (2003) demonstrated for instance, that the politeness occurring in many letters often originates from the more formal registers of the adapted templates. In another study, Stein & Baldauf (2000) found several phrasal expressions (such as “no tricks”, “no police”, “this is not a joke”) also to be text-type based in character as they frequently occur in text samples of the *extortion letter*. For the evaluation of features identified in any extortion letter, this means that even unusual expressions may be quite predictive with respect to the text-type (and therefore less useful for authorship attribution), but others could display a more individual function. These findings on elements based on the text-type not only contribute to the determination of idiolectal features but also can be of use for the more general aspect of intra- vs inter-author variation. An early study on extortion and threatening letters by Artmann (1996) laid the ground for a pragmatic approach to this specific text-type. He used speech act theory describing the letters’ function, structure and characteristic linguistic elements. Later his work was refined by Brinker (2000c, 2002) who identified so-called thematic patterns and explained them by relating to their underlying text functions, which are observed in specific themes or topics. As in other text-types, thematic patterns fall into two main categories: One category refers to those patterns that are obligatory with respect to the main illocutionary function, in the case of extortion, i.e. the DEMAND and the THREAT, because only by them the illocutionary act of extortion is linguistically completed. According to Brinker’s terminology, these patterns are called ‘obligatory basic functions’ (*obligatorische Grundfunktionen*). The other category refers to patterns that are optional because they do not contribute to the speech act of extortion itself. Brinker labels these ‘optional additional functions’ (*fakultative Zusatzfunktionen*).

2.2.1. Text-Thematic Patterns of the Extortion Letter⁸

The first pattern that is obligatory to any extortion is the *announcement of action*, i.e. the declaration of an action that threatens the addressee and usually describes a violent or damaging future act by the author that is evaluated as highly undesirable by the addressee (König & Pfister, 2017: 171),

- (1) if you want to stop the killing⁹
- (2) in order to prevent me from using this information against you.

⁸ Cf. Brinker, 2000c: 40–42 and Brinker, 2002: 54–55.

⁹ Examples 1, 3, 5 are retrieved from 2.bp.blogspot.com, 9 from imperva.com, 4, 6, 7, 10, 12, 14 from gramcluley.com, 8, 11, 13 from the JonBenét Ramsey ransom note.

The second obligatory pattern is the so-called *demand for action*, which usually involves paying a large sum of money:

- (3) then wire \$1,000,000.000
- (4) all you have to do is to pay me \$2000 [...].

A third function, closely related to the *demand for action*, is realized by the thematic pattern that specifies the handover, which is commonly referred to as *handover procedures*. This pattern is also called an ‘obligatory additional function’ (*obligatorische Zusatzfunktion*) because it is essential for the completion of the extortion:

- (5) to bank account 84-49-897 at Continental Illinois Bank Chicago
- (6) Payment must be received within 10 days of the post marked date of this letter’s envelope.

However, it is well-known by forensic linguists and investigators that many writers of extortion texts do not restrict themselves to the ultimate act of extortion but often dwell on other subjects like justifying themselves or holding the victim responsible for the crime. Because these text patterns are not necessary to complete the extortion but still occur frequently, they represent ‘optional additional functions’ (*fakultative Zusatzfunktionen*) and are considered relevant for insights into the author’s intentions and motivations (Dern, 2009: 163, 173; with instructive examples). The ‘optional additional functions’ can derive both from the *announcement of action* as well as from the *demand for action*. Brinker (2000: 40; 2002: 54, with reference to Artmann) identifies four supplementary functions and labels them as follows:

a) Holding the victim responsible for the course of action, i.e. *attribution of responsibility*:

- (7) That choice is completely yours.
- (8) It’s up to you now.

b) Affirming the gravity of the situation and/or the perpetrator’s determination, i.e. *assurance of determination*:

- (9) This is not a hoax
- (10) consider this: you received this letter via first class mail.

c) Instructing the victim not to take any steps against the perpetrator and to comply with the instructions, i.e. *request of compliance*

- (11) so I advise you not to provoke them.
- (12) if you don’t comply with my demand I am [...] going to humiliate you.

d) Presentation of the author to put more pressure on the victim, i.e. *self-presentation*

- (13) we are a group of individuals that represent a foreign faction
- (14) hello, you don’t know me but I know you very well.

2.2.2. The Text-Linguistic Approach to an Extortion Letter

From a text-linguistic perspective, the seven thematic patterns discussed above are core components because they constitute the text-type *extortion letter*. The analysis of such a letter by linguistic and stylistic features only makes sense when related to the thematic and functional structure of the text. Brinker has repeatedly (2000c, 2002) stressed the importance of this relationship for purposes of authorship analysis. According to him the way an author approaches the communicative task of extortion is influenced more likely by individual aspects – if not even aspects of the author’s personality – rather than by text-type dependent requirements.

The communicative task is accomplished through decision-making, i.e. the author decides which aspect is mentioned first, which one later, what possibly may be left out, how the instructions are arranged and how they are expressed. Brinker (2000c: 39, 41) emphasizes the relevance of indirect speech acts vs direct speech acts and the coherent and incoherent realization of logically dependent thematic patterns. Eventually, authors must use language to achieve their goals, and language is the instrument with which the perpetrator can commit the crime of extortion.

The text-linguistic approach to any text would at first identify the main illocutionary function of a text, how the text-theme unfolds through thematic patterns and how their functions contribute to the main function. Its stylistic analysis would examine the arrangement of these patterns within the text with respect to coherence and logic structure, analyze the language used and finally describe on all linguistic levels how function and themes connect and intertwine. By redefining the role of the surface structure elements according to a text-linguistic based concept of style, the applicable feature set expands. This is especially beneficial with regard to very short texts. Under these premises, a relatively small data set that does not lend itself to statistical analysis could still permit us to draw significant conclusions.

Of course, there are limitations to this type of approach. Brinker himself not only identified the thematic patterns presented above but also tried to derive some general text-type variations and their order from them, which led to a wide range of structures (Brinker, 2002: 53–57). So far, due to research desiderate, not much can be said about the typicality of a specific text structure, i.e., how frequent or rare an actual sequencing structure is, compared to other texts in a benchmark corpus.

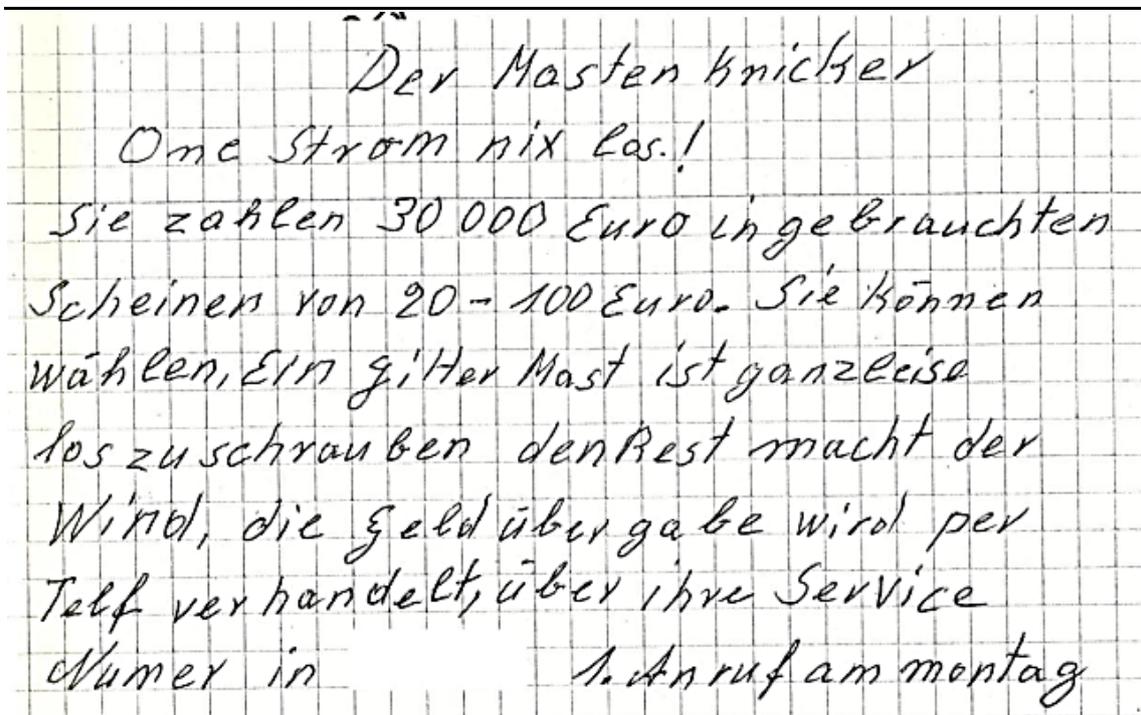
The approach to a writer’s style by integrating text-linguistics into its analysis is therefore particularly useful when comparing short texts in a series; it may also support identifying potential imitators. In the first case, it can support stylistic insights at the lexical and syntactical level; in the second, it can reveal differences at the textual level, even if the text refers to a similar or identical crime scenario.

3. A Case Study

For purposes of analysis, three very short texts from a series of seven extortion letters have been chosen. The anonymous letters were taken from the database of forensic texts of the German *Federal Criminal Police Office* (Bundeskriminalamt).¹⁰ The database, called *Kriminaltechnisches Informations-System Texte* (KIST), contains about 6.000 incriminatory and malicious forensic texts.

In these texts an anonymous author threatens a German Company to sabotage parts of the local infrastructure if the company does not pay him € 30,000. The analysis focuses both on the text patterns and on their stylistic features; the results of this analysis may be used for a potential authorship verification task at a later time.

Figure 1: Letter 'pylon' (anonymized)



¹⁰ Although most of the current material is computer-written and comes from digital sources, some offenders still write by hand.

Figure 2: Letter 'train' (anonymized)

Es wird Ernst!
 Sie haben die Wahl; entweder Sie
 zahlen 30000 Euro oder Sie verlieren
 einen Zug auf der Schiene von
 Die Übergabe
 wird durch Telf bekannt gegeben.
 Das Geld wird in gebrauchten Scheinen
 die nicht gekentzeichnet sind an
 einen Ort deponiert
 Der Schienen heiler

Figure 3: Letter 'power line' (anonymized)

Sie zahlen 30.000 Euro. Damit
 der Masten nicht umfallen es sind
 nur ein Teil Schrauben zu lösen, und
 die fällt aus.
 Die Geld übergabe erfolgt per totem
 Briefkasten Ort und Zeit wird am
 Freitag per Telf. bekannt gegeben.
 Der

3.1. Thematic Patterns, Speech Acts and Stylistic Devices

For providing the reader with a first impression of the letters' thematic structure, a sentence-by-sentence translation of all three texts is given in the tables 1–3 below, alongside a list of the thematic patterns:

Table 1: Text and thematic patterns of the letter 'pylon'

Translation	Thematic text pattern
The Pylon Breaker	self-presentation
no power no fun!	assurance of determination
You pay € 30,000	demand for action
in used notes from € 20–100	handover procedures
You can choose.	attribution of responsibility
A pylon is very quietly to unbolt ¹¹	announcement of action
the wind does the rest,	announcement of action
the handover will be negotiated via ph[one],	handover procedures
via your service hotline in [name of town]	handover procedures
1 st call on Monday	handover procedure

The text 'pylon' starts with a *self-presentation* ("the pylon breaker") followed by a slightly modified slogan assuring the author's determination. Then the *demand for action* is made, including a relatively detailed description of the money delivery. The following sentence expresses *attribution of responsibility* to the victim by explaining that he or she has a choice. This pattern is followed by two sentences referring to the announced action. The closing paragraph describes the way the handover will be negotiated and how the contact for further instructions will be made.

The second letter 'train' also starts with a standard phrase to affirm the extortionist's determination. Again, the *demand for action* is introduced by the *attribution of responsibility* and followed by *the announcement of action*. The next paragraph contains the *handover procedures*, including details about the notes, the sum and the way further contact will be made. The text closes with a signature comprising a *self-presentation*:

¹¹ The phrase can mean 'can be unbolted'. In contrast to English the *be-to construction* in German is neither very formal nor does it imply a formal arrangement. It can both express 'necessity' (*has to*) and 'possibility' (*can*): "Like *can*, which could be used instead, *be to* just expresses a present modal state, without even implying actualization of the residue-situation. [...]" (Declerck, 2010: 85; with reference to contemporary English). This definition is valid also for the German use of *ist zu* here.

Table 2: Text and thematic patterns of the letter ‘train’

Translation	Thematic text pattern
It's getting serious!	assurance of determination
you have the choice,	attribution of responsibility
either you pay € 30,000	demand for action
or you'll lose a train on the track from [A to B]	announcement of action
The handover will be announced by phone	handover procedures
The money will be in used notes	handover procedures
which are unmarked	handover procedures
deposited at a place	handover procedures
The Rail Wedger	self-presentation

In the third letter, the first sentence is the *demand for action* in the form of a statement. It is followed by *the announcement of action* referring to the event, its execution, and outcome. The next thematic pattern in line is the *handover procedures*. The vague place is mentioned before contact, day and medium of communication. The letter also closes with an alias used in the signature representing the pattern of *self-presentation*.

Table 3: Text and thematic patterns of the letter ‘power line’

Translation	Thematic text pattern
You pay € 30,000	demand for action
so that the pylons don't fall down	announcement of action
just some of the bolts are to unbolt ¹²	announcement of action
and the North-Southern power line will fail	announcement of action
The handover will be take place via a dead letter box	handover procedures
place and time will be announced by ph[one] on Friday	handover procedures
The Pylon Breaker	self-presentation

3.1.1. *Demand for Action and Announcement of Action*

Starting with the central pattern *demand for action*, a comparison of the letters shows that the author's demand is expressed by a declarative sentence rather than by an imperative one in all three letters. By changing the grammatical mood and by stating the expected behavior instead of demanding it, the commanding speech act acquires a more indirect character.

To express the commissive act of the *announcement of action* the author also decides for an indirect approach by realizing the commissive speech act indirectly: He uses passive

¹² Here the *be-to* construction expresses ‘necessity’ in German (‘have to be unbolted’).

voice combined with a modal *be-to construction* to describe the execution of the action in the first and the third letter; in this way the author avoids referring to the perpetrator of the crime. In the second letter, the author turns the focus away from himself in two steps: First, he takes the reader's perspective by using "you" instead of "I" and then he chooses the verb "to lose" instead of some action verb like "to do" or "to destroy". Whether the sabotaged train is a loss to the reader, the writer cannot say, but indicating it to the reader he is deliberately taking his perspective. Moreover, the author's change in perspective is just another variation to avoid any self-reference.

With the indirectly realized speech acts DEMAND and THREAT, even the extortion itself has to be inferred from the sequence of sentences in the letter 'pylon'. It lacks the conjunction "if" that would make the conditional relation between the *announcement of action* and the *demand for action* explicit. In the second letter ('train'), the link between DEMAND and THREAT is realized by the conjunction "either-or", which is compatible with but not identical with the conditional relation. Accordingly, in the third letter ('power line'), the author implies a link between the *demand for action* (*pay, so that the pylons don't fall down*) and the falling of the pylons by using a final subordinate clause, a relation which does not exist in reality. Shifting the responsibility for the power failure to the reader adds to the *announcement of action* the function of *attribution of responsibility*. This coordinates well with the fact that the execution itself is expressed by using a *be-to construction* (*just some of the bolts are to unbolt*) without any acting subject. Again, it is up to the reader to fill in the missing elements.

It is also worth mentioning that with respect to the narrative structure of all three letters, the thematic text pattern *announcement of action* lacks completeness. All texts select from narrative elements which would describe the action in full, i.e. execution, event, and outcome. In the text 'pylon' the execution is mentioned, but the event or the outcome is not. Similarly, in the text 'train' the reader does not know what will cause the train's derailment. In both texts, the reader must infer by way of implicature the damaging effect of sabotage. The third letter leaves it to the reader to fill in the gap between the failing of the power line, the unbolting of the pylons and their falling by his/her imagination. However, this lack of information is compensated by the text pattern of *self-presentation*, providing additional details in all three letters.

But there is a limiting aspect to the THREAT at least for the pylon breaker in the first letter: Usually, it would be strategically advantageous not to reveal the exact course of action, but in all letters the author deliberately provides *some* information. With respect to the illocutionary act of EXTORTION, a description of the *modus operandi* is not necessary but it is with respect to the perpetrator's potential power. Nevertheless, the mentioning of "the wind" doing "the rest" in the first letter obfuscates the identity of the agent. In this way, the author transfers to the elements powers that he earlier claimed he has. As a consequence, this feature of the *announcement of action* is actually a variation of the optional function of *attribution of responsibility*.

3.1.2. *Self-presentation and Assurance of Determination*

As the reader may have observed, the letters contain the multi-functional thematic pattern of *self-presentation* in the form of the nonce words *rail wedger* and *pylon breaker*, which are used as signatures to describe the author's alleged abilities concisely. In all three texts, this form of self-presentation is structurally derived from the obligatory *announcement of action* because it refers to the abilities the author needs to execute the sabotage. By using the alias "pylon breaker" in two of the letters ('pylon', 'power line'), the author informs us of his quite supernatural powers, in the other letter ('train'), only the alias provides information about a possible modus operandi (i.e., using a wedge for causing the train to derail). Therefore, the alias's main function is the *assurance of determination* by intimidating the victim. Once this type of alias is introduced, the author sticks to it, even as he changes from "pylon breaker" to "rail wedger".

The aforementioned *self-presentation* is expressed in the form of a deverbative noun. This type of word formation reduces the acting subject to the *-er* ending and allows the author to present himself in a more indirect way. We already observed this indirect presentation earlier within the thematic patterns *announcement of action* and *demand for action*.

At the same time, this lexical choice points to linguistic creativity as well as to a rather spontaneous approach to the underlying text theme. The same applies to the slogans in the first two letters in which the author modifies the original wording¹³ in order to describe his own planned sabotage. Bearing in mind that the main function of this optional pattern is *assurance of determination*, these lexical choices seem to serve the purpose well. They describe an omnipotent offender making fun of the serious situation, which would be interpreted by the reader as a display of the author's arrogant and insouciant attitude to the crime.

3.1.3. *Handover Procedures*

All letters seem to be initial letters because the *handover procedures* deal with the issue of making contact for the first time. Within this thematic pattern, the letters clearly show variation by referring in different ways either to the money (text 'pylon'), to both the money and the point of delivery (text 'train') or to the point of delivery only (text 'power line'). In terms of specificity the text 'train' remains vague on the point of delivery, the text 'power line' is quite specific by naming a drop site, and the text 'pylon' does not provide any such information. One wonders if these differences might correlate with the letters' time of origination. They also vary in sequencing: In the first text, the *demand for action* is directly followed by the money details. In the text 'train' the references to the

¹³ The original meaning is "without money ('dough') nothing works" – "ohne Moos nichts los". The author changes it to "ohne Strom (power) nichts los".

money details are placed between the announcement of making contact and the information about the place, and in the text ‘power line’ the details of the delivery are placed before those of further contact. The letters’ only constants are how the contact will be made (by phone) and how much money has to be paid.

Linguistically, the author shows a preference for the use of the passive voice and a more formal register in all three texts, when describing the *handover procedures*. One may find this type of vocabulary and syntactical structures in the register of formal letters, e.g. business letters or official letters. As we know from research on the text-type extortion letter authors often draw on this type of template to give a semi-official tone to the issue (Hoffmann, 2009: 310), so the author’s linguistic choices might have been under the influence of these specific genres. At the same time, the use of the passive voice is coherent with the overall tendency towards indirectness.

The author also varies in his interpretation of the envisaged contact by using different verbal expressions: In the first text, he thinks of it in terms of a preliminary negotiation through a first call and therefore, provides a detailed explanation about how he is going to make the call. In the other two texts, he just tells the victim that he is going to inform him/her about the handover details. This form of thematic shifting within the *handover procedures* might reflect that the perpetrator has not yet decided what attitude he wants to adopt towards his victim.

Looking at the thematic text patterns as a whole one can infer that indirectness is a conspicuous stylistic pattern here, emerging from the combined use of the corresponding linguistic features on the lexical, syntactical and textual level. However, there are differences in vagueness or indirectness for the different text patterns. While the *demand for action* is relatively explicit, the *announcement* in all three letters remains rather vague. The self-presentation is reduced not only in terms of word formation but also in terms of self-reference by not using the personal pronoun *I*.

The handover procedures also show a continuous transition from directness to indirectness. In contrast to the detailed description of the characteristics of the money, the place of handover lacks any detail. The same applies to how the author intends to contact the victim. Here the way of communication is clear, but the author does not give a date, only a day of the week, a particularity that could turn out to be a reference problem, as the author cannot control the exact time at which the victim will receive the letter, and the victim could misinterpret the date. Also, he does not mention the exact time of his phone call.

The description of the handover location is unusual, too. In the text ‘*pylon*’ the author does not refer to it at all, which is usually the case with an initial letter, but he also refers to it redundantly (“at a place”, in the text ‘*train*’) or illustrates the location’s characteristics (“via a dead letter box”, in the text ‘*power line*’) in the other letters. These different linguistic representations provoke in the reader the impression of indeterminacy. With regard to the text’s main pragmatic function (*appeal*), this is in stark contrast because the recipient – even if he or she is willing to do so – is not able to cooperate due to the lack of

information. Moreover, if we take into account, that these texts reflect the communicative strategy the author considers the most successful, then this indeterminacy has its right and serves a purpose. From the author's perspective, this could mean keeping all options open, even if risking that the money will not be delivered. From the recipient's perspective, this could lead to the conclusion that there is only a vague idea of how the handover should take place or that the author has not decided yet on how exactly to proceed. This is contradictory, but not at all surprising because to successfully complete the phase of negotiation and handover requires more detailed planning (Dern, 2009: 147); a task that some people are unable to accomplish.¹⁴

However, the subtle differences within the *handover procedures* and the relatively large amount of text dedicated to the *handover procedures* suggest a fairly stable focus of the perpetrator on the practical aspects of the extortion, i.e. gaining possession of the money.

3.2. Sequence

Based on the “pragmatic premise of intentionality of communicative action” (Fetzer, 2018: 416) the analysis of the text pattern arrangement and their comparison could provide insights into how the author tries to complete the challenging communicative task of extortion (Brinker, 2000c: 35) – insights which may not only be of linguistic but also of investigational interest. Table 4 on the next page illustrates the linguistic variation between the letters under comparison.

A closer look at this table reveals some correlation between the letters' sequencing: The pivotal speech act EXTORTION starts with the *demand for action* and then adds the *announcement of action* in all letters. The details of the money and the specifications of the contact, which are logically dependent, are realized in a coherent way but differ in order. The same applies to the remaining optional patterns. The author uses them like text modules in different arrangements and begins with a standardized phrase in two out of three texts. While in the text 'pylon' he introduces the *attribution of responsibility* before demanding the money, in the text 'train' he prefers the reverse pattern. However, in both cases, the author starts with a combination of two thematic patterns that represent optional functions (*Self-presentation, assurance of determination* compared to *assurance of determination, attribution of the responsibility*). The text 'power line' provides a scantier version of the extortion by leaving aside all optional patterns apart from *self-presentation*.

¹⁴ As a result, most perpetrators end their extortion prematurely; in the case of product extortion only 15% enter the handover phase (Dern, 2009: 149).

Table 4: Comparison of all letters

Text 1 'pylon'	Text 2 'train'	Text 3 'power line'
The Pylon Breaker	It's getting serious!	
no power no fun	you have got the choice	
You pay € 30,000	either you pay € 30,000	You pay € 30,000
in used notes from € 20 – 100	or you'll lose a train on the track from A to B.	so that the pylons don't fall down
you can choose	You will be informed about the delivery by phone	just some of the bolts are to unbolt
a pylon is very quietly to unbolt, the rest does the wind	The money will be in used notes, which are unmarked, placed at a place	and the North-Southern power line will fail
The handover is negotiated via phone, over your service hotline in A-town, 1. call on Monday	The Rail Wedger	The handover will be made via mail drop
		You will be informed about time and place on Friday via phone
		The Pylon Breaker

It is interesting to note that the text 'power line' does not introduce other patterns instead, so the displayed communicative pattern may well be an idiosyncratic feature of the author's language use:

Table 5: Comparison of the letters' sequencing

Text 'pylon'	Thematic text pattern	Text 'train'	Thematic text pattern	Text 'power line'	Thematic text pattern
The Pylon Breaker	self-presentation				
no power no fun	assurance of determination	It's getting serious!	assurance of determination		
you have got the choice	attribution of responsibility	either you pay € 30,000	demand for action	You pay € 30,000	demand for action
You pay € 30,000	demand for action	you can choose	attribution of responsibility		
		The Rail Wedger	self-presentation		
				The Pylon Breaker	self-presentation

These findings make it plausible that the variation observed in the sequencing is part of the author's intra-author variation. However, the author seems to be very consistent in the selection of patterns. All texts do without the optional pattern *request of compliance* and do not vary on the functions of the patterns realized. Accordingly, in all three letters, the text pattern of *self-presentation* is only used to illustrate the act of sabotage but not to justify the crime.

4. Conclusion

This article has aimed to present the text-linguistic approach as a useful method for the forensic-linguistic analysis of forensic texts, particularly of short ones. The combined analysis of the thematic text patterns, their pragmatic realizations, their sequence and their linguistic features not only connects the surface structure elements to the deeper layers of the text but also offers a concept of style which integrates the textual level into the linguistic analysis. As the whole linguistic design of a text is considered to serve the author's communicational intentions, its linguistic characteristics are possible cues to the author's thoughts and motivations. Text-linguistics may also work well as a linguistic tool for authorship identification purposes, helping the expert linguist to solve problems relating to consistency (intra-author variation), resemblance and population (inter-author variation).

The analysis of the texts followed Brinker's (2000c, 2002) research on extortion letters based on the concept of the thematic pattern. The combined examination of the optional text patterns, their arrangement and their language has provided the following insights: The author's dominant preference for indirectness has been identified by analyzing the illocutionary speech acts of both the optional and the obligatory text functions. In all three letters, the author decided for a more indirect realization of the respective speech acts, also making repeated use of implicatures as well as of the corresponding stylistic devices on the different levels of language such as impersonal constructions or passive voice. It has also been pointed out that some of the text patterns lack crucial content information (*outcome of the sabotage, point of delivery*). In this way, the author added an impression of vagueness to the situation. Furthermore, the specific linguistic design of the optional patterns (*slogan, ad-hoc formation*) revealed the author's personal detachment. Finally, in two of the letters, the *announcement of action* developed also the function of attributing the responsibility to the victim or even the elements. It can be inferred from these findings that what is missing in all letters is self-commitment, in other words, the main conceptual strategy of this author is to pass the responsibility for the crime to someone else.

Concerning the relationship between the single text pattern and the choice of specific linguistic means, the latter should not be considered as arbitrary but rather bound to specific thematic patterns. The registers of the thematic patterns differ in being either more formal (*handover procedures*) or more colloquial with variant expressions (*assurance of determination, self-presentation*), but the combination of a pattern with a specific register remains constant across all texts.

Furthermore, the analysis of the sequencing has indicated that the patterns seem to follow a certain sequence arrangement: While some patterns vary in sequence, others share the same relative order in all three texts. Another interesting result is the fact that the texts do not vary with regard to the selection of the patterns. Although the vocabulary of the pattern differs slightly and the sabotage relates to different scenarios, their overall

forms, as well as their functions, remain the same across the letters. These findings, however, not only suggest that the author is being consistent in his communicative design of the extortion but also that the text pattern in itself is an important condition for intra-author variation and therefore, a worthwhile topic for further exploration.

References

- Adamzik, Kirsten (2018). Textlinguistik. In Liedtke & Tuchen (Eds.), *Handbuch Pragmatik* (pp. 53–64). Stuttgart: J. B. Metzler.
- Anonymous (n. d.). Death of JonBenét Ramsey. Evidence. Ransom note. Retrieved August 26, 2020, from en.wikipedia.org/wiki/Death_of_JonBenét_Ramsey.
- Anonymous (2012). Accusations. Retrieved October 25, 2020, from introductiontoproject.blogspot.com/2012/11/accusations.html.
- Anonymous (2016). Ransom note. Retrieved January 5, 2021, from imperva.com/blog/how-to-respond-to-ddos-ransom.
- Artmann, Peter (1996). *Tätertexte. Eine linguistische Analyse der Textsorten Erpresserbrief und Drohbrief*. Inauguraldissertation Universität Würzburg. München.
- Bierwisch, Manfred (1992). Recht linguistisch gesehen. In Grewendorf (Ed.), *Rechtskultur als Sprachkultur. Zur forensischen Funktion der Sprachanalyse* (pp. 42–68). Frankfurt/M.: Suhrkamp (stw; 1030).
- Brinker, Klaus (2000a). Textstrukturanalyse. In Brinker, Antos, Heinemann & Sager (Eds.), *Handbuch Text- und Gesprächslinguistik. Ein internationales Handbuch zeitgenössischer Forschung* (pp. 164–175). Berlin/New York: de Gruyter.
- Brinker, Klaus (2000b). Textfunktionale Analyse. In Brinker, Antos, Heinemann & Sager (Eds.), *Handbuch Text- und Gesprächslinguistik. Ein internationales Handbuch zeitgenössischer Forschung* (pp. 175–186). Berlin/New York: de Gruyter.
- Brinker, Klaus (2000c). Zum Problem der Autorenerkennung aus textlinguistischer Sicht (am Beispiel von Erpresserschreiben). In Baldauf (Ed.), *2. Symposium Autorenerkennung des Bundeskriminalamtes vom 03. bis 05. April 2000* (pp. 34–53). Wiesbaden: Bundeskriminalamt.
- Brinker, Klaus (2002). Textsortenbeschreibung auf handlungstheoretischer Grundlage (am Beispiel des Erpresserbriefs). In Adamzik (Ed.), *Texte – Diskurse – Interaktionsrollen. Analysen zur Kommunikation im öffentlichen Raum* (pp. 41–59). Tübingen: Stauffenburg.
- Brinker, Klaus, Cölfen, Hermann & Pappert, Steffen (2018). *Linguistische Textanalyse. Eine Einführung in Grundbegriffe und Methoden*. 9., durchgesehene Auflage. Berlin: Erich Schmidt.
- Carter, Ronald & Nash, Walter (1990). *Seeing Through Language: A Guide to Styles of English Writing*. Oxford: Blackwell.
- Chaski, Carole E. (2012). Author Identification In The Forensic Setting. In Tiersma & Solan (Eds.), *The Oxford Handbook of Language and Law* (pp. 489–503). Oxford: OUP.
- Clark, Billy (2018). Stylistics and relevance theory. In Burke (Ed.), *The Routledge Handbook of Stylistics* (pp. 155–174). London/New York: Routledge.
- Cluley, Graham (2016). Here's what an Ashley Madison blackmail letter looks like. Retrieved October 01, 2020, from grahamcluley.com/ashley-madison-blackmail-letter.
- Coulthard, Malcom & Johnson, Alison (Eds. 2010). *An Introduction to Forensic Linguistics: Language in Evidence*. Reprint. London/New York: Routledge.

- Declerck, Renaat (2010). Future time reference expressed by *be to* in Present-day English. *English Language and Linguistics*, 14(2), 271–291. DOI: [10.1017/S1360674310000080](https://doi.org/10.1017/S1360674310000080).
- Dern, Christa (2003). Unhöflichkeit ist es nicht. Sprachliche Höflichkeit in Erpresserschreiben. *Deutsche Sprache*, 31(3), 127–141. DOI: [10.37307/j.1868-775X.2003.02.03](https://doi.org/10.37307/j.1868-775X.2003.02.03).
- Dern, Christa (2009). *Autorenerkennung. Theorie und Praxis der linguistischen Tatschreibenanalyse*. Stuttgart: Boorberg.
- Esser, Jürgen (2009). *Introduction to English Text-linguistics*. Frankfurt a. M.: Peter Lang.
- Ehrhardt, Sabine (2018). Authorship attribution analysis. In Visconti (Ed.), *Handbook of Communication in the Legal Sphere* (pp. 169–200). Boston/Berlin: de Gruyter.
- Feilke, Helmuth (2000). Die pragmatische Wende in der Textlinguistik. In Brinker, Antos, Heinemann & Sager (Eds.), *Handbuch Text- und Gesprächslinguistik: Ein internationales Handbuch zeitgenössischer Forschung* (pp. 64–82). Berlin/New York: de Gruyter.
- Fetzer, Anita (2018). Discourse analysis. In Jucker, Schneider & Bulbitz (Eds.), *Methods in Pragmatics* (pp. 395–423). Berlin: de Gruyter Mouton.
- Fix, Ulla (2004). Stil gibt immer etwas zu verstehen. *Der Deutschunterricht*, 56(1), 41–50.
- Fobbe, Eilika (2011). *Forensische Linguistik. Eine Einführung*. Tübingen: Narr.
- Fobbe, Eilika (in press) (2021). *Stilkonzepte in der automatisierten Autorschaftsattributions im forensischen Kontext*. In Luttermann & Busch (Eds.), *Recht und Sprache: Konstitutions- und Transferprozesse in nationaler und europäischer Dimension* (Rechtslinguistik, 11). Berlin/Münster/London: LIT.
- Grant, Tim (2007). Quantifying the Evidence in Forensic Authorship Analysis. *The International Journal of Speech, Language and the Law*, 14(1), 1–26. DOI: [10.1558/ijssl.v14i1.1](https://doi.org/10.1558/ijssl.v14i1.1).
- Grant, Tim (2008). Approaching Questions in Forensic Authorship Analysis. In Gibbons & Turell (Eds.), *Dimensions of Forensic Linguistics* (pp. 215–230). Amsterdam/Philadelphia: John Benjamins.
- Hansen, Sandra (2009). Verwaltungssprache in Erpresserbriefen. In Bachmann-Stein & Stein (Eds.), *Mediale Varietäten – Analysen von gesprochener und geschriebener Sprache und ihre fremdsprachlichen Potenziale* (pp. 295–312). Landau: Verlag Empirische Pädagogik.
- Iqbal, Farkhund, Hadjidj, Rachid, Fung, Benjamin C. M. & Debbabi, Mourad (2008). A novel approach of mining write-prints for authorship attribution in e-mail forensics. *Digital investigation* 5, 42–51. DOI: [10.1016/j.diin.2008.05.001](https://doi.org/10.1016/j.diin.2008.05.001).
- Kniffka, Hannes (2007). *Working in Language and Law. A German perspective*. Basingstoke/New York: Palgrave MacMillan.
- König, Ekkehard & Pfister, Manfred (2017). *Literary Analysis and Linguistics*. Berlin: Erich Schmidt.
- Levinson, Steven C. (1983). *Pragmatics*. Cambridge: CUP.
- McMenamin, Gerald R. (2010). Forensic stylistics. Theory and practice of forensic stylistics. In Coulthard & Johnson (Eds.), *The Routledge Handbook of Forensic Linguistics* (pp. 487–507). London/New York: Routledge.
- Püschel, Ulrich (2009). Stil: Theorie und Praxis. *Zeitschrift für Angewandte Linguistik*, 50, 47–68.
- Sandig, Barbara (2006). *Textstilistik des Deutschen*. 2nd ed. Berlin/New York: de Gruyter.
- Shuy, Roger (2006). *Linguistics in the Courtroom. A Practical Guide*. Oxford: OUP.
- Spillner, Bernd (2009). Verfahren stilistischer Textanalyse. In Fix, Gardt & Knape (Eds.), *Rhetorik und Stilistik. Ein internationales Handbuch historischer und systematischer Forschung* (pp. 1739–1782). Berlin/New York: de Gruyter Mouton.
- Stamatatos, Efstathios (2009). A Survey of Modern Authorship Attribution Methods. *Journal of the American Society for Information Science and Technology*, 60(3), 538–556. DOI: [10.1002/asi.21001](https://doi.org/10.1002/asi.21001).
- Stamatatos, Efstathios (2018). Masking Topic-Related Information to Enhance Authorship Attribution. In *Journal of the Association for Information Science and Technology* 69(3), 461–473. DOI: [10.1002/asi.23968](https://doi.org/10.1002/asi.23968).

- Stein, Stephan & Baldauf, Christa (2000). Feste sprachliche Einheiten in Erpresserbriefen. Empirische Analysen und Überlegungen zu ihrer Relevanz für die forensische Textanalyse. In *Zeitschrift für Germanistische Linguistik* 28(3), pp. 377–403. DOI: [10.1515/zfgl.2000.28.3.377](https://doi.org/10.1515/zfgl.2000.28.3.377).
- Toolan, Michael (2014). The Theory and Philosophy of Stylistics. In Stockwell (Ed.), *The Cambridge Handbook of Stylistics* (pp. 13–31). Cambridge: CUP.
- Wright, David (2013). Stylistic variation within genre conventions in the Enron email corpus: developing a text sensitive methodology for authorship research. *The International Journal of Speech, Language and the Law* 20, 1, 45–75. DOI: [10.1558/ijssl.v20i1.45](https://doi.org/10.1558/ijssl.v20i1.45).

Note: JLL and its contents are Open Access publications under the [Creative Commons Attribution 4.0 License](https://creativecommons.org/licenses/by/4.0/).



Copyright remains with the authors. You are free to share and adapt for any purpose if you give appropriate credit, include a link to the license, and indicate if changes were made.

Publishing Open Access is free, supports a greater global exchange of knowledge and improves your visibility.