

# J. Berenike Herrmann, Karina van Dalen-Oskam, Christof Schöch Revisiting Style, a Key Concept in Literary Studies

**Abstract:** Language and literary studies have studied style for centuries, and even since the advent of ›stylistics‹ as a discipline at the beginning of the twentieth century, definitions of ›style‹ have varied heavily across time, space and fields. Today, with increasingly large collections of literary texts being made available in digital form, computational approaches to literary style are proliferating. New methods from disciplines such as corpus linguistics and computer science are being adopted and adapted in interrelated fields such as computational stylistics and corpus stylistics, and are facilitating new approaches to literary style.

The relation between definitions of style in established linguistic or literary stylistics, and definitions of style in computational or corpus stylistics has not, however, been systematically assessed. This contribution aims to respond to the need to redefine style in the light of this new situation and to establish a clearer perception of both the overlap and the boundaries between ›mainstream‹ and ›computational‹ and/or ›empirical‹ literary stylistics. While stylistic studies of non-literary texts are currently flourishing, our contribution deliberately centers on those approaches relevant to ›literary stylistics‹. It concludes by proposing an operational definition of style that we hope can act as a common ground for diverse approaches to literary style, fostering transdisciplinary research.

The focus of this contribution is on literary style in linguistics and literary studies (rather than in art history, musicology or fashion), on textual aspects of style (rather than production- or reception-oriented theories of style), and on a descriptive perspective (rather than a prescriptive or didactic one). Even within these limits, however, it appears necessary to build on a broad understanding of the various perspectives on style that have been adopted at different times and in different traditions. For this reason, the contribution first traces the development of the notion of style in three different traditions, those of German, Dutch and French language and literary studies. Despite the numerous links between each other, and between each of them to the British and American traditions, these

---

**J. Berenike Herrmann:** Department of German Philology, Georg-August-University Göttingen,  
E-Mail: bherrma1@gwdg.de

**Karina van Dalen-Oskam:** Department of Literary Studies, Huygens Institute for the History of the Netherlands, University of Amsterdam, E-Mail: karina.van.dalen@huygens.knaw.nl

**Christof Schöch:** Department of Literary Computing, Würzburg University,  
E-Mail: christof.schoech@uni-wuerzburg.de

three traditions each have their proper dynamics, especially with regard to the convergence and/or confrontation between mainstream and computational stylistics. For reasons of space and coherence, the contribution is limited to theoretical developments occurring since 1945.

The contribution begins by briefly outlining the range of definitions of style that can be encountered across traditions today: style as revealing a higher-order aesthetic value, as the holistic ›gestalt‹ of single texts, as an expression of the individuality of an author, as an artifact presupposing choice among alternatives, as a deviation from a norm or reference, or as any formal property of a text. The contribution then traces the development of definitions of style in each of the three traditions mentioned, with the aim of giving a concise account of how, in each tradition, definitions of style have evolved over time, with special regard to the way such definitions relate to empirical, quantitative or otherwise computational approaches to style in literary texts. It will become apparent how, in each of the three traditions, foundational texts continue to influence current discussions on literary style, but also how stylistics has continuously reacted to broader developments in cultural and literary theory, and how empirical, quantitative or computational approaches have long existed, usually in parallel to or at the margins of mainstream stylistics. The review will also reflect the lines of discussion around style as a property of literary texts – or of any textual entity in general.

The perspective on three stylistic traditions is accompanied by a more systematic perspective. The rationale is to work towards a common ground for literary scholars and linguists when talking about (literary) style, across traditions of stylistics, with respect for established definitions of style, but also in light of the digital paradigm. Here, we first show to what extent, at similar or different moments in time, the three traditions have developed comparable positions on style, and which definitions out of the range of possible definitions have been proposed or promoted by which authors in each of the three traditions.

On the basis of this synthesis, we then conclude by proposing an operational definition of style that is an attempt to provide a common ground for both mainstream and computational literary stylistics. This definition is discussed in some detail in order to explain not only what is meant by each term in the definition, but also how it relates to computational analyses of style – and how this definition aims to avoid some of the pitfalls that can be perceived in earlier definitions of style. Our definition, we hope, will be put to use by a new generation of computational, quantitative, and empirical studies of style in literary texts.

# 1 Introduction

In emerging interdisciplinary fields such as Digital Humanities (DH), even basic terminology can be treacherous. Technical terms may look unambiguous, but in fact have different meanings in each of the original disciplines. In fact, there may be little consensus on a term's definitions even within research traditions. We have experienced this to be the case when discussing ›style‹, a notion that is not at all unanimously defined in the study of different languages and literatures. This situation has, as a rule, provoked two opposing types of reactions:

- (a) hard-boiled debates about how to define style properly on a theoretical level;
- (b) the avoidance of attempting a proper conceptualization, or even operational definition, of style.

Both of these are problematic. While reaction (a) is often fruitless (at least when not coming to an end, for example by operationalization), reaction (b) makes it hard to make valid claims (when not properly defining the object of study). When talking about style with regard to literary texts, one is thus faced with a (communicative) problem that endangers, or at least slows, cross-disciplinary collaboration and the development of new research possibilities. We are certainly not the first scholars to note this problem. Yet, regarding the rapid development of digital studies of style of the past few years, it is high time to revisit the definition of style for literary stylistics, taking into account existing research traditions as well as the new developments in digital humanities research.

This article centers on a discussion of the notion of style for research that is empirical in the widest sense, i. e., understood as the acquisition of knowledge through qualitative or quantitative observation of stylistic features of text.<sup>1</sup> Our

---

<sup>1</sup> We understand ›empirical‹ in the wide sense – as the »acquisition of knowledge through observation« (»Wissensgewinn durch Beobachtung«, Eibl 2013, 23). At the most general level, this wide notion of empirical subsumes both the ›quantitative‹ and ›qualitative‹ empiricism (cf. Groeben 2013). A ›quantitative‹, experimental research methodology revolves around an ideal of reducing intuition and speculation to zero, while ›qualitative‹ approaches explicitly incorporate a role of intuition in the research process. In order to ensure quality, stylisticians concerned with research methodology have established basic criteria for ›the stylistic method‹, such as ›rigorous, retrievable, and replicable‹ (Simpson 2004; see also Wales 2001, 373). To Simpson (2004, 4), *rigorous* »means that it [the stylistic method] should be based on an explicit framework of analysis«; *retrievable* means that »the analysis is organised through explicit terms and criteria, the meanings of which are agreed upon by other students of stylistics«; and *replicable* »means that the methods should be sufficiently transparent as to allow other stylisticians to verify them, either by testing them on the same text or by applying them beyond that text«. While the qualitative and the quantitative empiricism diverge for example in the role

aim is twofold: (1) We provide a review (from the broad perspective of a ›digital stylistics‹) of the ways in which style has been defined since 1945 in Dutch, French, and German language and literary studies. Here, we trace back the individual traditions with the aim of a synthetic overview across traditions; (2) we introduce an operational definition of style that incorporates a minimal common ground for interdisciplinary empirical research and the application of new, digital methods. The rationale is to work towards a common ground for literary scholars and stylisticians when talking about style, across different traditions of scholarship, in light of the new, digital paradigm.

We decided to approach our subject in this way, focusing on German, French, and Dutch studies, because we can tap into our experience as scholars of these disciplines. The tradition(s) of stylistics from the Anglo-Saxon countries will hence not be at the center of this overview. However, as they are both eminently influential and productive in international and interdisciplinary (digital) stylistics, they will serve as a backdrop against which the three traditions will be described.<sup>2</sup> Although our focus on three traditions involves a somewhat arbitrary limitation, it has a number of benefits:

- Choosing a comparative approach ›by tradition‹ (instead of solely a systematic synthesis) enables us to show the internal development of each tradition. At the same time, it allows us to explore how the traditions have diverged in giving prominence to stylistic research, which helps to explain why in certain areas less fundamental work is done than elsewhere.
- The exercise allows gauging how dissimilar or how similar definitions of style have been in international comparison. With our approach we can highlight divergences and convergences, as well as common references between the distinct discourses.
- Stylistics needs to factor in differences between particular language communities, with their particular language and usage systems (cf. Sandig 2006). Literary traditions have to be added to the picture. It thus appears sensible to approach (literary) style through the lense of distinct scholarly

---

assigned to intuition, they confer in being data-based, transparent, and demanding tests of the relation between data and conclusions drawn from them. In DH, both approaches can be observed. In our notion of ›empirical‹, we explicitly include both quantitative and qualitative stylistics.

<sup>2</sup> The Anglo-Saxon literature on (literary) style and stylistics of the past 30 years is extensive. For an overview of the different research traditions, see the latest overviews by Crystal (2010), Nørgaard et al. (2010), Simpson (2004), and also Wales (2001). Selected approaches will be described below, or at least mentioned *en passant*, especially in the sections on digital approaches.

discourses across languages, and at the same time transgress the border between these discourses.

- Together with the Anglo-Saxon one(s) as a reference point, the three chosen discourses on style are quite representative for (European) stylistics. Other ›national‹ traditions may be easily compared to the picture drawn.

We hence do not attempt to provide ›the history‹ of style for any of these three traditions (or four, considering the Anglo-Saxon one). Rather, we depict their respective internal development in order to delineate those approaches that have influenced contemporary positions on literary style.

For reasons of space and coherence, we decided to focus on the developments in stylistics of the past seventy years (since 1945). This decision is supported by the new-formation of university life after WWII in Germany, but also the other countries. It can also be defended with regard to the development of 20th-Century stylistics that spread through Europe gradually: the impact of Russian formalism and early structuralism was in large parts not received until well after the Second World War. In the following, however, reference will be made to important definitions older than seventy years where needed.

Attempting an overview of definitions of style for more than one country or discipline on less than fifty pages produces imbalances of various kinds, as will our foray into the German, French and Dutch traditions that for example depicts more German (speaking) and French definitions than Dutch ones. In part, this reflects the sheer size of the respective discourse communities, in part the interest that has gone into researching aspects of literary style in the respective language.<sup>3</sup>

A millennium before its career as an – at times hotly – debated notion in the humanities, the term *style* had a quite mundane beginning as a concrete tool: Latin *stilus* denominated »[a]n instrument made of metal, bone, etc., having one end sharp-pointed for incising letters on a wax tablet, and the other flat and broad for smoothing the tablet and erasing what is written«.<sup>4</sup> By metonymy, *style* acquired the meaning of *modus scribendi/dicendi* (cf. Sowinski 2007, 1393), the *way of writing/saying something*, which has stayed the center of the notion's meaning in language and literary studies. At the most general level, *style* can mean ›features‹ that ›characterize‹ somebody or something. Style has hence been

---

<sup>3</sup> Dutch and Flemish scholars who have predominantly worked on languages other than Dutch and at the same time have not been discussed in the discourse on Dutch literary style are not in the focus of this contribution. However, many such scholars appear in the Anglo-Saxon tradition (e. g., Louwerse, Hakemulder, Van Peer, or Steen). For a more detailed overview of the study of style in the Netherlands see Fagel-de Werd 2015.

<sup>4</sup> *Oxford English Dictionary Online*, <http://www.oed.com/view/Entry/192315>.

restricted to the factual linguistic features of some text; it has also been more abstractly used to capture thought, personality, and ideology (often with little concern for linguistic features); it has been conceptualized to denote unity – or difference. It has been treated from a normative/evaluative approach (which we will not cover in this contribution) as well as from a descriptive one (for a comprehensive overview of the current research on style, see the two volumes of Fix et al. 2008/2009).

Before moving on, we would like to note that, presupposing (which we do in this article) a focus on *style as language use at the textual level*, and excluding prescriptive approaches as well as views of style centering on readers' reception<sup>5</sup>, there seems to be a limited number of recurring types of definitions of style; these are definitions of style as:

1. constituting a higher-order artistic value (assessed through aesthetic experience),
2. a holistic gestalt of single texts,
3. an expression of individuality, subjectivity and/or emotional attitude of an author or speaker,
4. an artifact that presupposes (hypothetical or factual) selection/choice among a set of (more or less synonymous) alternatives,
5. a deviation from some type of norm, involving (quantitative or cognitive) contrast,
6. any property of a text that can be measured computationally.

This set of definitions of style will re-appear in one or another form in the following reviews of thinking about style in German, Dutch and French linguistics and literary studies, be it explicitly or implicitly, and we will return to it in a synthesis section.

---

<sup>5</sup> This contribution centers on style definitions that are valid at the textual level of usage. While the analytical distinction between the levels of reader response and textual structure is eminently important, we are aware that many scholars have not made the distinction (either implicitly or explicitly). For reasons of focus and theoretical rigor, we will not consider those approaches that deal explicitly with style at the level of the reader's processing, such as psychological theories of style processing and cognitive-linguistic theories in their pure form. We do, however, include pragmatic theories – since they are pitched to a substantial degree at the textual level of usage.

## 2 German Linguistics and Literary Studies

In Germany, Switzerland and Austria, the *Werkimmanenz* (Kayser 1948, Staiger 1955), a text-centered approach that largely does without extra-textual resources, has influenced literary studies of style to the day. In its classical version, *Werkimmanenz* treated style as a synthetic category that could only be grasped through affectively driven aesthetic experience. The basic unit of analysis was not an objective ›text‹, but something of a higher value, the ›literary work‹, which possessed an almost spiritual power that lay in its ›unity‹ (*Einheitlichkeit*,<sup>6</sup> Staiger 1955). Underlying *Werkimmanenz*'s approach was hence a traditional conception of aesthetic experience as a precondition for aesthetic judgment, that – going all the way back to Kant's *Critique of Judgment* (first 1790) – explicitly involves affective experience. To Staiger, style is the ›Ineffable-Identical‹ (*Unaussprechlich-Identische*, Staiger 1955, 23), ›wherein a perfect work of art – or the entire work of an artist or his time – agrees in all its aspects‹ (ibid., 14). To him, art structures are ›perfect when they are stylistically unanimous‹ (*stilistisch einstimmig*, ibid.). Since the identification of style is restricted to ›perfect‹ works of art, *Werkimmanenz* essentially conflates the aesthetic experience with aesthetic judgment. What is more, the scholar ›feels‹ when a stylistic analysis fails – because of a lack of stylistic unanimity (ibid., 13). Kayser (1948) proposes style as a holistic entity as well, but in contrast to Staiger, his method engages to a greater extent in the elaborate description of textual elements. His aesthetic account of style influenced a whole new text-centered aesthetic approach to text interpretation in literary studies, with his monograph seeing twenty editions until 1992. Style, according to Kayser (ibid., 300), could be identified in particular works, but also in specific poets or epochs. Although some aspects of *Werkimmanenz* actually correspond with a descriptive-empirical concentration on textual, formal features, its definition of style is not empirically testable, for it stays essentially fuzzy (›the ineffable-identical‹). Also, its method depends on the gifted scholar's introspective-aesthetic capacities for identifying style, which eventually means that it conflates aesthetic experience and aesthetic judgment (agreeing with Staiger's view of style as an epitome of the perfect work of art). From an empirical point of view, the prior is problematic because of the lack of intersubjectivity/objectivity, transparency, and reliability in the identification and analysis of style. The latter is problematic because it involves an implicit quality judgment. Today, the assumption that the identification of relevant style ele-

---

<sup>6</sup> Translations from the original language by the authors. References indicate the original edition. Holds for all following quotations except when indicated differently.



ments cannot be separated from interpretation (and implicit value judgments) is still widespread in German literary studies (e. g., Anderegg 2008), with the »intensive immersion in the individual work of art« (Szondi 1967, 21) as the basic mode of insight.

To many, the most central figure in German stylistics is the Austrian Romance scholar Spitzer (1961). Spitzer, linguist by training, was an early proponent of a literary-linguistic approach to literary texts, seeing stylistics as the »most viable bridge between language and literature« (1961, 3–4). His method incorporates an elaborate linguistic analysis followed by a literary interpretation, aiming at »capturing the language of poets in their artistic intentions, characterizing it, and tracing it back to the soul, which is linguistically expressed by the poets« (*das Seelische, das die Dichter sprachlich ausdrücken*, *ibid.*, 4). He combined the description of linguistic »systems of expression« (*ibid.*, IX) with that of conceptual-cognitive »stylistic characteristics of certain attitudes« (*ibid.*). He maintained a perspective on style as an artifact used purposefully by authors. His approach, described as »hermeneutic stylistics«, was taken up by many, albeit often without explicit reference to Spitzer.

Where Spitzer promoted a deliberate reliance on subjectivity in the interpretation, structuralist approaches took a strong stance against it. In the 1960s, they took over the debate on style and literature, promoting strictly scientific-empirical criteria and applying linguistic measures. Style was now seen as a textual (or reception) phenomenon that could be objectively and exactly described. One important dimension of the discussion was the call for less reliance on impressionistic value judgments (Riffaterre 1973), which for many meant refraining from interpretation in general (e. g., Spillner 1974, see also Schmidt 1979).

Especially in Western Germany, style was now related to the concept of *selection* (e. g., Sanders 1973), where a sender (hypothetically) selects facultative variants on the paradigmatic axis and combines these on the syntagmatic axis. Then, the receiver decodes the syntagma and classifies the separate elements with regard to paradigms: this is where the receiver is taken to assign stylistic value to the elements (*ibid.*, 69). The problem of how exactly the stylistic quality of language use (*parole*) is constituted in relation to the language system (*langue*) was resolved in a different way in Eastern Germany: functionalist accounts (e. g., Michel et al. 1968, Riesel 1959) saw style as mediated by social norms. Style, according to them, refers to the relation between the linguistic sign and the human being – not to the sign or the sign system itself – and this relation is mediated by »societal norms of usage« (Michel et al. 1968, 34–35). It is these norms that assign stylistic value to the *parole*, which in turn carries out distinct functions.



A second influential set of structuralist theories received in German literary studies are *deviance* theories of style.<sup>7</sup> The basic idea is that literary works deviate from some assumed norm. However, since ideas about norms may imply the existence of a ›solid‹ system of language, any deviationist account has had to take a stance with regard to a sufficient notion of norm/rule. Riffaterre's (1973) influential version of the deviation theory proposed as a solution a *contextual* notion of style. It focuses on deviation within the context constituted by the text itself, incorporating the reader's anticipation as constituting a contextually primed norm. Similar in its focus on ›context‹ is Enkvist's difference-theoretic variationist notion of style, stressing that »the essence of variation, and thus of style, is difference, and differences cannot be analysed and described without comparison« (Enkvist 1973, 21). Describing a particular style, one matches the text under scrutiny »against another body of texts which we might label as *norm*, this norm being chosen because it is contextually relevant as a background for the text« (ibid., 25–26, emphasis in original).

Despite evident continuities, by the late 1970s, mainstream literary studies declared ›dead‹ the interdisciplinary stylistics project called *LiLi* (*Literature and Linguistics*, cf. Fix 2010), and either returned to (neo-)scholastic rhetoric as a reference point (cf. the influential textbook by Plett [latest edition 2001] that made available the Anglo-Saxon study of rhetoric in German studies) or moved on to post-structuralist and cultural-studies perspectives, which did not focus on patterns and functions of linguistic forms in the strict sense. Generally, it was felt that the linguistic (i. e. structuralist) definition of both literature and style did not capture enough of the complexity and contextuality of literature (cf. Fix 2010). Selection theories that (at least hypothetically) assign intentionality to subjects in the literary communication process became especially unfashionable, much as the talk of linguistic or stylistic ›norms‹. At the same time, (Chomskyan) linguistics focused on relatively small units of analysis, and aspects of ›style‹ of rich literary texts became rather unattractive.

However, in the 1980s, pragmatic stylistics appeared, somewhat removed from mainstream literary studies, with a »holistic perspective on linguistic behavior in all its complexity« (Püschel 2008, 175). Mostly empirical, often along the lines of (qualitative) text- and discourse linguistics, pragmatic stylistics tends to literary,

---

<sup>7</sup> A third strand of structuralist stylistics received in Germany are connotation theories (see also French section below). One proponent in Germany is Graubner (1973), who differentiated distinct types of lexical connotations, which correspond largely with specific communicative functions of language in the sense of Bühler and Jakobson, but incorporate among other things »Stilzüge«, text-immanent coherent style effects (such as *solemnity*, *joy*, or *irony*).

but also non-literary texts. Scholars followed up on the functional dimensions of style that were already present in Russian formalism and were now invigorated by linguistic pragmatics (Austin, Searle). One central, but by no means new, tenet of pragmatic stylistics is that style is understood as an artifact, produced in interaction with stylistic rules or patterns and general conventions of language use (cf. Sowinski 2007). Style is hence not restricted to literary or openly creative, artistic language or text types; the fundamental conventionality of language use predicts that even slight variations in phrasing can have an effect on how utterances are perceived. This includes rhetorical figures, but is not limited to them. Therefore, style is not seen as *ornatus* (a facultative embellishment), but as an unavoidable dimension of language use. Pragmatic accounts, by contrast to many structuralist ones, hence emphasize strongly the communicative function of style, with its capacity of transporting evaluations and shaping perspectives.

This approach is pursued for example by Sandig (2006, first edition 1986), who maintains a broad definition of style. In her textlinguistic approach, style is seen as the »meaningful functional and situational variation in the use of language and other communicatively relevant types of signs« (Sandig 2006, 1). Style is modeled as *performance* (*Durchführen*), »the way in which a socially relevant and meaningful communicative action is carried out« (ibid., 9). Here, the structuralist perspective of choice is still important: »in uttering (a sentence, text, etc.), individual elements are used from the provided inventories: [they are] preferred to other options that would also be provided by the language« (Sandig 1986, 43). Sandig's approach can be compared to Fix's theory that puts a focus on the product of the performance, the aspect of »*Gestalten*« (cf. Fix 2007). The predicate *gestalten* [configure, shape] refers to a text's particular gestalt/sum of elements. Fix maintains a holistic notion of style and a method that strives for a unifying interpretation (ibid., 83).

In pragmatic approaches, every aspect of the linguistic action that can be perceived as »made« pertains to the subject of stylistics. Therefore, the wide pragmatic notion of style has been criticized for inducing a »hypertrophy of stylistic issues« (cf. Püschel 2008, 175). Many pragmatic approaches to style incorporate the empirical examination of style phenomena, describing »typical« or individual dimensions of style, genres and registers, incorporating formal linguistic descriptions of style in texts in terms of lexis, grammar, wording and style figures, topic, text patterns (text types) and aspects of their materiality (cf. Sandig 2006, 2–3). Most are qualitative approaches, but there are also quantitative ones, often incorporating aspects of cultural studies and critical theory. Some, however, run the danger of circularity when presupposing intentions and functions for the analysis of formal features, which in turn are used to buttress the assumptions about intentions and functions.

Finally, an increasingly popular approach to style is the computational one, assessing features of writing style on the ›text surface‹ that can be processed by the computer (e. g., frequencies of words, sentences, but also word classes and other descriptive categories). Despite its marginal position in literary studies, an entry ›Statistische Literaturanalyse‹ is featured by the *Reallexikon*, and the use of statistical methods is seen as helping the ›objectivization of stylistic evidence‹ (Pieper 2003, 502). In general, statistical approaches to style adhere to the maxims of the scientific method, with three levels of empirical adequacy: of observation, of description, and of explication (ibid., 501).

One important sub-strand of statistical style studies in Germany is stylometry, or ›mathematical‹ literary studies (cf. Fucks 1955, Kreuzer/Gunzenhäuser 1965). It has traditionally concentrated on issues of authorship,<sup>8</sup> but its methods have been applied also to questions of literary history, to the description of author, genre and period style (e. g., Jannidis/Lauer 2014), and descriptive accounts of individual characters' style. An early account of the empirical psychology of style is Busemann (1948), who incorporated statistical measures of style. Another subfield of computational stylistics is corpus stylistics, with one of the few proponents in the German-speaking field being Müller (2012), whose study on creative metaphor applies a corpus-linguistic version of deviance theory to style.

For mathematical stylistics, ›style means the totality of quantitatively assessable elements in the so-called ›formal structure‹ of a text‹ (Kreuzer 1965, 11). This totality, however, has had to remain hypothetical, since there is ›no complete list of the linguistic properties of a text‹ (Leech/Short 2007, 56). Stylisticians of all denominations hence ›have to select the features to study‹ (ibid.). In German studies, for example, Jannidis and Lauer (2014) select the feature ›word frequencies‹ in their literary-historical analysis, proposing to describe traditional categories such as ›styles of authors‹, ›writing styles of larger groups or literary epochs‹, and ›narrative‹ and ›dramatic style‹ (ibid., 50). Herrmann (2013) uses a reference corpus of newer German literature as a basis for singling out linguistic features characteristic of Franz Kafka's prose.

---

<sup>8</sup> Naturally, after Foucault and Barthes, who stressed the role of discursive and intertextual forms of meaning creation, quantitative authorship attribution in literary texts has had to deal with the epistemological status of authorship itself.

### 3 Dutch Linguistics and Literary Studies

Literary stylistics has never been at the forefront in The Netherlands and Flanders (cf. Fagel 2008). Only a handful of individual scholars have published about the topic in a more than superficial manner. The first important publication in our time range is a short book written by Stutterheim (1947). He states that stylistic research has changed in the last decades: A very strict and concrete way of describing minutely anything relating to the language has given way to a more impressionistic, and thus subjective, description. For Stutterheim, style is an individual, original, use of language, which establishes itself as a deviation from normal language use, but it is difficult to indicate at which point exactly language use becomes style. In evaluating linguistic expressions, the scholar makes optimal use of his sensitiveness to language, his impressions and intuition. The stylistician describes the many ways in which language can be used to express emotion. Stutterheim shows influence from Spitzer, but only acknowledges him by mentioning his *Stilstudien* (first 1928) in his bibliography.

In contrast to Stutterheim's integrated approach from both linguistic and literary perspectives, the next important publication on style, by Hellinga and Van der Merwe Scholtz (1955), clearly distinguishes the roles of each of these disciplines: The linguistic scholar gives a description of individual instances of language use. The literary scholar then labels these instances, evaluates them, and compares them with other instances. Hellinga and Van der Merwe Scholtz favor the linguistic approach, in which style is seen as deviation not from a linguistic norm, but from the current conventions of a language system, depending on the context. They refer to publications by Kayser, Staiger, and Spitzer, but take a much more descriptive and objective stance in their own work than these scholars do.

Of main importance for the current approach to style is an article written by Anbeek and Verhagen (2001). They refer to the fact that many literary scholars refuse to analyze style by stating that the uniqueness of style can only become clear once the norm has been established. They propose a contrastive approach, inspired by Leech and Short's (2007) *Style in Fiction* and as such, introduce this scholarly work to a wider Dutch audience. As an example, Anbeek and Verhagen contrast two recent novels by looking at sentence length, structure of the sentences, and semantics of the described situation in the sentences. Their description is predominantly qualitative and only has some very basic quantitative information as a kind of aside, and thus merges linguistic observations and interpretation of the text. The strength of this approach for literary studies lies, according to Anbeek and Verhagen, in the fact that it can lead to a reinterpretation of the concept of style: »In this way the stylistic analysis primarily reveals a

connection between characteristic uses of language and characteristic elements of the contents of the text, and not between language use and author« (Anbeek/Verhagen 2001, 23).

This new take on style has influenced later scholars. Van Driel (2007) analyzes a small set of stylistic features in Middle Dutch epic poetry by contrasting the different texts. Fagel (2009, 2010) focuses on the analysis of a set of all-pervasive linguistic elements in a very small corpus in a qualitative (impressionistic, non-quantitative) way. The results are then seen as a hypothesis that needs quantitative testing, which is the topic of Fagel et al. (2012). One of their observations is that some of their intuitions proved to be wrong, but led them to look at the data in a different way to try to find out why the results were unexpected. This led to new observations, from which they learned that »Thus, apart from an interpretative function, quantitative analysis in stylistic research can also have a heuristic function« (ibid., 196).

Most literary scholars, when dealing with style, do not go into the problems of the concept but refer to the definition of style in the *Algemeen letterkundig lexicon* (General literary lexicon) edited by Van Bork et al. (2012) and its precursors. The definition, translated from the lemma ›stijl‹ is:

General term for outward characteristics of a way of writing or speaking; the phenomenon thus belongs to elocutio. Style first of all concerns the characteristic way in which someone expresses himself in language, be it prose or verse. This kind of expression can be any imaginable feature of language: choice of words, the use of tropes, sentence structure and composition, structure, tone of voice, connotation/association, narrative technique, etc. This sense of personal style can be broadened to the ways in which groups of people, i. e. authors, express themselves (e. g. the style of the ›Poets of 1880‹) or even whole people (idiomatic language use, e. g. the differences between typically French or German style). Furthermore the term is also used for the characteristic expression of a certain time period (period style, e. g. baroque style) or a genre (genre style, e. g. epistolary style).

This definition is clearly based on a transdisciplinary approach, referring to Spitzer, Stutterheim, Anbeek and Verhagen and several publications from ›literary linguistics‹ (stylistics; e. g., Lambrou/Stockwell 2007) and from ›traditional‹ literary studies. The central piece of the definition, »the characteristic way in which someone expresses himself in language«, however, uses the adjective »characteristic« in a rather vague way, inviting different interpretations. This ambiguity would ideally be avoided in a new definition.

Stylometry (computational stylistics) is a recent addition to Dutch literary studies. Like German linguistics and literary studies, this discipline until now has mainly focused on authorship attribution (Kestemont 2011, Van Dalen-Oskam/Van Zundert 2007). In her inaugural lecture *De stijl van R* (*The style of R*), dealing with a.

o. a computational approach to the stylistic voices of letter writers in a set of Eighteenth-century Dutch epistolary novels, Van Dalen-Oskam writes about style:

In Digital Humanities, ›style‹ is seen as anything that can be measured in the linguistic form of a text, such as vocabulary, punctuation marks, sentence length, word length, the use of character strings. [...] Every word and every feature contributes to the general outlook of the text; any other ratio in frequencies, any difference in mean sentence length, every individual punctuation use results in a different outlook of the text. In short: everything is important. (Van Dalen-Oskam 2012, 6)

## 4 French Linguistics and Literary Studies

In French linguistics and literary studies, stylistics and the notion of style have been for many decades matters of intense debate, as evidenced by extensive publications on the matter as well as numerous synthetic works on stylistics (e. g., Guiraud 1963, Larthomas 1998). In the 1950s, well into the 1960s and even beyond, the influence of two important branches of earlier stylistics continues to be felt: first, the narrowly defined ›stylistique de l'expression‹ by Swiss linguist Bally (1921), which focused on the general potentialities of (the French) spoken language to convey more or less emotionally charged messages. Second, the ›stylistique idéaliste‹ associated with Spitzer (1961, 1988), whose particular way of combining attention to linguistic detail with a hypothesis about the unifying principle in an author's literary work has been discussed above.

The renewal brought to linguistics and literary studies by structuralism made itself felt in French stylistics as well. A famous application of structuralist principles to stylistic analysis is no doubt Jakobson and Lévi-Strauss's reading of *Les Chats* by Baudelaire, which shows how different levels of description of the poem all interrelate to make the poem ›un objet absolu‹, i. e. a closed system of interrelated elements (Jakobson/Lévi-Strauss 1962, 17). The commutative method developed by Dupriez (1971) is also structuralist in inspiration, making use of the idea of the paradigmatic and syntagmatic axes of any statement. The *Essais de Stylistique Structurale* by Riffaterre (1971), transgressing in fact the text as only object of stylistics and opening it up again to the reader's perspective, has also been avidly received in France.

Towards the end of the 1960s, however, it seemed to many that stylistics (as the linguistic description of language use, especially in literary texts) was ›more or less dead‹ (*à peu près morte*, Arrivé 1969, 3; he retracts that statement at the

end of his article).<sup>9</sup> If not dead, stylistics seemed to have been replaced by other adjacent fields such as semiotics, narratology, and poetics. At the very least, the usefulness or even possibility of the linguistic description of literary texts as a specific field was questioned in the light of positions defending either the irreducibility of individual texts (or of specific stylistic phenomena) to any conceptual scheme, or the lack of specificity of literary language use as opposed to everyday language use (see Arrivé 1969).

Despite its disappearance from the front of the scene in the 1970s and 1980s, stylistics in fact continued to be widely practiced in France, not least because it continued to be an important element of the reading practices ingrained in school and university culture, something which also meant that school- and university-level textbooks continued to be published. However, a truly renewed academic interest in stylistics only became felt at the beginning of the 1990s, when several important publications appeared. Style, as Compagnon (1998) has shown, is one of a number of concepts in literary studies that survived the *années théoriques* of post-structuralism and continue to thrive and be useful.

What definitions of style, then, have been prominent in the field over the last decades? Probably one of the oldest definitions of style in the French tradition is that of style in ›parole‹ as a choice among the variation allowed by ›langue‹ when conveying a specific meaning. This definition of style is present in Bally's works and was perpetuated well into the second half of the twentieth century. Bally distinguishes between an abstract or logical content on the one hand, and the specific, more or less intense ›valeur affective‹ a given linguistic expression may add to that content (Bally 1921, 1–17).<sup>10</sup> Such a theory implies a decoupling of thought and language, of meaning and style, and therefore relies on the possibility of synonymy: one given abstract meaning can be expressed in several ways, which change not the meaning, but only the emotional attitude carried with that meaning. In this sense, this view of style is a more modern version of the much older idea of style as an ornament or additional coloring added to a preexisting meaning which could also have been expressed in a neutral form.

Closely connected to Bally's view of style is a somewhat more general view of meaning in language based on the distinction between denotation and connotation in semiotics (e. g. Barthes 1964, 130–132). This is sometimes seen as a

---

<sup>9</sup> Barthes, as early as 1953, dismissed ›style‹ as being too bound up with a writer's body and individual history to be of aesthetic interest, and proposed to focus on ›écriture‹ instead (cf. Barthes 1972).

<sup>10</sup> This view of style is perpetuated by introductions like the one by Cressot (1947), and methods of stylistic analysis such as the commutative analysis proposed by Dupriez (1971).



solution to the synonymy issue: synonymy in the context of style as choice can be redefined as synonymy of denotation only (the denotative or core meaning is stable across stylistic variants), while there can still be variation of connotation (the stylistic variation may induce a variation in secondary meaning). Especially from a linguistic point of view, this distinction is a sufficient way of reconciling the concept of »style as choice« with the idea that differences in form or style also change the meaning of a linguistic expression. However, proponents of literary stylistics, especially since the advent of semiotics and post-structuralism, have vehemently opposed this view, arguing for »the inseparability, the consubstantiality of thought and sign« (*l'indissolubilité, la consubstantialité de la pensée et du signe*, Gauthier 1970, 92): form and meaning cannot be separated at all, and any variation in form or style will create a fundamental variation in meaning.

Closely related to the first definition of style as choice, and in a way a further development of it, is the definition of style as a deviance from a norm. This is the view on style proposed, for example, by Bruneau (1951), who identifies the norm with the affordances and principles of grammaticality a given language offers, and defines style as the »écart«, i. e. the voluntary deviations from such a norm. »Stylistics is the science of deviations« he states programmatically (ibid., 6). More recent approaches to style have heavily criticized the notion of style as deviation from a norm, precisely because it is problematic what such a standard, neutral norm could be (cf. Guenier 1969). Bruneau himself, although he is attentive to the historic dimension of language, does not recognize this as a problem for the definition of the norm. For the purposes of a quantitative, empirical stylistics, the impossibility of defining and establishing one generally relevant norm makes implementing such approaches impractical, but it does not exclude using comparative or reference corpora and assessing relative deviations from them. Also, limiting stylistic phenomena to relatively rare, voluntary deviations from grammaticality leaves the largest part of literary and non-literary texts without style. This view is incompatible with computational and statistical approaches considering subtle differences in relative frequencies of many formal phenomena to be as much a sign of style as particular voluntary errors.

As a reaction to such criticism, an important more recent variant is the approach to style as differential deviation, where the deviation is not meant to be relative to a neutral standard or norm, but relative to a purposefully selected reference corpus which presents a specific relation of identity and difference to the texts under scrutiny: it may be texts written at the same time but by different authors, or by the same author but belonging to a different genre, etc. (see above). This variant of style as deviation has recently been employed, for example, by Pincemin (2012) in the context of »textométrie«, or the computational stylistic analysis of literary texts. Such a differential and comparative definition of style

avoids the problem of the supposedly neutral norm and has the advantage of being well-suited to quantitative, formalized approaches to style.

Not dissimilar to Staiger's »ineffable-identical« (see above), and still present, if marginally, in recent years, is the definition of style as the manifestation of an irreducible individuality. Jenny (1993) is an example of this position. He argues that literary style is in essence a »singularité«: this means that stylistics is concerned precisely with phenomena that eschew any attempt to categorize, class and analyze them, because style's absolute uniqueness makes it irreducibly to any category or class. Such a view of literary style is a fundamental challenge to any methodologically sound approach, because here, stylistic phenomena cannot be grouped and compared to each other or to similar phenomena in other texts. Jenny is very conscious of this issue, which is of course an even larger challenge to explicit formalization or modeling of stylistic features and operationalization of methodologies with algorithmic implementations. Despite being a less radical proponent of stylistics, Molinié (e. g. 1996) is equally interested in the stylistic specificities of literary texts as a specific type of discourse whose formal characteristics at various levels he aims to describe.

In parallel to all of the above, there is a long tradition of computational approaches to literary stylistics in France, going back to early quantitative analyses of the vocabulary of playwright Corneille by Muller (1964), followed for example by analyses of Racine's plays by Bernet (1983). These studies have been descriptive in nature, for the most part, outlining aspects like vocabulary richness or the prevalence of specific word classes over the career of a given author, and sometimes comparing one author to another. Jean-Paul Benzécri (1973) and others have introduced more complex statistical methods to the field, and pioneering work has been done in this area in numerous studies by Brunet since the 1970s (see Brunet 2009 and 2011). More recently, and making use of the availability of larger collections of text and ever more sophisticated algorithms, systems for the automatic identification of metric structures in verse (Beaudouin/Yvon 1996) and highly controversial studies attempting authorship re-attributions based on relative word frequencies for Corneille and Molière (e. g. Labbé/Labbé 2006, see also Schöch 2014) as well as individual stylistic studies (e. g. Magri-Mourgues 2011) have been published. The most recent avatar of computational analysis of literary style in France is called »textométrie« and, building on solid conceptual bases (e. g. Pincemin 2008, 2012), develops methods and tools (such as TXM) which are used by a growing community of practice.

## 5 Synthesis of Definitions

This exercise has shown why style has been seen as a thorny concept of language and literature studies. However, despite the multitude of definitions we have reviewed, there are actually only a few basic ways in which style has been approached across traditions. In the following, we take up the six types of definitions of style mentioned initially, and briefly describe where the three traditions of stylistics have converged and diverged in the past seventy years with respect to these schematic definitions.

- (1) With respect to our first category, that of style as a higher-order artistic value assessed through aesthetic experience, a clear tradition can be seen in assigning affective-aesthetic value to ›style‹. It appears especially in the German-speaking field (first Kayser, Staiger), with intensive immersion in the individual work of art (Szondi) as the basic mode of insight. The identification of style has here involved a form of aesthetic (value) judgment that explicitly includes affective, and hence subjective experience. As a rule, this conception of style involves evaluation on part of the scholar (on the basis of some idea of literary quality). Today, it can be encountered not only in journalistic literary criticism, but also in German *Literaturwissenschaft*, and is common in Anglo-Saxon literary criticism (cf. Crystal 2010, Wales 2001). Similar in Dutch and French studies are those approaches that incorporate impressionistic aspects in their method (e.g., Stutterheim). The affective-aesthetic approach has, however, been less common in French and Dutch stylistics, although Spitzer's view of explicitly incorporating subjective appreciation is widespread in France.
- (2) Our second category, that of style as the holistic gestalt of single texts, can be found in all three traditions. It was central to Werkimmanenz, and, in a more descriptive variant, is today a major pillar of pragmatic/textlinguistic approaches (which are pursued in the Anglo-Saxon context as well). It is, however, implicit in any approach that works with the literary text as a unit of analysis, assuming that there is some stylistic unity of the text or even, as in Leo Spitzer, a given author's entire oeuvre. Style has accordingly been defined as a text-constituting element (e.g., Fix 2008). One relatively marked example from French stylistics is Jenny's (1993) contention that literary style is in essence a ›singularité‹ that cannot be reduced to descriptive categories or classes.
- (3) The third category, style as an expression of individuality, subjectivity, or emotions of the author, captures an important aspect of style studies in the three traditions (and the Anglo-Saxon ones). Traditional approaches to style as expression of subjectivity strongly marked style studies between the 18th

and 20th Century (e.g., Spitzer). For example, Bally's stylistic analysis, applied to spontaneous spoken language, is entirely focused on style as an expression of affect or emotions. Similarly, Stutterheim describes the many ways in which language can be used to express emotion. In computational authorship attribution studies, and many quantitative approaches to style derived from it, the individuality of an author's style plays an important role (e. g., Brunet, Busemann, Jannidis/Lauer).

- (4) The fourth category, style as an artifact that presupposes (hypothetical or intentional) selection from among a set of alternatives, is quite heterogeneous. While an emphasis on ›selection‹ first of all reflects a structuralist assumption of a language system and its actualization in usage (Sanders), many, especially newer, selection theories have put a strong focus on pragmatic aspects of style (e. g., Fix, Michel et al., and Sandig in the German tradition; Bally, Dupriez in the French one). It is central in (stylistic) approaches in the Anglo-Saxon tradition, as for example reflected in Crystal's and Wales's definitions (see also Simpson 2004).
- (5) The fifth category, style as deviation, has been one of the strongest strands of stylistics across all three traditions (as in the Anglo-Saxon ones). However, as the ›selection‹ theories, this category subsumes very distinct approaches. They range from truly normative-prescriptive ideas of how to write (which are not the object of this article) and assumptions of static language norms (that are today untenable) to views of flexibly changing norms, or contrasts, of different kinds (e. g., Enkvist, Hellings/Van der Merwe Scholtz, Leech/Short, Riffaterre). The latter can be either defined in terms of social-communicative conventions, of cognitive (genre) expectations, or of statistical reference points (e. g., comparing literary corpora with each other or to a representative corpus of (non-)literary language, e. g., Brunet, Müller). A key concept from early structuralism that reentered stylistics through the Anglo-Saxon tradition is ›foregrounding‹ (cf. Wales 2001, 156–157).
- (6) Style as any linguistic feature that can be formally defined and measured computationally frequently implies a quantitative analysis of literature.<sup>11</sup> In this context, stylistic analysis can focus on a multitude of style markers (Stamatatos 2009) whose absolute or relative frequencies or values can be

---

**11** For the Anglo-Saxon context, Wales notes a basic definition of style at the level of text: »the set or sum of linguistic features that seem to be characteristic: whether of a register, genre, or period, etc.« (Wales 2001, 371), and holds that stylistic analysis does not necessarily involve computation. However, to many scholars, computation appears to be a logical next step (e. g., Crystal).

assessed. The list of style features is open, pending new features and measures to be developed in combined annotation and machine learning efforts. Frequency patterns can be compared within one text, across several texts, or relative to reference corpora (Herrmann, Jannidis/Lauer). The computational approach allows views of ›norms‹ as well as of style as a result of selection operations. Quantitative studies of formal style markers are practiced also in French (Brunet, Magri-Mourgues, Schöch) and Dutch studies (Kestemont, Van Dalen-Oskam).

## 6 A New Definition of Style

Bringing together what we have observed in this exercise, in order to ease the way for future research into style, we propose the following new definition:

*Style is a property of texts constituted by an ensemble of formal features which can be observed quantitatively or qualitatively.*

The following comments are intended to explicate key aspects of this definition and relate it to previous definitions reviewed above. By »text« we mean complete texts or fragments of texts; single such texts or collections of them (e. g. with a common author, genre, or time period); and finally, literary or non-literary texts.<sup>12</sup> In this respect, therefore, our definition differs from some previous definitions in that it does not make presuppositions as to the literary or non-literary status of the texts under scrutiny, and is not limited to a given author's style. By »ensemble« we mean that style is constituted by the combination of many possible features and should be seen as a complex system, with features situated at different linguistic levels. Such an ensemble does not necessarily exhibit a coherent unity; rather, it can have various degrees of unity or harmony, or, on the contrary, contrasts or incoherence. In this respect, our definition differs from more normative views of style which see coherent style as a sign of aesthetic quality. By »formal features«, we mean linguistic features at the level of characters, lexicon, syntax, semantics (Stamatatos 2009, 4), but also features going beyond the sentence, such as narrative perspective or textual macro-structure; we differ from some previous definitions in that we conceive of stylistic features as explicitly defined and clearly identifiable. Features relevant to style are not limited to deviations from grammaticality or some supposedly neutral norm.

---

<sup>12</sup> Textuality is determined by prototypical features such as cohesion, coherence, and situational factors (cf. Winko 2008).

Finally, by »quantitatively or qualitatively«, we mean that a certain style can be described using methods based on computing frequencies, relations, and distributions of features and relevant statistics, as well as methods based on precise observation and description of individual occurrences. In fact, most actual research is likely to practice a mixed method, in which the direction of research may vary: Qualitative observations may be confirmed by quantitative ones, after careful modeling and testing on a larger amount of material, but quantitative findings may as well be followed up by qualitative analyses of smaller samples, be they considered representative or atypical. In this respect, our definition differs from previous definitions in its attempt to include both perspectives, albeit in the context of formal stylistic features. Our definition is designed to work as a common ground for research in a new paradigm of style studies that is emerging from literary studies as much as from stylistics and computational linguistics, clarifying what can actually be meant when talking about style. In the following, we will contextualize the definition further, putting an emphasis on issues relevant to computational studies of style.

Our definition of style is broader and more abstract than most of the earlier definitions. Although the reported numbers of potential stylistic features and measures vary (some mentioning e. g., »nearly 1,000 different measures«, cf. Stamatatos 2009, 1), and despite their vast heterogeneity, most style markers have so far been relatively simple in nature. Among such style features are frequencies and frequency distributions of characters, words, lemmata, word classes or syntactical structures, taken by themselves or in sequences (n-grams); and the length, and distribution of lengths, of words, sentences, paragraphs or other units. Higher-order stylistic features are derived from basic style markers using various statistical techniques and measures: among the more commonly considered ones are vocabulary richness (e. g., type-token ratio or Yule's *K*) and various types of entropy scores.

Among the well-understood methods in quantitative stylistics are keyness measures, various distance measures (such as Burrows's Delta; Burrows 2002), principal component analysis, and multi-dimensional register analysis (cf. Biber 1988, Biber/Finegan 1994), allowing to assess the stylistic similarity of texts, as well as many more exploratory and classificatory techniques from the area of (supervised and unsupervised) machine learning. More recently, machine learning, and natural language processing (NLP) have advanced the development even further, with efficient techniques for representing and classifying large volumes of text and powerful machine learning algorithms to handle multi-dimensional and sparse data (e. g. Diederich et al. 2003 on Support Vector Machines), and more expressive visualizations (e. g. Schöch/Pielström 2014 on Principal Component Analysis).

In the light of recent advances, it seems both useful and possible to add the levels beyond the sentence among the style markers. In quantitative stylistics, this is new. While Stamatatos (2009) does not include this type of style markers in his overview, textlinguistic and discourse linguistic accounts have long focused on such features, as have literary studies (but almost solely from a qualitative perspective). Among these are, for example, proportions of text type (argumentative, narrative, descriptive), ratio of narrative vs. dialogue passages, or types of focalization. Recent work on automatically identifying reported speech (Brunner 2013) or descriptions (Jautze et al. 2013) using rule-based as well as machine-learning techniques, are examples of such new avenues for quantitative stylistic research. Despite such advances, we contend with stylisticians such as Leech and Short (2007) that one cannot really get a hold on all formal features that make up the style of some unit, but that one has to select certain features: for practical reasons, but also because of methodological limitations. This is true, for example, for stylistically relevant phenomena such as metaphors or irony which, despite some recent advances, are currently analyzed in a qualitative rather than quantitative paradigm and remain a challenge to formal modeling, reliable automatic identification and computational measurement.

In our approach, style is not something unique to literary works; rather, every text has a certain kind of style. However, the described ensemble of formal features may be interpreted from a literary perspective, taking into account the ensemble's internal relations which may be rich (or not) and complex (or not). Furthermore, our definition departs from the principle of necessity of contrast advocated by some earlier scholars: style can also be observed in and described for one text or one text sample without a necessary explicit comparison or contrast with other units of analysis (such a ›unit‹ being, for example, an oeuvre, book, chapter, but also genre, epoch, or author's gender). At the same time, the definition explicitly allows for such contrastive approaches.

In our definition, style can be associated with categories such as genre, epoch, author, and many more. In many cases, correlations between specific style markers or groups of style markers with these categories may be observed. What is more, even in the absence of conscious intentions, causal relationships may be hypothesized: genre can cause style (e. g. by means of conventions: form and themes), authors can cause style (e. g. by means of idiosyncrasies), theme and topic can cause style. The interpretability of style relative to categories such as authorship, literary genre, or literary period, is hence paramount. This means that any stylistic phenomenon can ultimately be considered the trace of or the index towards such categories (or, in other terms, may be ›characteristic‹ of them).

Our definition allows to determine the style of texts, authors, genres, periods etc. in terms of a quantitative profiling of formal features. Frequencies of selected



elements can be quantitatively related *across* a collection of individual texts, for example by using the collection itself as the relative norm, or using a large(r) reference corpus as a backdrop. Quantitative comparison is also possible *within* single texts, which allows for example distributional profiles of selected traits, as well as probabilistic assessments of co-occurrence of elements. Stylistic computations hence allow statements about style in some unit of analysis (e. g., text part, author, genre) in comparison with a wide range of reference data (other text parts, other authors, other genres). This can be related to the ›classic‹ structuralist and formalist approaches to literature, in particular to the notions of deviance/norm (foregrounding), and, of course, texts as semiotic systems of discrete elements. Ultimately, such an approach does not only facilitate text classification, but also stylistic description, and, if wished, inferences or assumptions about aesthetic effects and communicative intentions.

In this contribution, we have juxtaposed hermeneutic approaches with (qualitative and quantitative) empirical approaches. While the differences between the first one and the latter two are quite obvious, quite substantial divergences between the latter two may be stated as well: for example qualitative stylistics is interested in rare special phenomena and quantitative stylistics builds on subtle differences in high-frequency phenomena. However, they may be able to converge around a definition of style such as ours. It should be mentioned that qualitative literary stylistics have already taken a corpus stylistic turn, with e.g. Leech and Short (2007) advancing quantitative stylistics specifically for fiction. Here, frequency-based methods are applied, such as collocation and keyword analysis (cf. Biber 2011, Mahlberg 2013), but also more advanced statistical procedures such as multi-dimensional analysis in register variation (e. g., Biber 1988). These measures allow identifying non-coincidental, i. e. statistically significant, associations and frequency clusters (for a quantitative cross-register comparison of metaphor across word classes in English, see Steen et al. 2010), and they facilitate quantitative variation perspectives as well as qualitative ones.

Finally, the tentative convergence of linguistic stylistics, literary stylistics and computational stylistics also raises the question of the relation between evidence and interpretation, and calls for a careful assessment of the interpretive and evidence-based aspects of computational analyses of style. One key principle of corpus stylistics is the assumption of an equivalence between frequency and significance in language data (cf. Sinclair 1991):<sup>13</sup> Scholars such as Fischer-

---

<sup>13</sup> But see Leech and Short (2007, 57), who state that »there is no direct relation between statistical deviance and stylistic significance«, contending that »literary considerations must guide us in selecting what features to examine«.

Starcke (2009) argue that the »fact that the patterns are objective features of the text gives their subjective interpretation a firm basis« (ibid., 495; see also Wales 2001, 373 for the debate between literary criticism and ›objective‹ approaches in the Anglo-Saxon context). Style manifests itself in lexico-grammatical and other textual features that have been established by rhetoric and linguistic scholarship as meaningful and functional units of language use and structure, and that have a basis in the linguistic processing behavior by actual interlocutors. Seen from this perspective, style could be termed an ›objective‹ quality of texts. However, the way researchers use such evidence in order to make more general arguments about literary texts inevitably involves interpretation. As Jannidis and Lauer put it with regard to their stylometric analysis of German literary history, the »interpretations of the results of quantitative studies [...] are hermeneutic acts of sense making« (2014, 50). Inversely, a style's distinctive characteristic may in principle be captured by descriptive labels, such as »poetic«, »colloquial« or »aphoristic«. However, a careful definition of the relationship between formal style markers and descriptive categories is paramount, as is providing evidence for the distinctiveness of such style markers with regard to a scrutinized text or text collection. This challenge of striving for a high level of intersubjectivity and reproducibility while at the same time proposing meaningful and enriching perspectives on literary texts has frequently been noted, be it by stylisticians such as Simpson (2004), or computational stylisticians and digital humanists (e. g., Craig 1999, Ramsay 2011). The first aspect demands a clear conceptualization of style as well as valid operationalizations that allow reliable identification of stylistic features; the latter aspect requires bold synthesis and minute contextualization of results. We hope that our definition of style can be the basis for an encounter of linguistics, literary studies and computer science as much as an encounter of evidence and interpretation – offering a fruitful avenue for computational investigations into the rich nature and history of style in literary texts.

## References<sup>14</sup>

- Anbeek, Ton/Arie Verhagen, Over stijl, *Neerlandistiek.nl* 1:1 (2001), <http://www.neerlandistiek.nl/01.01/>.
- Anderegg, Johannes, Literaturwissenschaftliche Stilauffassungen, in: Ulla Fix/Andreas Gardt/Joachim Knappe (ed.), *Rhetoric and Stylistics. An International Handbook of Historical and Systematic Research*, Vol. 1, Berlin/New York 2008, 1076–1092.

---

<sup>14</sup> All online references were last consulted on July 30, 2014.

- Arrivé, Michel, Postulats pour la description linguistique des textes littéraires, *Langue française* 3:1 (1969), 3–13.
- Bally, Charles, *Traité de stylistique française* [1909], Heidelberg <sup>2</sup>1921.
- Barthes, Roland, *Le Degré zéro de l'écriture* [1953], Paris 1972.
- Barthes, Roland, Éléments de sémiologie, *Communications* 4 (1964), 91–135.
- Beaudouin, Valérie/Francois Yvon, The Metrometer: A Tool for Analysing French Verse, *Literary and Linguistic Computing* 11:1 (1996), 23–31.
- Benzécri, Jean-Paul, *L'analyse des données*, Paris 1973.
- Bernet, Charles, *Le Vocabulaire des tragédies de Racine. Analyse statistique*, Geneva/Paris 1983.
- Biber, Douglas, *Variation across Speech and Writing*, Cambridge 1988.
- Biber, Douglas, Corpus Linguistics and the Study of Literature: Back to the Future?, *Scientific Study of Literature* 1:1 (2011), 15–23.
- Biber, Douglas/Edward Finegan, *Multi-Dimensional Analyses of Authors' Styles: Some Case Studies from the Eighteenth Century*, Oxford 1994.
- Bruneau, Charles, La stylistique, *Romance Philology* 5:1 (1951), 1–14.
- Brunet, Etienne, *Comptes d'auteurs, tome I: Études statistiques, de Rabelais à Gracq*, ed. Damon Mayaffre, Paris 2009.
- Brunet, Etienne, *Ce qui compte: Écrits choisis, tome II: Méthodes statistiques*, ed. Céline Poudat, Paris 2011.
- Brunner, Annelen, Automatic Recognition of Speech, Thought, and Writing Representation in German Narrative Texts, *Literary and Linguistic Computing* 28:4 (2013), 563–575.
- Burrows, John, ›Delta‹: A Measure of Stylistic Difference and a Guide to Likely Authorship, *Literary and Linguistic Computing* 17:3 (2002), 267–287.
- Busemann, Adolf, *Stil und Charakter. Untersuchungen zur Psychologie der individuellen Redeform*, Meisenheim 1948.
- Compagnon, Antoine, Style, in: A.C., *Le démon de la théorie: littérature et sens commun*, Paris 1998, 179–208.
- Craig, Hugh, Authorial Attribution and Computational Stylistics: If You Can Tell Authors Apart, Have You Learned Anything about Them?, *Literary and Linguistic Computing* 14:1 (1999), 103–113.
- Cressot, Marcel, *Le Style et ses techniques. Précis d'analyse stylistique*, Paris 1947.
- Crystal, David, Stylistic Identity and literature, in: *The Cambridge Encyclopedia of Language* [1987], Cambridge <sup>3</sup>2010, 68–81.
- Diederich, Joachim/Jörg Kindermann/Edda Leopold/Gerhard Paass, Authorship Attribution with Support Vector Machines, *Applied Intelligence* 19:1–2 (2003), 109–123.
- Dupriez, Bernard, Une stylistique structurale est-elle possible?, *Le Français moderne* 4 (1971), 336–344.
- Eibl, Karl, Ist Literaturwissenschaft als Erfahrungswissenschaft möglich? Mit einigen Anmerkungen zur Wissenschaftsphilosophie des Wiener Kreises, in: Philip Ajouri/Katja Mellmann/Christoph Rauen (ed.), *Empirie in der Literaturwissenschaft*, Münster 2013, 19–45.
- Enkvist, Nils Erik, *Linguistic Stylistics*, The Hague 1973.
- Fagel, Suzanne, Structure versus Style: The Disappearance of Stylistics in the Netherlands 1940–1970, *Online Proceedings of the Annual Conference of the Poetics and Linguistics Association (PALA)*, 2008, <http://www.pala.ac.uk/resources/proceedings/2008/fagel2008.pdf>.
- Fagel, Suzanne, The art of ordinary words and sentences: sentence complexity and madness in works by the Dutch authors Maarten Biesheuvel and Jan Arends, *Online proceedings of the Annual Conference of the Poetics and Linguistics Association (PALA)*, 2009, <http://www.pala.ac.uk/resources/proceedings/2009/fagel2009.pdf>.

- Fagel, Suzanne, The stylistic creation of an antihero: tense and aspect in Arnon Grunberg's ›The Asylum Seeker‹, *Online proceedings of the Annual Conference of the Poetics and Linguistics Association (PALA)*, 2010, <http://www.pala.ac.uk/resources/proceedings/2010/fagel2010.pdf>.
- Fagel, Suzanne/Ninke Stukker/Loes van Andel, Hoe telbaar is stijl?: een kwantitatieve analyse van observatie en participatie in de stijl van Arnon Grunberg, *Nederlandse Letterkunde* 17:3 (2012), 178–203.
- Fagel-de Werd, Suzanne, *De stijl van gewoon proza*, Leiden 2015, <http://hdl.handle.net/1887/31606>.
- Fischer-Starcke, Bettina, Keywords and Frequent Phrases of Jane Austen's *Pride and Prejudice* A Corpus-Stylistic Analysis, *International Journal of Corpus Linguistics* 14:4 (2009), 492–523.
- Fix, Ulla, *Stil – ein sprachliches und soziales Phänomen. Beiträge zur Stilistik*, Berlin 2007.
- Fix, Ulla, Text und Textlinguistik, in: Nina Janich (ed.), *Textlinguistik. 15 Einführungen*, Tübingen 2008, 15–34.
- Fix, Ulla, Literaturwissenschaft und Linguistik – Das Konzept »LiLi« aus heutiger linguistischer Sicht, *Journal of Literary Theory* 4:1 (2010), 19–40.
- Fix, Ulla/Andreas Gardt/Joachim Knappe (ed.), *Rhetoric and Stylistics. An International Handbook of Historical and Systematic Research*, 2 Vol., Berlin/New York 2008/2009.
- Fucks, Wilhelm, *Mathematische Analyse von Sprachelementen, Sprachstil und Sprachen*, Köln 1955.
- Gauthier, Michel, À propos de stylistiques, *Langue française* 8 (1970), 90–97.
- Graubner, Hans, Stilistik, in: Heinz Ludwig Arnold/Volker Sinemus (ed.), *Grundzüge der Literatur- und Sprachwissenschaft*, Vol. 1, München 1973, 164–186.
- Groeben, Norbert, Was kann/soll Empirisierung (in) der Literaturwissenschaft heißen?, in: Philip Ajouri/Katja Mellmann/Christoph Rauen (ed.), *Empirie in der Literaturwissenschaft*, Münster 2013, 47–74.
- Guenier, Nicole, La pertinence de la notion d'écart en stylistique, *Langue Française* 3 (1969), 34–45.
- Guiraud, Pierre, *La stylistique*, Paris 1963.
- Hellings, W.G./H. van der Merwe Scholtz, *Kreatiewe analise van taalgebruik*, Amsterdam/Pretoria 1995.
- Herrmann, J. Berenike, Computing Kafka – How Keyness and Collocation Analysis Help Explain Paradoxical Style, talk and poster presented at International Herrenhausen Conferene (*Digital Humanities Revisited*, Hanover, Dec. 2013).
- Jakobson, Roman/Claude Lévi-Strauss, »Les Chats« de Charles Baudelaire, *L'Homme* 2:1 (1962), 5–21.
- Jannidis, Fotis/Gerhard Lauer, Burrows's Delta and Its Use in German Literary History, in: Matt Erlin/Lynne Tatlock (ed.), *Distant Readings. Topologies of German Culture in the Long Nineteenth Century*, Rochester 2014, 29–54.
- Jautze, Kim/Corina Koolen/Andreas van Cranenburgh/Hayco de Jong, From high heels to weed attics: a syntactic investigation of chick lit and literature, *Proceedings of the Second Workshop on Computational Linguistics for Literature*, 2013, <http://www.aclweb.org/anthology/W13-1410>.
- Jenny, Laurent, L'objet singulier de la stylistique, *Littérature* 89:1 (1993), 113–124.
- Kayser, Wolfgang, *Das sprachliche Kunstwerk. Eine Einführung in die Literaturwissenschaft*, Bern 1948.
- Kestemont, Mike, What Can Stylometry Learn from Its Application to Middle Dutch Literature?, *Journal of Dutch Literature* 2:2 (2011), 46–65.

- Kreuzer, Helmut, *Mathematik und Dichtung. Zur Einführung*, in: H.K./Rul Gunzenhäuser (ed.), *Mathematik und Dichtung. Versuche zur Frage einer exakten Literaturwissenschaft*, München <sup>4</sup>1965, 9–20.
- Kreuzer, Helmut/Rul Gunzenhäuser (ed.), *Mathematik und Dichtung: Versuche zur Frage einer exakten Literaturwissenschaft*, München <sup>4</sup>1965.
- Labbé, Cyril/Dominique Labbé, A Tool for Literary Studies: Intertextual Distance and Tree Classification, *Literary and Linguistic Computing* 21:3 (2006), 311–326.
- Lambrou, Marina/Peter Stockwell (ed.), *Contemporary Stylistics*, London 2007.
- Larthomas, Pierre, *Notions de stylistique générale*, Paris 1998.
- Leech, Geoffrey N./Mick Short, *Style in Fiction. A Linguistic Introduction to English Fictional Prose* [1981], New York <sup>2</sup>2007.
- Magri-Mourgues, Véronique, Analyse textométrique et interprétation littéraire – Hyperbase, Rousseau et les Lumières, *Travaux neuchâtelois de linguistique* 55 (2011), 77–93.
- Mahlberg, Michaela, Corpus Analysis of Literary Texts, in: Carol A. Chapelle (ed.), *The Encyclopedia of Applied Linguistics*, London 2013, doi:10.1002/9781405198431.wbeal0249.
- Michel, Georg/Günter Starke/Franz Graehn, *Einführung in die Methodik der Stiluntersuchung. Ein Lehr- und Übungsbuch für Studierende*, Berlin 1968.
- Molinié, Georges, Le champ stylistique, *L'Information Grammaticale* 70:1 (1996), 21–24.
- Müller, Ralph, *Die Metapher. Kognition, Korpusstilistik und Kreativität*, Paderborn 2012.
- Muller, Charles, *Étude de statistique lexicale: le vocabulaire du théâtre de Pierre Corneille*, Paris 1967.
- Nørgaard, Nina/Beatrix Busse/Rocío Montoro, *Key Terms in Stylistics*, London 2010.
- Pieper, Ursula, Statistische Literaturanalyse, in: Jan-Dirk Müller et al. (ed.), *Reallexikon der deutschen Literaturwissenschaft*, Vol. 3, Berlin/New York 2003, 500–503.
- Pincemin, Bénédicte, Modélisation textométrique des textes, in: Serge Heiden/Bénédicte Pincemin (ed.), *Actes des 9es Journées internationales d'Analyse statistique des Données Textuelles*, Tome 2, Lyon 2008, 949–960.
- Pincemin, Bénédicte, Sémantique interprétative et textométrie, *Texto!* 17:3 (2012), <http://www.revue-texto.net/index.php?id=3049>.
- Plett, Heinrich F., *Einführung in die rhetorische Textanalyse* [1971], Hamburg <sup>9</sup>2001.
- Püschel, Ulrich, Stilistik der deutschsprachigen Länder vom Beginn des 20. Jahrhunderts bis zur Gegenwart, in: Ulla Fix/Andreas Gardt/Joachim Knappe (ed.), *Rhetoric and Stylistics. An International Handbook of Historical and Systematic Research*, Vol. 1, Berlin/New York 2008, 165–179.
- Ramsay, Stephen, *Reading Machines: Toward an Algorithmic Criticism*, Urbana 2011.
- Riesel, Elise, *Stilistik der deutschen Sprache*, Moskau 1959.
- Riffaterre, Michael, *Essais de stylistique structurale*, Paris 1971.
- Riffaterre, Michael, *Strukturelle Stilistik*, trans. Wilhelm Bolle, München 1973.
- Sanders, Willy, *Linguistische Stiltheorie. Probleme, Prinzipien und moderne Perspektiven des Sprachstils*, Göttingen 1973.
- Sandig, Barbara, *Stilistik der deutschen Sprache*, Berlin/New York 1986.
- Sandig, Barbara, *Textstilistik des Deutschen*, Berlin/New York <sup>2</sup>2006.
- Schmidt, Siegfried J., »Bekämpfen Sie das häßliche Laster der Interpretation! Bekämpfen Sie das noch häßlichere Laster der richtigen Interpretation!«, in: Wolfgang Frier/Gerd Labrousse (Hg.), *Grundfragen der Textwissenschaft. Linguistische und literaturwissenschaftliche Aspekte*, Amsterdam 1979, 279–309.

- Schöch, Christof, Corneille, Molière et les autres. Stilometrische Analysen zu Autorschaft und Gattungszugehörigkeit im französischen Theater der Klassik, in: Christof Schöch/Lars Schneider (ed.), *Literaturwissenschaft im digitalen Medienwandel* (Beihefte zu Philologie im Netz), 2014, 130–157, <http://web.fu-berlin.de/phn/beiheft7/b7t08.pdf>.
- Schöch, Christof/Steffen Pielström, Für eine computergestützte literarische Gattungsstilistik, *Jahrestagung der Digital Humanities im deutschsprachigen Raum*, 2014, [http://dig-hum.de/sites/dig-hum.de/files/Schoch-Pielstrom\\_2014\\_Gattungsstilistik.pdf](http://dig-hum.de/sites/dig-hum.de/files/Schoch-Pielstrom_2014_Gattungsstilistik.pdf).
- Simpson, Paul, *Stylistics. A Resource Book for Students*, London 2004.
- Sinclair, John M., *Corpus, Concordance, Collocation: Describing English Language*, Oxford 1991.
- Sowinski, Bernhard, Stil, in: Gert Ueding/Gregor Kalivoda (ed.), *Historisches Wörterbuch der Rhetorik*, Vol. 8, Darmstadt 2007, 1393–1419.
- Spillner, Bernd, *Linguistik und Literaturwissenschaft. Stilforschung, Rhetorik, Textlinguistik*, Stuttgart 1974.
- Spitzer, Leo, *Stilstudien. Erster Teil: Sprachstile. Zweiter Teil: Stilsprachen* [1928], München 1961.
- Spitzer, Leo, 1988. *Études de style*, Paris 1988.
- Staiger, Emil, *Die Kunst der Interpretation. Studien zur deutschen Literaturgeschichte*, Zürich 1955.
- Stamatatos, Efstathios, A Survey of Modern Authorship Attribution Methods, *Journal of the American Society for Information Science and Technology* 60:3 (2009), 538–556.
- Steen, Gerard J./Aletta G. Dorst/J. Berenike Herrmann/Anna A. Kaal/T. Krennmayr, Metaphor in Usage, *Cognitive Linguistics* 21:4 (2010), 765–796.
- Stutterheim, C.F.P., *Stijllee*, Den Haag 1947.
- Szondi, Peter, Über philologische Erkenntnis, in: P.C., *Hölderlin-Studien*, Frankfurt a. M. 1967, 9–30.
- Van Bork, G.J./D. Delabatista/H. van Gorp/P.J. Verkruijsse/G.J. Vis, in: G.J. van Bork, D. Delabatista, H. van Gorp, P.J. Verkruijsse, and G.J. Vis (ed.), *Algemeen letterkundig lexicon*, Dbnl, 2012, [http://www.dbnl.org/tekst/dela012alge01\\_01/dela012alge01\\_01\\_03175.php](http://www.dbnl.org/tekst/dela012alge01_01/dela012alge01_01_03175.php).
- Van Dalen-Oskam, Karina, De stijl van R. Inaugural lecture, Amsterdam 2012, [http://www.oratiereeks.nl/upload/pdf/PDF-1290Weboratie\\_Van\\_Dalen-Oskam.pdf](http://www.oratiereeks.nl/upload/pdf/PDF-1290Weboratie_Van_Dalen-Oskam.pdf).
- Van Dalen-Oskam, Karina/Joris van Zundert, Delta for Middle Dutch: Author and Copyist Distinction in ›Walewein‹, *Literary and Linguistic Computing* 22:3 (2007), 345–362.
- Van Driel, Joost, *Prikkeling der zinnen. De stilistische diversiteit van de Middelnederlandse epische poëzie*, Zutphen 2007.
- Wales, Katie, *A Dictionary of Stylistics*, Harlow 2001.
- Winko, Simone, Textualitätsannahmen und die Analyse literarischer Texte, *Zeitschrift für Germanistische Linguistik* 36 (2008), 427–443.