Mapping the aesthetic space of literature "from below"

Christine A. Knoop[a,*,1], Valentin Wagner[a,1], Thomas Jacobsen[b], Winfried Menninghaus[a]

[a] Max Planck Institute for Empirical Aesthetics, Grüneburgweg 14, 60322 Frankfurt, Germany
[b] Helmut Schmidt University, University of the Federal Armed Forces Hamburg, Allgemeine und Biologische Psychologie, Holstenhofweg 85, 22043 Hamburg, Germany

ARTICLE INFO

ABSTRACT

The present paper aims to elucidate the conceptual structure of the aesthetics of literature. Following Fechner's "aesthetics from below" (1876) and adopting a method introduced by Jacobsen, Buchta, Kohler, and Schroeger (2004), we asked 1544 German-speaking research participants to list adjectives that they use to label aesthetic dimensions of literature in general and of individual literary forms and genres in particular (novels, short stories, poems, plays, comedies). According to our analyses of frequency, mean list rank, and the Cognitive Salience Index, *beautiful* and *suspenseful* rank highest across all target categories. For plays/comedies, *funny* and *sad* turned out to be the most relevant terms; for novels and short stories, *suspenseful,interesting* and *romantic;* and for poetry *romantic*, along with the music-related terms *harmonious*, *rhythmic*, and *melodious*. A comparison of our results with analogous studies for visual aesthetics and music yielded a comprehensive map of the distribution of aesthetic appeal dimensions across sensory modalities and aesthetic domains, with poetry and music showing the greatest overlap.

© 2016 The Authors. Published by Elsevier B.V. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

1. Introduction

Reading literature, just like viewing visual artworks (Leder, Belke, Oeberst, & Augustin, 2004) or listening to music (Brattico, Bogert, & Jacobsen, 2013), involves perceptual, cognitive, affective and evaluative processing dimensions (Jacobs, 2015). Aesthetic appreciation integrates all these processing dimensions with a focus on evaluating the perceived aesthetic appeal of a work of literature, and the felt hedonic reward of the exposure (cf. Kant's theory of aesthetic "judgment" and "liking"; Kant, 2007). When communicating the perceived aesthetic appeal of a work of literature to others, we project our experience onto the space of available verbal concepts and select those we find most appropriate to capture the nature of our impression. The resulting attributions of particular aesthetic appeal dimensions to particular works or genres of literature reflect the conceptual representations (1) of our aesthetic appreciation of these works or (2) of the expectations we have of certain genres. These attributions, i.e. the selective use of specific verbal terms, allow for straightforward empirical investigation. They also represent valid options to be considered when selecting rating scale items for empirical research on literature.

Dating back as far as the 18th century, literary history records a number of attempts to establish rating scales for evaluating the perceived aesthetic appeal of the work of individual authors (see Spoerhase, 2014, for a review). These historical efforts have a number of analogs in recent literary scholarship that attempt, in light of historical, psychological and

---

* Corresponding author. Fax: +49 69 8300479 599.
  E-mail address: christine.knoop@aesthetics.mpg.de (C.A. Knoop).
[1] Christine A. Knoop and Valentin Wagner are joint first authors.

sociological findings and based on theoretical reflection, to establish models of aesthetic appreciation of literature; such models seek to identify dimensions of evaluation that are either characteristic for individual works of art or that recur in the history of literary reception (cf., e.g., von Heydebrand & Winko, 1996), sometimes with a specific focus on individual genres (cf., e.g., Ribeiro, 2012). In the empirical and experimental research on art reception, a number of scales have been developed to capture specific aspects of aesthetic experience in reading literature (see, for instance, Green & Brock, 2000; Koopman, 2013, 2015; Kuijpers, Hakemulder, Tan, & Doicaru, 2014; Kuiken, Campbell, & Sopčák, 2012; Miall & Kuiken, 1995). Nevertheless, empirical studies of literature have not yet systematically investigated the verbal concepts readers use for communicating their aesthetic perceptions and evaluations of literature and the range and distribution of these concepts as dependent on specific genres of literature.

For an aesthetics "from below" as proposed by Fechner (1876), it is crucial to collect data on how readers themselves verbally represent their aesthetic experiences and expectations rather than to rely on expert evaluations and experimenter-selected rating scales only. Importantly, however, these different approaches are by no means mutually exclusive. In fact, previous research in music reception (Zentner, Grandjean, & Scherer, 2008) has shown that a combined bottom-up and top-down approach was best suited to capture musical experiences. Similarly, a recent study on films revealed that participants faced with experimenter-selected scales gave consistently high ratings for items which only few of them had mentioned unprompted when asked to verbalize their responses in their own terms (Wassiliwizky, Wagner, Jacobsen, & Menninghaus, 2015). Thus, a selection of dimensions of aesthetic appreciation that exclusively relies on a bottom-up retrieval process may be subject to serious limitations. However, collecting spontaneous bottom-up conceptualizations of the aesthetics of literature does have potential not only for selecting items for rating scales in future studies, but also for providing a deeper comparative understanding of the individual domains of aesthetics through the lens of their respective conceptualizations.

Similar to other art domains, aesthetic perceptions and evaluations of works of literature tend to be shaped by, and entail, comparisons to past experiences with particular literary forms or genres (see, e.g., Picon, 1953; Schmidt, 2007; von Heydebrand & Winko, 1996; Weninger, 1994; Zwaan, 1994). In turn, the expectations derived from these comparisons influence the choice of terms with which readers verbally describe their aesthetic perceptions and evaluations of literature. The same holds for expectations derived from knowledge readers have acquired academically or via paratext or excerpts (cf. Dixon, Bortolussi & Sopcak, 2015). Based on these assumptions, the goal of the present paper is to elucidate the conceptual structure of the aesthetics of literature. This effort has the potential to reveal both form- and genre-specific variations in the relevant vocabulary and interdependencies and similarities between genre-specific expressions. Our approach is well established as a language-based way of collecting associations to gain insight into the mental representation of conceptual structure (cf. Fehr & Russell, 1984; Kuehnast, Wagner, Wassiliwizky, Jacobsen, & Menninghaus, 2014). Moreover, by comparing our data with those from studies conducted in other aesthetic domains (Augustin, Wagemans et al., 2012; Istók et al., 2009; Jacobsen et al., 2004), we wish (1) to determine terms that are used specifically for literature and (2) to identify overlaps between aesthetic appeal dimensions mentioned for certain literary genres and other domains, including fashion, architecture, design, physical attractiveness, and so on (see Augustin, Carbon et al., 2012, for a similar approach comparing visual arts, film, and music). This approach allows for a comparative mapping of the domain- and genre-specific aesthetic expectations of non-experts across the broader field of aesthetic appreciation.

2. Capturing aesthetic perceptions and evaluations

Jacobsen et al. (2004) successfully used a free listing task to collect terms used for designating aesthetically relevant dimensions of objects at large. Using the same methodology, Istók et al. (2009) conducted a study on music, while Augustin, Wagemans et al. (2012) collected terms describing aesthetic appeal dimensions regarding eight different groups of visual objects (buildings, cars, clothing, faces, interior designs, landscapes, geometric shapes and patterns, and visual art).

Our study adopts the bottom-up, exploratory "aesthetics from below" approach used in these three preceding studies on other aesthetic domains. This places us in a position not only to add the missing data regarding literature, but also to be the first to compare the available data for the different aesthetic domains. Literature is traditionally comprised of three major forms: poetry, prose, and drama (cf. Hegel, 1975). We included all of these forms in our study, yet treated them differently for a variety of reasons. Our general subsample for plays encompassed the dramatic form as a whole, but the data were further specified by the inclusion of a comedies-only subsample. For prose, we collected subsample data for two of the most popular prose genres: novels and short stories. We did not ask for dimensions of aesthetic appeal associated with prose as a general category because "prose" also encompasses scientific and essayistic writing, while the interest of our study is limited to the more narrowly defined field of literature. In contrast, we included poetry exclusively as a broad category, without further generic subdistinctions. We did so anticipating that (1) unlike prose, "poetry" is a fairly meaningful category even for non-readers of poetic genres, and (2) terms designating special genres of poetry (such as the elegy, the ode, or the hymn) may seem too specific to the broader public. Regarding plays, we intended to collect data for the all-inclusive category "play/drama", which includes the polar genres "tragedy" and "comedy"; however, due to a communication error between the teams who collected the data, one team gathered data for the genre of comedy only. (Fortunately, the data we did obtain at least partly compensated for this error, since the difference between the entries for the general category "play/drama" and those for the subgenre "comedy" largely appear to reflect the role of tragedy. See the discussion section for details.)

3. Terminological mapping of literature

3.1. Method

3.1.1. Participants

One thousand five hundred and forty-four students of different academic disciplines participated in our study (994 women and 542 men, 8 undisclosed); the mean age was 23.5 years (SD = 7.28, min = 17, max = 84). Students were chosen as respondents for the pragmatic reason of availability, but also to keep the design as comparable as possible to the studies conducted in other aesthetic domains (Augustin, Wagemans et al., 2012; Istók et al., 2009; Jacobsen et al., 2004), which were also based on student samples. Of the participants, 875 were tested in Hamburg and 669 in Berlin, with the subsample sizes varying between 134 and 423 participants.[2]

3.1.2. Procedure

The data were collected in several lecture classes of different academic disciplines, including biology, cultural anthropology, economics, law, pedagogy, history, linguistics, literary studies, and psychology, to add variance to the samples and reduce group-specific effects.[3] To obtain the aesthetic perception and evaluation terms used for different forms and genres of literature, we collected data for the six subsamples Literature (in general), Poems, Novels, Short stories, Plays (in general), and Comedies. The study was conducted as a paper-and-pencil survey; the participating students received the following instruction: "Please write down terms that could be used to describe the aesthetics of literature. Please use adjectives only. You now have 2 minutes." Depending on the subsample, the word "literature" was replaced by "poems", "novels", "short stories", "plays (comedies or tragedies)" or "comedies".

The wording of the instruction was chosen for the following reasons. First, we wanted to avoid the intricacies of coding and evaluating qualitative data of a free response format and profit from the straightforward quantitative analysis of word frequency and list rank. Therefore the instruction ruled out full-sentence descriptions of participants' conceptions regarding aesthetically relevant features of the respective genres. At the same time, the instruction involved the notion that full-length accounts of aesthetic expectations are likely to include multiple dimensions rather than just one. Therefore, while extracting the terms participants consider relevant for designating aesthetic appeal dimensions does not amount to any concrete description of an aesthetic experience, it provides us with a list of words to be used as part of such descriptions. Moreover, the wording of our instruction was kept as similar as possible to the instruction used in the studies on other aesthetic domains (Augustin, Wagemans et al., 2012; Istók et al., 2009; Jacobsen et al., 2004) to ensure that we would be able to compare our data with theirs in a meaningful way.[4]

On top of collecting the word lists we logged the participants' gender, age, and academic discipline, and asked them to indicate whether they considered themselves experts on literature. All participants remained entirely anonymous. The data were collected in German; the results displayed in the figures and tables in this article are translations of the original material (see Appendix A, Table A1).

3.2. Results

3.2.1. Sample statistics

We processed and analyzed the data in several steps. In the first step, we corrected spelling mistakes and computed the absolute and mean numbers of entries. Our participants generated 9669 answers in total, corresponding to 2131 different words. The total figure also includes illegible entries and terms listed twice by the same participant. The number of answers per participant ranged between 1 and 21 ($M = 6.26$; SD = 3.76). Regarding the number of entries, an effect of the specificity of the reference categories became evident ($F(5,1538) = 11.49$, $p < .001$): the mean number of entries for literature in general (4.95) was smaller than those for the more specific reference categories, i.e., novels (7.18), poems (6.50), plays (6.71), and comedies (6.79). Moreover, the number of mean entries for short stories (5.78) was smaller than for novels and comedies (these post-hoc comparisons were based on Tukey's "Honestly Significant Difference" method).[5]

---

[2] The study was conducted in Berlin and Hamburg because the participating researchers had their respective institutional affiliations in these two cities, and thus could enlist the help of both research assistants and lecturers willing to let us conduct our study in their classes.

[3] To secure variety in our samples, we preselected classes from different research fields (natural sciences, social sciences, humanities) and, within these fields, different disciplines. The final selection of classes was contingent upon the agreement of the respective lecturers to allocate time for our study during class.

[4] From a theoretical point of view, it would have been more precise to ask for terms designating "perceived aesthetic appeal dimensions" rather than just "aesthetics". However, we chose the simpler option because we anticipated that this proxy would be easier to understand for our participants.

[5] Differences of the 15 comparisons with 95%CI and adjusted p-values: poems-literature = 1.55 [0.74; 2.37], $p < .001$; novels-literature = 2.23 [1.12; 3.34], $p < .001$; short stories-literature = 0.83 [−0.12; 1.78], $p = .130$; plays-literature = 1.76 [0.73; 2.79], $p < .001$; comedies-literature = 1.84 [0.97; 2.71], $p < .001$; novels-poems = 0.68 [−0.37; 1.72], $p = .437$; short stories-poems = −0.73 [−1.60; 0.15], $p = .164$; plays-poems = 0.21 [−0.75; 1.16], $p = .990$; comedies-poems = 0.28 [−0.50; 1.07], $p = .908$; short stories-novels = −1.40 [−2.56; −0.25], $p = .007$; plays-novels = −0.47 [−1.68; 0.75], $p = .880$; comedies-novels = −0.39 [−1.48; 0.69], $p = .907$; plays-short stories = 0.93 [−0.14; 2.00], $p = .128$; comedies-short stories = 1.01 [0.09; 1.93], $p = .022$; comedies-plays = 0.08 [−0.92; 1.08], $p = 1.0$.

3.2.2. Results for the individual subsamples

In the second step, terms mentioned by fewer than 5% of the participants in the individual subsamples were excluded from the analyses (for a similar cut-off procedure, see Istók et al., 2009; Jacobsen et al., 2004; van Goozen & Frijda, 1993). This procedure reduces variability that might reflect idiosyncratic uses. For the remaining 52 terms (out of the original 2131) we computed relative frequency (i.e., the number of participants mentioning a given term divided by subsample size), mean list rank (based on the raw lists, before excluding terms), and the Cognitive Salience Index (CSI; Sutrop, 2001) for each individual subsample. The CSI is the quotient of relative frequency and mean list rank and is bounded between 0 and 1, with higher values reflecting more salient terms. The results for the individual subsamples are depicted in Fig. S1, with the terms ranked for each subsample.

Overall, *beautiful* and *suspenseful* turned out to be listed most frequently (423 and 269; 27.4% and 17.4%, respectively). *Beautiful* and *boring* are the only terms mentioned in each sample by at least 5%; *boring* also turned out to be the only term of unambiguously negative valence. The terms *suspenseful* and *interesting* rank particularly highly in the narrative (plot-based) genres; these include dramatic plot, in accordance with Aristotle's concept of *mythos* (Aristotle, 2005) as a variant of narrative plot.

3.2.3. Comparing the subsamples

To further examine the commonalities and differences between the subsamples, we preprocessed the data slightly differently and retained in all subsamples all entries that were mentioned by at least 10% of the participants in one subsample. This leaves 22 terms, as opposed to the 52 terms for the 5% cut-off. The new frequency patterns were examined by cross-tabulation (using the χ^2-test) to test for differences between the samples. For all 15 comparisons of the frequency patterns, the χ^2-test revealed significant differences between the subsamples ($p < .001$, Bonferroni-corrected for multiple comparisons). In the next step, we calculated the overlapping coefficient (OVL; Inman & Bradley, 1989; Marx, 1976a, 1976b) for each pair of subsamples by adding up the lower relative frequencies in the two distributions for each of the mentioned words. The OVL ranged between .32 and .66, with the smallest overlap between poems and comedies and the largest between novels and short stories (see Table 1). We further analyzed the OVL matrix using classical multidimensional scaling (MDS; also called principal coordinate analysis; Gower, 1966) and hierarchical cluster analysis (HCA, employing Ward's criterion). Thus, we identified a pattern of clusters in both a cluster analysis and an MDS based on the OVL matrix (see Figs. S2A and B), with literature in general and poetry in particular showing vast similarities: they are both frequently associated with terms such as *poetic*, *harmonic*, *rhythmical*, and *harmonious* (see Fig. 1).

The samples for novels and short stories show substantial overlap. Two important exceptions stand out: not surprisingly, *short* and *succinct* rank highly only in the short story sample, whereas *romantic*, *thrilling*, *riveting*, and *exciting* are listed far more frequently with reference to novels. Plays and comedies also show a considerable overlap; nevertheless, positive emotion terms are more frequent in the comedies-only sample, whereas negative emotion terms play a larger role in the general category of plays, most likely reflecting the traditional hegemony of the tragedy over the comedy.

Terms used to denote dimensions of aesthetic appeal can be expected to have an evaluative connotation (cf. Juslin, 2013); accordingly, most of the terms which were listed by our participants are evaluative in nature (Jacobsen et al., 2004, report the same for their data). However, a few of the terms obtained in our study are less obviously evaluative. *Rhythmical*, *short*, and *sad* in particular at first appear to be nothing but purely descriptive attributes. Nevertheless, ratings of sadness have been found to correlate positively with ratings for aesthetic appreciation in several studies (Hanich, Wagner, Shah, Jacobsen, & Menninghaus, 2014; Taruffi & Koelsch, 2014); similarly, ratings confirm that rhythmically regular poems correlate positively with aesthetic liking (Obermeier et al., 2013). And even the term *short* has, in fact, a tradition as a distinct aesthetic merit term in rhetoric (see Section 3.3.4). Therefore, there are reasons to assume that the use of seemingly purely descriptive terms may also, in the context of aesthetics, include an evaluative dimension.

Table 1
Descriptive statistics of the subsamples.

| #	Subsamples	Location	Number of participants	Number of answers	Number of terms	Mean answers per participant	Number of terms listed by more than 5% of the subsample	Number of terms listed by more than 10% of the subsample	1	2	3	4	5
1	Literature	Berlin	274	1356	591	4.95	8	7					
2	Poems	Hamburg	423	2751	826	6.50	18	9	0.64				
3	Novels	Hamburg	134	962	411	7.18	23	11	0.63	0.42			
4	Short stories	Berlin	223	1288	570	5.78	16	10	0.54	0.34	0.66		
5	Plays	Berlin	172	1154	505	6.71	17	11	0.42	0.36	0.58	0.50	
6	Comedies	Hamburg	318	2158	791	6.79	16	10	0.50	0.32	0.60	0.57	0.58
	overall		1544	9669	2131	6.26							
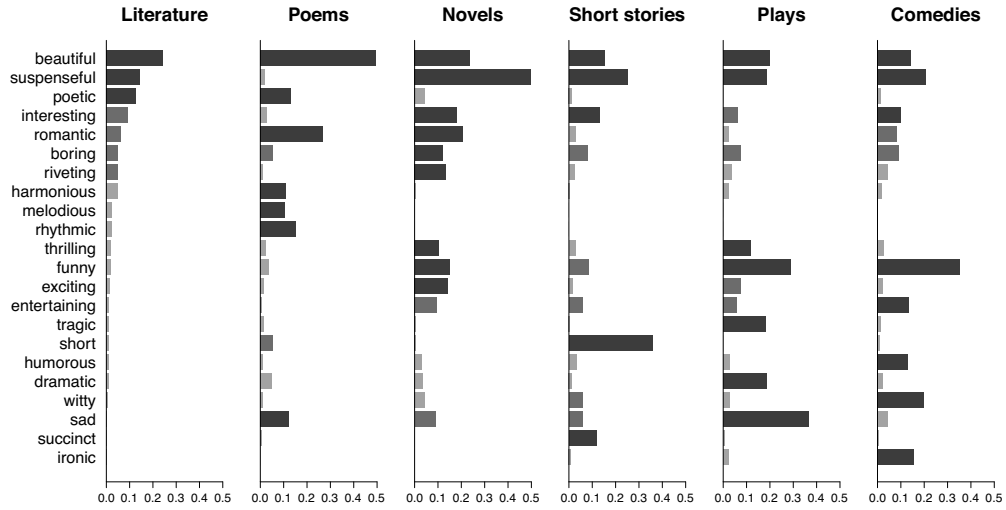
Fig. 1. Relative frequencies plotted for terms mentioned by at least 10% of the participants in one subsample, ordered by frequency in the literature subsample. The grey scale indicates whether the term was mentioned/listed by less than 5% (light grey), 5% to 10% (grey), or more than 10% (dark grey) of the participants of each individual sample.
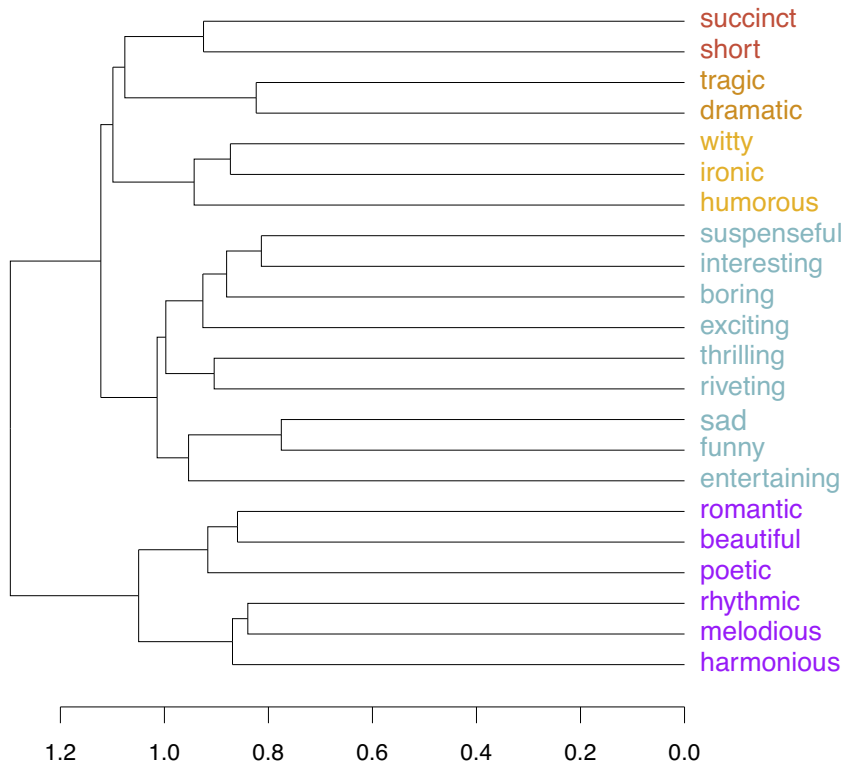

Fig. 2. Dendrogram of the Hierarchical Cluster Analysis for the 22 adjectives.
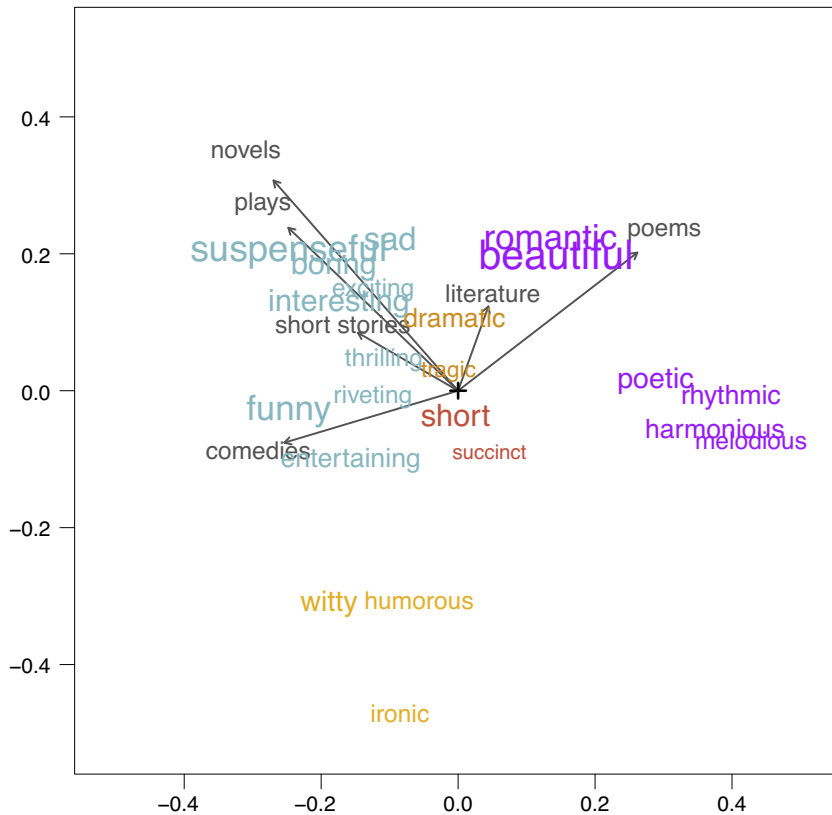
Fig. 3. Two-dimensional Multidimensional Scaling solution. The five clusters are coded by color, the location of the six subsamples in this space is indicated by the black vectors.

3.2.4. Overall mapping

In order to integrate all genre-specific findings into a semantic "map" of the field of literary aesthetics, we calculated a dissimilarity matrix based on the co-occurrence of the terms on the lists of all participants and analyzed this matrix using HCA and MDS. As a dissimilarity measure we used the Jaccard Index (Real & Vargas, 1996), which takes into account only positive matches and non-matches.

Three distinct clusters emerged from our analysis: one comprising the term *beautiful* along with sound- and prosody-related terms (e.g., *melodious*, *rhythmical*), a second comprising plot- and emotion-related terms (e.g., *suspenseful*, *thrilling*, *sad*), and a third that seemed to be more heterogeneous at first, but that could in a second step be divided into three subclusters that appear to be related to the specific natures of short stories, comedies, and plays in general (see Fig. 2).

The dissimilarity matrix was also fed into a classical MDS procedure; Fig. 3 depicts the two-dimensional MDS solution.[6] Furthermore, based on the MDS coordinates of the terms and their frequencies in the samples, we calculated points that represent the localization of the subsamples in this two-dimensional aesthetic appreciation space and plotted them as vector arrows into the MDS plot. The first (horizontal) dimension shows a clear bipolar structure with terms designating "poetic" qualities and musicality on the one side (with the term "poetic" linguistically not being restricted to the verbal features of poetry but also including "poetic" feelings, atmospheres, moods, and related phenomena in a broader sense that extends to other art forms) and narrative and conceptual structure on the other. The second (vertical) dimension distinguishes between terms referring to emotional and potentially immersive states at one end and humorous and intellectual states at the other.

---

[6] The Goodness-of-Fit (GOF) for the two-dimensional solution is only .17, and thus remains below conventional criteria. Yet in the specific MDS-procedure we applied, adding further dimensions – while increasing the GOF – left the first two dimensions unchanged (cf. Gower, 1966). Furthermore, only the first two dimensions, which by definition capture more variance than the subsequent ones, appear to reflect more general aspects of the aesthetic space of literature, whereas the other dimensions show a stronger focus on more particular aspects. In any event, the two-dimensional solution presented in Fig. 3 captures only a limited amount of the variance included in our data.

3.3. Discussion

The results of this study highlight the variance and complexity of how the perceived aesthetic appeal dimensions of literature are mapped onto verbal concepts by readers. Many of the terms listed did not make it past the cut-off procedures described above; by implication, the terms that passed the cut-off procedure by no means represent the entire variance in aesthetic appeal dimensions associated with the categories under scrutiny. Nonetheless, the terms that ended up being named by at least 10% (22) or 5% (52) of the participants in at least one sample still offer a substantial – and, as will become evident in study 2, highly distinctive – range of aesthetic appeal dimensions regarding literature. In the following, we will individually discuss the most high-ranking terms.

3.3.1. Beautiful

The term *beautiful* – being listed most frequently overall – proves its preeminence in aesthetics in the field of literature no less than in other domains that have been studied, despite numerous claims made since the 18th century that literature no longer primarily aims at beauty (cf. Jauss, 1991; Lessing, 1984). To be sure, the German word *schön* has a considerably broader application and higher usage frequency than the English term *beautiful* (Wortschatzprojekt Universität Leipzig, 2015). However, contrary to an indiscriminate bias toward beauty, our data reveal differences regarding the role of the term *beautiful* in the different aesthetic domains, with the patterns of co-occurrence providing more specific insight into what readers might mean when labeling literary texts as *beautiful* (see Fig. 3): the term clusters most strongly with *romantic*, followed by *poetic*, *rhythmic*, *melodious*, and *harmonious*. Apparently, the attribution of beauty to a literary text relies more strongly on the prosodic qualities of language and on "romantic" or "poetic" feelings, moods, and atmospheres than on narrative plot in the stricter sense. In line with this pattern of co-occurrence, beauty seems to be preeminently associated with qualities of poetry. The more colloquial meanings of *poetic*, i.e., soulful and enchanting, hint at the affective qualities of the term *beautiful*. *Romantic* is the only one of the clustering terms that could be read as referring to certain content features (typical of the Romantic period), but it may equally pertain to a characteristically Romantic form, as well as (in the everyday sense of the word) to powerfully sentimental, evocative, or atmospheric effects that could be due to both form and content.

3.3.2. Suspenseful

The term *suspenseful* is the second most frequently mentioned term overall; its highest scores occur in the plot-based samples. The fact that most of the forms and genres we examined are plot-based may therefore serve as an explanation for the term's overall high salience. In the samples for novels, short stories, and comedies, *suspenseful* is listed more often than *beautiful*.

Narrative suspense makes readers fear or desire particular plot outcomes, which can vary in likelihood during the different stages of a narrative; this creates a need for resolution that accompanies the entire suspenseful trajectory (cf. Anz, 1998; Berlyne, 1960; Carroll, 1996; Fill, 2007; Lehne & Koelsch, 2015; Löker, 1976; Wulff, 1996; Zillmann, 1980). Importantly, this tense feedback loop between changing degrees of uncertainty and anticipation of the potential resolution strongly depends on the way the content is narratively arranged. In fact, artistic narrative arrangements can even build suspense when readers already know the outcome of the plot (cf. Gerrig, 1989; Hoeken & van Vilet, 2000; Lehne & Koelsch, 2015; Yanal, 1996). Suspense also serves to increase the general emotional susceptibility of the audience (cf. Oatley, 1999; Vorderer, Wulff, & Friedrichsen, 1996). Consistent with these theoretical assumptions, the term *suspense* clusters strongly with emotion terms such as *thrilling* and *riveting* in our sample.

3.3.3. Interesting

The term *interesting* is the fourth highest scoring term in our literature, novel, and short story subsamples. In theoretical aesthetics, the attribute "interesting" has been acknowledged as an essential category since the 18th century (cf. Diderot, 1995; Garve, 1974; Ostermann, 1997; Sulzer, 1967). Friedrich Schlegel in particular developed a concept of aesthetic evaluation in which the interesting, not the beautiful, is the main reference point (Schlegel, 2001). More recent empirical research supports the notion of the interesting as a prime category of aesthetic appreciation (Silvia, 2005, 2010).

Our poetry subsample scores very low on the dimension *interesting*, which may be due to various reasons. Firstly, since the 18th century the interesting has been considered to appeal to the intellect rather than the heart, which is why it was primarily applied to novels and plays; poetry, by contrast, long continued to be largely conceived as focusing on highly personal feelings and on beautiful poetic diction (Hegel, 1975; Lukács, 1971; Schlegel, 2001). Secondly, in accordance with Schlegel's (2001) theoretical reflections, Silvia (2010) suggests that a high degree of novelty is an important factor for finding something interesting. However, the poetry with which average readers typically come into contact (in school, at university, in printed anthologies, on greeting cards, at wedding or birthday celebrations) is often representative of traditional lyrical forms and motifs; contact with lyrical texts that a contemporary reader would deem innovative, and therefore more interesting, is rare by comparison (cf. Shetley, 1993).

3.3.4. Other terms

Overall, we found a high number of emotion-related terms. Needless to say, emotion generally plays an enormous role in the aesthetic experience with and the aesthetic evaluation of literature. The particularly high scores of emotion terms in the

lists for plays in general and comedies in particular confirm the notion that drama as a genre aims primarily at evoking strong affective responses in audiences (Aristotle, 2005; Zillmann, 1980, 1983).

Two frequently used quality judgments regarding works of literature, namely, attributions of goodness or badness ("a good novel", a "bad play"), are entirely lacking from our data. Attributions of the adjectives *good* and *bad* reflect a particularly abstract and unspecific type of aesthetic judgment that fails to provide hints regarding either textual features or experiential response dimensions; the explicit instruction to list aesthetics-specific adjectives may have prevented our participants from listing this simple dichotomy. The absence of these terms may thus well be a task effect; confirming this assumption, the terms "good" and "bad" were likewise not listed in the studies on other aesthetic domains that used similar instructions (Augustin, Wagemans et al., 2012; Istók et al., 2009; Jacobsen et al., 2004). Similarly, the term *pleasant* does not appear in our results, despite being one of the more frequently employed items in aesthetic rating scales. This might, again, be due to the relatively unspecific nature of the term, designating as it does a broad range of positive experiences that could be mediated by a great variety of emotions and other response dimensions, as well as aesthetically relevant text features.

The analysis of the terms listed for poetry reveals a strong bias for music-related terms, most notably *rhythmical*, *melodious*, and *harmonious*. These results coincide with a long-standing tradition of equating poets with singers and poems with songs (e.g., Hegel, 1975; Herder, 1998; Nietzsche, 1993), which became a cornerstone of the Romantic concept of poetry and has shaped prototypical notions of poetry ever since. Our bottom-up data demonstrate the endurance of this concept far beyond a scholarly awareness of the underlying intellectual traditions. At the same time, terms designating narrative and conceptual properties, such as *thrilling*, *riveting*, and *suspenseful*, which prove very important for the novel, are rarely mentioned in the poetry subsample; this supports the notion that both formal and content-related characteristics of literary genres cluster with specific aesthetic evaluations (cf. Ribeiro, 2012).

As we expected, the data obtained for novels and short stories cluster closely. Yet we also found some differences that are worth a closer look. First, the frequencies for the adjective *short* set the genres neatly apart. One might dismiss this as an indication that our participants confounded the name of the genre with its potential aesthetic merits; however, shortness (gr. *brachytes*, lat. *brevitas*) is in fact a well-established category designating the rhetorical achievement of condensing a message and making it unusually compact (Quintilian, 1953; for an empirical study on effects of rhetorical brevity in the processing of proverbs, cf. Menninghaus et al., 2015). In fact, a short story that is not short would violate genre expectations, including aesthetic reward expectations. Thus, the high scores for *short* may not, after all, be a mere task effect.

Second, the terms *romantic*, *thrilling*, *riveting*, and *exciting* were listed very frequently for novels but do not seem to be expected properties of short stories—which might baffle readers of Poe, Maupassant, or Chekhov. One possible explanation for this could be found in the German curricula: the German short story after 1945 is featured prominently in German secondary schools; therefore, most or all of our participants will have been exposed not only to examples of the genre itself, but also to a catalog of its specific features: this type of short story commits to straightforward, open storylines while forgoing complicated chains of cause and effect, highly emotional suspense–resolution sequences, and even unambiguous endings (Weyrauch, 1989). Knowledge of this particular tradition and resulting genre expectations might be reflected in the lower ratings for *exciting*, *thrilling*, and *riveting* in our short story subsample compared to the novel subsample.

3.3.5. The two dimensions of the MDS solution

Terms referring to poetic qualities of a text and terms referring to its narrative and/or thematic structure seem to form the opposite ends of the horizontal axis of our MDS solution (see Fig. 3). These poles are most markedly represented by the terms *beautiful*, which refers mostly to the poetic and formal appeal of texts, and *suspenseful*, which primarily refers to plot trajectories. However, this result should not be read as supporting a form-content dissociation. In fact, the plot side of the spectrum actually represents not just content alone, but a particular type of interaction between form and content.

The vertical axis of our MDS graph (see Fig. 3) shows terms that refer to more intellectual aspects of reading (*witty*, *humorous*, *ironic*) on the one hand, and terms that pertain to the text's emotional content or the affective responses it may evoke (e.g., *suspenseful*, *sad*, *boring*) on the other (for a basis for this distinction between emotions perceived as being represented by a work of art and the actual felt emotions of the audience, cf. Juslin, 2013; see also Regel, Gunter, & Friederici, 2011; Nagels et al., 2013). Notably, in this graph comedies and plays in general (which by definition include comedies) are clearly separated: compared to comedies, plays in general are a lot farther removed from the pole that entails *witty*, *humorous*, and *ironic*, and far closer to the pole that entails *suspenseful*, *sad*, *boring*. We surmise that the differences between the responses given for comedies only and for plays in general primarily reflect the role of tragedy. In accordance with this interpretation, the category of plays in general includes the term *tragic*, which is listed much more frequently than *humorous* and *witty* (although less frequently than *funny*), and also the term *sad*, which is the highest-scoring term in the entire subsample. In all likelihood, separate data for tragedy would have further enlarged this distance between plays in general and comedies, because the inclusion of comedies in the category of plays in general should keep this distance at a more moderate level. Thus, the lack of a separate data set for tragedies is at least partly compensated by the differences between our data for comedies and for plays in general. In light of this interpretation of the data, it appears that from an affective perspective, the two dramatic genres of comedy and tragedy are perceived to have two entirely different affective signatures, not simply in valence (positive − negative) and affective content (funny − sad), but also regarding the quality of the evoked affects and the extent to which they are linked to cognitive processes (cf., e.g., Taylor, 1988; Morreall, 1983).

Table 2
Overview of the comparison between studies.

Study	Year	Object class	N	Language	Answer format	Design	Procedure	Time
Jacobsen et al. (2004)	2004	Objects	311	German	Adjectives	–	Paper & pencil	2 min
Istók et al. (2009)	2009	Music	300	Finnish	Adjectives	–	Paper & pencil	5 min
Augustin, Wagemans et al. (2012)	2012	Buildings	178	Dutch	No restriction	Within participant	Online	No restriction
		Cars	177					
		Clothing	175					
		Faces	175					
		Geometric shapes & patterns	173					
		Interior design	175					
		Landscapes	177					
		Visual art	177					
Our study		Literature	274	German	Adjectives	Between participant	Paper & pencil	2 min
		Poems	423					
		Novels	134					
		Short stories	223					
		Plays (tragedies or comedies)	172					
		Comedies	318					

While the terms on the more intellectual end of the spectrum have a primary semantic focus on appeal dimensions ascribed to the text (e.g., witty, humorous, ironic), the affective end of the spectrum entails a large number of terms that place a primary focus on the text's (potential) effect on the reader (e.g., *suspenseful, exciting, thrilling*).

4. Comparing literature with other domains

4.1. Studies and data

In order to compare our results regarding the domain of literature with other (aesthetic) domains, we conducted analyses including the results of our study and the results published by Jacobsen et al. (2004) on aesthetic objects at large, Istók et al. (2009) on music, and Augustin, Wagemans et al. (2012) on eight different classes of visual objects: buildings, cars, clothing, faces, interior design, landscapes, geometric shapes and patterns, and visual art. Although the studies all employed a free listing method, there are several differences between them: they differed regarding the language in which they were conducted (German, Finnish, or Dutch), restrictions on the answer format (only adjectives vs. unconstrained), and the time accorded for the task (2 min, 5 min, or unconstrained); for an overview, see Table 2.

The terms mentioned by at least 5% in the respective samples, along with their frequencies, were the basis for the following analyses (these data are available as part of the published studies; for our data, see Section 3.2.2). Based on these frequency data, we calculated the OVL for each sample pairing (see Table S1), which we fed into an HCA and an MDS procedure.

4.2. Results

Unsurprisingly, individual samples tend to show the greatest overlap with other samples from the same domain; however, we identified a few interesting exceptions. Literature in general primarily intersects with novels and poems, but when compared to nonliterary domains, it shows the greatest overlap with visual arts. However, the literary subdomain of poems – again, when compared to nonliterary domains – shows the greatest overlap with music, followed by the visual arts and visual aesthetic objects in general. The terms mentioned for music have the greatest overlap with those listed for poetry, followed by visual art. Visual art, in turn, showed the greatest overlap with clothing, followed by literature in general, and then visual aesthetic objects in general, faces, and cars. Hence, our data also reveal some considerable similarities between the terms used to conceptualize the aesthetic appeal dimensions of visual art, music, poetry, and literature in general.

Moreover, we transformed the OVL matrix into a dissimilarities matrix and fed it into an HCA procedure, using Ward's method. The result shows two clearly separate clusters, one of which comprises most samples of the visual domain. The other cluster includes the literary samples, as well as the music and landscape samples. Both clusters can be further divided into two subclusters: within the visual cluster, manufactured artifacts are grouped together, and so are faces, geometrical
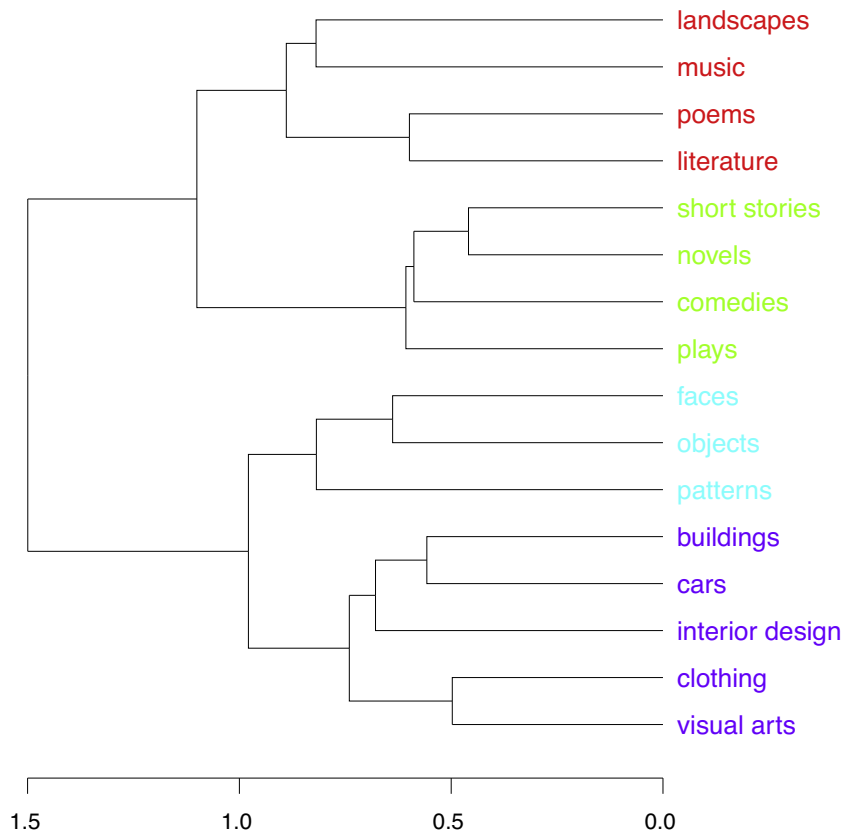
Fig. 4. Dendrogram of the Hierarchical Cluster Analysis for the 16 object classes in the compared studies.

patterns, and visual aesthetic objects. Within the literary cluster, one subcluster comprises the narrative genres, and the other comprises literature in general, poems, music, and landscapes (Fig. 4).

An MDS based on the dissimilarities matrix largely mirrors the result of the cluster analysis (see Fig. 5).[7] Again, we find a cluster comprised of the plot-based literary genres, and one comprising literature in general, poems, and music, the last three clustering somewhat more loosely. There are slight differences in the two non-literature-related clusters: one comprises manufactured artifacts of practical use, including geometrical patterns, but not the visual arts; the other contains the visual arts, landscapes, faces, and visual aesthetic objects in general. An interpretation of the two dimensions of this MDS is not straightforward.

4.3. Discussion

A particularly interesting result of the comparison of our data with those of previous studies is the strong overlap between the terms used for music and poetry; this underscores our finding that the phonological (including prosodic) properties of poetry, along with its emotional aspects, are deemed its most distinctive characteristics. Moreover, the content of poetry – consisting, as it typically does, of situational miniatures and depictions of momentary mood states (cf. Schlaffer, 2012) – usually does not involve anything resembling a full-blown narrative plot. As a result, poetry appears to favor a comparison with music over a comparison with other literary genres that are based on plot. Of course, a number of subgenres might render this clear-cut solution somewhat complicated: the narratively structured ballad, the metered and sometimes even rhymed tragedy, free verse. However, while our data for poetry do not cover these and other subgenres, they do reflect not merely the expectations of our participants, but also important – if not predominant – features of poetry.

---

[7] The two-dimensional MDS solution only yielded a GOF of .41. However, as was the case for the first MDS reported in this paper, the pattern found for the two first dimensions did not change as we added more dimensions to the MDS solution. We here report this MDS solution only as an additional visualization of the data underlying the cluster analysis.
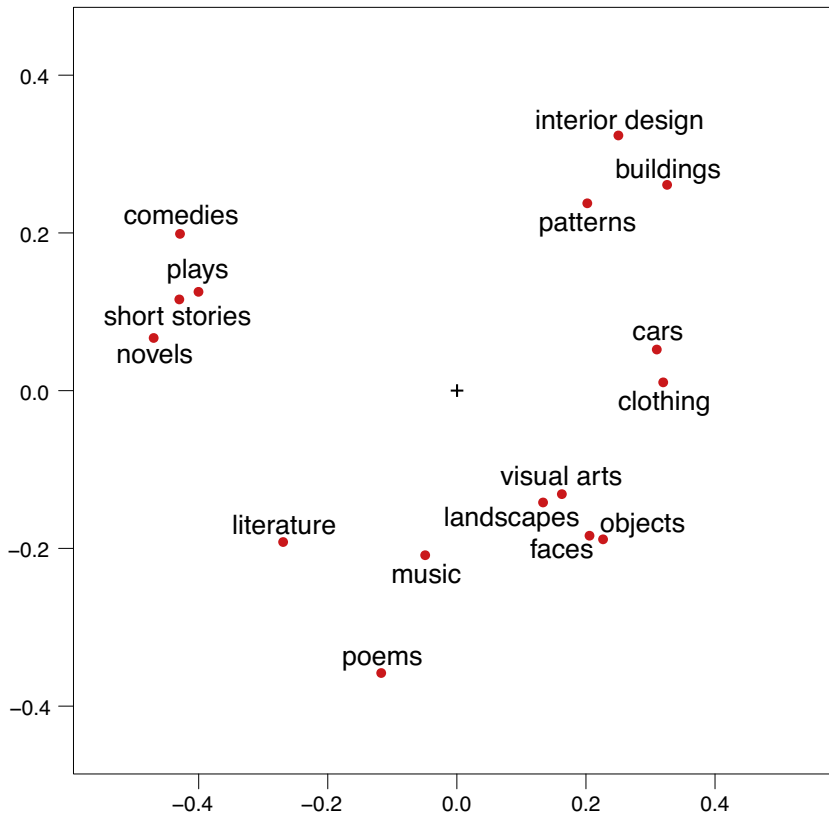
Fig. 5. Two-dimensional Multidimensional Scaling solution for the 16 object classes of the compared studies.

Generally, the artistic domains (literature, visual arts, music) share many of the terms used to conceptualize their aesthetic appeal; this corresponds to the way in which 18th century aesthetics comprised all arts, which until then had been considered exclusively as individual *artes*, under the new all-encompassing singular of *art* (cf. von Schelling, 1989). Moreover, our finding that literature in general, when compared to nonliterary domains, has the greatest overlap with visual arts can be read as supporting the long-standing *ut pictura poesis* tradition (cf. Horace, 2011; Lessing, 1984).

In some respects, however, we found clear domain-based differences: *ugly*, for instance, is listed only for visual objects and music. It appears that ugliness in literature, if it can be found at all, is a matter of semantic content rather than aesthetics (Eco, 2007). The clearest antonym to *beautiful* in the literature sample is the term *boring* (cf. Lorand, 1994) – which in turn is barely relevant for the other domains.

5. General discussion

5.1. Key results

(1) In all studies to date (including ours) that analyze the linguistic terms used to designate expected or perceived aesthetic appeal dimensions, the concept of beauty has been found to be prevalent in all tested domains. At the same time, *beautiful* is by no means the most frequently mentioned term in each of the domains in our study. It scored highest in the subsamples for poetry and literature in general; yet in the plot-based subsamples, other terms turned out to be more significant: *suspenseful* in the case of the novel as well as the short story, *funny* and *sad* in the subsample of plays in general, and *suspenseful*, *funny*, *witty*, and *ironic* in the comedy subsample. Nevertheless, counting all terms listed in the entire study, *beautiful* is mentioned most frequently overall. Regarding only the categories of literature, which were investigated in the present study, attributions of the term *beautiful* cluster strongly with terms referring to the "poetic" qualities of diction and emotional tonality, and far less with terms that are more closely associated with plot trajectories. Overall, the selection of verbal concepts we gathered reveals that, in many instances, concepts with a long-standing tradition in aesthetic theory play an important role in non-expert conceptualizations of the aesthetics of literature; this appears to apply far beyond any scholarly awareness of this tradition.

(2) The comparison of previous studies with our own findings underscores the commonalities between music and poetry that can already be extracted from our own data alone. The consistent data pattern confirms that these commonalities are not only seen by experts familiar with the longstanding "Romantic" tradition of theorizing poetry, but are seemingly also part of the way readers conceive of poetry to this very day.

(3) Regarding the aesthetics of literature, our data show two preeminent factors: plot-based appeal dimensions, which are primarily important for novels, short stories, and plays, and "poetic" appeal dimensions, which are mostly based on prosody- and music-related aspects as well as on special emotional tonalities. This second group of appeal dimensions is by no means exclusively related to poetry (where it is particularly strong), but extends into the broader sample of literature in general.

(4) Our approach "from below" identifies verbal concepts which, together, form a conceptual map of prototypical aesthetic perceptions and evaluations regarding particular literary genres. The participants in our study did not evaluate individual works of literature, but listed expected dimensions of aesthetic appreciation regarding literary genres; furthermore, they chose their own wording rather than responding to expert-selected items. No data of this type has been collected in previous empirical research on literature.

5.2. Limitations

First, it is important to note that the appeal dimensions extracted in our study rely on abstract representations of genre concepts. In all likelihood, these abstract concepts are idealized entities that do not cover the full historical variance of the concrete works of literature found across multiple languages and cultures. Moreover, our data exclusively reflect appeal dimensions identified by a sample of today's readers and are likely to be subject to ongoing change. Therefore, the range of dimensions of appeal identified here must be considered as a first, rather general guidepost.

Second, the four studies on the conceptual structure of specific aesthetic domains that we compared were conducted in different languages: German, Finnish, and Dutch (see Table 2). All data, including our own, have been translated into English. Therefore, any comparison must bear in mind language-specific differences in usage frequency, emotional valence, habitual connotations, and, not least, semantics between these linguistic groups (for other procedural differences see Table 2).

Third, collecting free listing data by definition does not provide any information about the participants' previous genre exposure and preferences, though these may have influenced the results we received (regarding the potential impact of genre exposure, see Fong, Mullin, & Mar, 2013). The task is also not suited to capture previous experiences that may have prompted our participants to include particular items in their lists. Moreover, we can not extrapolate which replies may have been primed by theoretical concepts of the relevant genres (e.g., concepts our participants acquired in school) and which ones were a result of own reading experiences.

5.3. Future research

To acquire a deeper understanding of how recipients perceive and conceptualize the dimensions of aesthetic appeal for different art forms, and to provide a map of the dimensions of aesthetic appreciation, additional data for other artistic domains will be indispensable. This includes multimodal domains like film, opera, performance art, and dance, which in many ways bridge the gap between literature and music, literature and theatrical performance, and literature and the visual arts. At the other end of the scale, it would be beneficial to make finer distinctions within the individual domains and subdomains – for instance, a subdomain like the novel could be divided into smaller and more specific subgenres, such as science fiction novels, detective novels, romance novels, and so on. Furthermore, cross-linguistic as well as cross-cultural comparisons for the same artistic domain would be very useful in order to establish the extent of variety we should accept when translating items. Lastly, a study such as ours, which is based on general, open questions, could be complemented with a corpus-based approach by extracting adjectives appraising literature from contemporary and historical newspapers and journals, online and print reviews, blurbs, and other types of paratext. Another possibility would be empirical research that collects and analyzes more comprehensive comments on the topic of literature from research participants (for a study along these lines, cf. Dixon & Bortolussi, 2009). In this way, the more general aesthetic appeal dimensions we recorded would be complemented by terms that readers use to conceptualize their impressions regarding a variety of specific literary texts.

Acknowledgments

Appendix A.

Original German instruction

"Bitte schreiben Sie Wörter auf, die man zur Beschreibung der Ästhetik von Literatur verwenden kann. Bitte benutzen Sie nur Adjektive (Eigenschaftswörter). Sie haben ab jetzt 2 Minuten Zeit."

Table A1
Terms listed by at least 5% of the participants in one subsample.

English (translation)	German (original)
Aesthetic	ästhetisch
Amusing	amüsant
Beautiful	schön
Boring	langweilig
Brief	knapp
Charming	lieblich
Comical	komisch
Creative	kreativ
Creepy	gruselig
Deeply moving	ergreifend
Diverse	abwechslungsreich
Diverting	kurzweilig
Dramatic	dramatisch
Emotional	emotional
Entertaining	unterhaltend
Entertaining	unterhaltsam
Euphonious	wohlklingend
Exciting	aufregend
Fantastic	phantastisch
Fluent	fließend
Funny	lustig
Harmonious	harmonisch
Humorous	humorvoll
Imaginative	phantasievoll
Instructive	lehrreich
Interesting	interessant
Ironic	ironisch
Long	lang
Long-winded	langatmig
Melancholic	melancholisch
Melodic	melodisch
Melodious	klangvoll
Metaphorical	metaphorisch
Moving	bewegend
Open	offen
Passionate	leidenschaftlich
Poetic	poetisch
Pointed	pointiert
Profound	tiefgründig
Realistic	realistisch
Rhyming	reimend
Rhythmic	rhythmisch
Riveting	fesselnd
Romantic	romantisch
Sad	traurig
Sarcastic	sarkastisch
Satirical	satirisch
Sentimental	gefühlvoll
Short	kurz
Stimulating	anregend
Succinct	prägnant
Surprising	überraschend
Suspenseful	spannend
Thrilling	mitreißend
Touching	berührend
Tragic	tragisch
Witty	witzig

Appendix B. Supplementary data

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.poetic.2016.02.001.

References

Anz, T. (1998). Spannungskunst und Glückstechniken. *Literatur und Lust*. Munich: C.H. Beck, 150–171.

Aristotle (2005 [original approx. 335 BC]). *Poetics*. (J. Sachs, Trans) Newburyport, MA: Focus.

Augustin, M. D., Carbon, C. C., & Wagemans, J. (2012). Artful terms: a study on aesthetic word usage for visual art versus film and music. *I-Perception*, *3*(5), 319–337. http://dx.doi.org/10.1068/i0511aap.

Augustin, M. D., Wagemans, J., & Carbon, C. C. (2012). All is beautiful? Generality vs. specificity of word usage in visual aesthetics. *Acta Psychologica*, *139*(1), 187–201. http://dx.doi.org/10.1016/j.actpsy.2011.10.004.

Berlyne, D. E. (1960). *Conflict, arousal, and curiosity*. New York: McGraw-Hill.

Brattico, E., Bogert, B., & Jacobsen, T. (2013). Toward a neural chronometry for the aesthetic experience of music. *Frontiers in Psychology*, *4*(206) . http://dx.doi.org/10.3389/fpsyg.2013.00206.

Carroll, N. (1996). Toward a theory of film suspense. *Theorizing the moving image*. Cambridge: Cambridge University Press, 94–115.

Diderot, D. (1995 [original 1767]). *On Art II: The Salon of 1767*. (J. Goodman, Trans.) New Haven: Yale University Press.

Dixon, P., & Bortolussi, M. (2009). Readers' knowledge of popular genre. *Discourse Processes*, *46*(6), 541–571. http://dx.doi.org/10.1080/01638530902959570.

Dixon, P., Bortolussi, M., & Sopčák, P. (2015). Extratextual effects on the evaluation of narrative texts. *Poetics*, *48*, 42–54. http://dx.doi.org/10.1016/j.poetic.2014.12.001.

Eco, U. (2007). *On Ugliness*. (A. McEwen, Trans.) London: Harvill Secker.

Fechner, G. T. (1876). *Vorschule der Ästhetik*. Leipzig: Breitkopf & Härtel.

Fehr, B., & Russell, J. A. (1984). Concept of emotion viewed from a prototype perspective. *Journal of Experimental Psychology*, *113*, 464–486. http://dx.doi.org/10.1037/0022-3514.60.3.425.

Fill, A. (2007). *Das Prinzip Spannung: Sprachwissenschaftliche Betrachtungen zu einem universalen Phänomen*. Tübingen: Narr Francke Attempto.

Fong, K., Mullin, J. B., & Mar, R. A. (2013). What you read matters: the role of fiction genre in predicting interpersonal sensitivity. *Psychology of Aesthetics, Creativity and the Arts*, *7*(4), 370–376. http://dx.doi.org/10.1037/a0034084.

Garve C. (1974 [original 1779]). *Popularhistorische Schriften über literarische, ästhetische und gesellschaftliche Gegenstände*. Stuttgart: J. B. Metzler.

Gerrig, R. J. (1989). Suspense in the absence of uncertainty. *Journal of Memory and Language*, *28*, 633–648. http://dx.doi.org/10.1016/0749-596X(89) 90001-6.

Gower, J. C. (1966). Some distance properties of latent root and vector methods used in multivariate analysis. *Biometrika*, *53*(3–4), 325–338. http://dx.doi.org/10.1093/biomet/53.3-4.325.

Green, M. C., & Brock, T. C. (2000). The role of transportation in the persuasiveness of public narratives. *Journal of Personality and Social Psychology*, *79*(5), 701–721. http://dx.doi.org/10.1037//0022-3514.79.5.701.

Hanich, J., Wagner, V., Shah, M., Jacobsen, T., & Menninghaus, W. (2014). Why we like to watch sad films: the pleasure of being moved in aesthetic experiences. *Psychology of Aesthetics, Creativity and the Arts*, *8*(2), 130–143. http://dx.doi.org/10.1037/a0035690.

Hegel, G.W. F. (1975 [original 1835]). *Aesthetics: Lectures on Fine Art*. (T. M. Knox., Trans.) Oxford: Clarendon Press.

Herder, J.G. (1998 [original 1792–1800]). *Schriften zu Literatur und Philosophie*. (*Werke*, vol. 8). Ed. H. D. Irmscher. Frankfurt am Main: Deutscher Klassiker Verlag.

Hoeken, H., & van Vilet, M. (2000). Suspense, curiosity, and surprise: how discourse structure influences the affective and cognitive processing of a story. *Poetics*, *27*(4), 277–286. http://dx.doi.org/10.1016/S0304-422X(99)00021-2.

Horace (2011 [original approx. 19 BC]). The art of poetry, In: R. Cowan (ed.). *Horace: Satires and Epistles*. (J. Davie, Trans.) New York: Oxford University Press, 106–118.

Inman, H. F., & Bradley, E. L. (1989). The overlapping coefficient as a measure of agreement between probability distributions and point estimation of the overlap of two normal densities. *Communications in Statistics—Theory and Methods*, *18*(10), 3851–3874. http://dx.doi.org/10.1080/03610928908830127.

Istók, E., Brattico, E., Jacobsen, T., Krohn, K., Muller, M., & Tervaniemi, M. (2009). Aesthetic responses to music: a questionnaire study. *Musicae Scientiae*, *13*(2), 183–206. http://dx.doi.org/10.1177/102986490901300201.

Jacobs, A. M. (2015). Neurocognitive poetics: methods and models for investigating the neuronal and cognitive-affective bases of literature reception. *Frontiers in Human Neuroscience*, *9*, 186. http://dx.doi.org/10.3389/fnhum.2015.00186.

Jacobsen, T., Buchta, K., Kohler, M., & Schroeger, E. (2004). The primacy of beauty in judging the aesthetics of objects. *Psychological Reports*, *94*(3, Pt. 2), 1253–1260. http://dx.doi.org/10.2466/pr0.94.3c.1253-1260.

Jauss, H. R. (1991). *Die nicht mehr schönen Künste: Grenzphänomene des Ästhetischen*. Munich: Fink.

Juslin, P. N. (2013). From everyday emotions to aesthetic emotions: Towards a unified theory of musical emotions. *Physics of Life Reviews*, *10*(3), 235–266. http://dx.doi.org/10.1016/j.plrev.2013.05.008.

Kant, I. (2007 [original 1790]). *The Critique of Judgment*. (J. C. Meredith, Trans.) Oxford: Oxford University Press.

Koopman, E. M. (2013). The attraction of tragic narrative: catharsis and other motives. *Scientific Study of Literature*, *3*(2), 178–208. http://dx.doi.org/10.1075/ssol.3.2.03koo.

Koopman, E. M. (2015). Why do we read sad books? Eudaimonic motives and meta-emotions. *Poetics*, *52*, 18–31. http://dx.doi.org/10.1016/j.poetic.2015.06.004.

Kuehnast, M., Wagner, V., Wassiliwizky, E., Jacobsen, T., & Menninghaus, W. (2014). Being moved: linguistic representation and conceptual structure. *Frontiers in Psychology*, *5*, . http://dx.doi.org/10.3389/fpsyg.2014.01242.

Kuijpers, M. M., Hakemulder, F., Tan, E. S., & Doicaru, M. M. (2014). Exploring absorbing reading experiences: developing and validating a self-report scale to measure story world absorption. *Scientific Study of Literature*, *4*(1), 89–122. http://dx.doi.org/10.1075/ssol.4.1.05kui.

Kuiken, D., Campbell, P., & Sopčák, P. (2012). The experiencing questionnaire: locating exceptional reading moments. *Scientific Study of Literature*, *2*(2), 243–272. http://dx.doi.org/10.1075/ssol.2.2.04kui.

Leder, H., Belke, B., Oeberst, A., & Augustin, D. (2004). A model of aesthetic appreciation and aesthetic judgments. *British Journal of Psychology*, *95*(4), 489–508. http://dx.doi.org/10.1348/0007126042369811.

Lehne, M., & Koelsch, S. (2015). Towards a general psychological model of tension and suspense. *Frontiers in Psychology*, *6*(79) . http://dx.doi.org/10.3389/fpsyg.2015.00079.

Lessing, G.E. (1984 [original 1766]). *Laocoön: An Essay on the Limits of Painting and Poetry*. (E. A. McCormick, Trans.) New York: Bobbs-Merrill.

Löker, A. (1976). *Film and suspense*. Istanbul: Matbassi.

Lorand, R. (1994). Beauty and its opposites. *The Journal of Aesthetics and Art Criticism*, *52*(4), 399–406. http://dx.doi.org/10.2307/432027.

Lukács, G. (1971 [original 1916]). *Theory of the Novel*. (A. Bostock, Trans.) London: Merlin Press/Cambridge, MA.: MIT Press.

Marx, W. (1976a). Die Messung der assoziativen Bedeutungsähnlichkeit. *Zeitschrift für Experimentelle und Angewandte Psychologie*, *23*(1), 62–76.

Marx, W. (1976b). Die statistische Sicherung des Überlappungs-Koeffizienten. *Zeitschrift für Experimentelle und Angewandte Psychologie*, *23*(2), 267–270.

Menninghaus, W., Bohrn, I., Knoop, C. A., Kotz, S., Schlotz, W., & Jacobs, A. M. (2015). Rhetorical features facilitate prosodic processing while handicapping ease of semantic comprehension. *Cognition*, *143*, 48–60. http://dx.doi.org/10.1016/j.cognition.2015.05.026.

Miall, D. S., & Kuiken, D. (1995). Aspects of literary response: A new questionnaire. *Research in the Teaching of English*, *29*, 37–58.

Morreall, J. (1983). Humor and emotion. *American Philosophical Quarterly*, *20*(3), 297–304.

Nagels, A., Kauschke, C., Schrauf, J., Whitney, C., Straube, B., & Kircher, T. (2013). Neural substrates of figurative language during natural speech perception: an fMRI study. *Frontiers in Behavioral Neuroscience*, *7*(121) . http://dx.doi.org/10.3389/fnbeh.2013.00121.

Nietzsche, F. (1993 [original 1872]). *The Birth of Tragedy Out of the Spirit of Music*. Ed. M. Tanner. (S. Whiteside, Trans.) London: Penguin.

Oatley, K. (1999). Why fiction may be twice as true as fact: fiction as cognitive and emotional simulation. *Review of General Psychology*, *3*(2), 101–117. http://dx.doi.org/10.1037/a0037999.

Obermeier, C., Menninghaus, W., von Koppenfels, M., Raettig, T., Schmidt-Kassow, M., Otterbein, S., et al. (2013). Aesthetic and emotional effects of meter and rhyme in poetry. *Frontiers in Psychology*, *4*(10) . http://dx.doi.org/10.3389/fpsyg.2013.00010.

Ostermann, E. (1997). Das Interessante als Element ästhetischer Authentizität. In J. Berg, H.-O. Hügel, & H. Kurzenberger (Eds.), *Authentizität als Darstellung* (pp. 197–215).Hildesheim: Universitätsverlag.

Picon, G. (1953). *L'écrivain et son ombre: introduction à une esthétique de la littérature*. Paris: Gallimard.

Quintilian (1953). *The Instituto oratoria of Quintilian*. (H. E. Butler, Ed. and trans.) London: W. Heinemann/Cambridge, MA: Harvard University Press.

Real, R., & Vargas, J. M. (1996). The probabilistic basis of Jaccard's Index of Similarity. *Systematic Biology*, *45*(3), 380–385. http://dx.doi.org/10.2307/2413572.

Regel, S., Gunter, T. C., & Friederici, A. D. (2011). Isn't it ironic? An electrophysiological exploration of figurative language processing. *Journal of Cognitive Neuroscience*, *23*(2), 277–293. http://dx.doi.org/10.1162/jocn.2010.21411.

Ribeiro, A. C. (2012). Aesthetic attributions: the case of poetry. *The Journal of Aesthetics and Art Criticism*, *70*(3), 293–302. http://dx.doi.org/10.1111/j.1540-6245.2012.01521.x.

Schlaffer, H. (2012). *Geistersprache: Zweck und Mittel der Lyrik*. Munich: Hanser.

Schlegel, F. (2001 [original 1797]). *On the Study of Greek Poetry*. (S. Barnett, Trans.) Albany: State University of New York Press.

Schmidt, R. (2007). Der literarische Kanon: Ein Organ des Willens zur Macht oder des Gewinns an Kompetenzen? In N. Saul, & R. Schmidt (Eds.), *Literarische Wertung und Kanonbildung* (pp. 9–21).Würzburg: Königshausen und Neumann.

Shetley, V. (1993). *After the death of poetry: poet and audience in contemporary America*. Durham: Duke University Press.

Silvia, P. J. (2010). Confusion and interest: the role of knowledge emotions in aesthetic experience. *Psychology of Aesthetics, Creativity and the Arts*, *4*(2), 75–80. http://dx.doi.org/10.1037/a0017081.

Silvia, P. J. (2005). What is interesting? Exploring the appraisal structure of interest. *Emotion*, *5*(1), 89–102. http://dx.doi.org/10.1037/1528-3542.5.1.89.

Spoerhase, C. (2014). Das Maß der Potsdamer Garde: Die ästhetische Vorgeschichte des Rankings in der europäischen Literatur- und Kunstkritik des 18: Jahrhunderts. *Jahrbuch der Deutschen Schillergesellschaft*, *58*, 90–126.

Sulzer, J.-G. (1967 [original 1773]). *Allgemeine Theorie der schönen Künste*, 2. Hildesheim: Olms.

Sutrop, U. (2001). List task and a cognitive salience index. *Field Methods*, *13*(3), 263–276. http://dx.doi.org/10.1177/1525822X0101300303.

Taruffi, L., & Koelsch, S. (2014). The paradox of music-evoked sadness: an online survey. *PLoS One*, *9*(10) . http://dx.doi.org/10.1371/journal.pone.0110490 e110490.

Taylor, S. (1988). 'Vis comica' and comic vices: catharsis and morality in comedy. *Forum for Modern Language Studies*, *24*, 321–331.

van Goozen, S., & Frijda, N. H. (1993). Emotion words used in 6 European countries. *European Journal of Social Psychology*, *23*(1), 89–95. http://dx.doi.org/10.1002/ejsp.2420230108.

von Heydebrand, R., & Winko, S. (1996). *Einführung in die Wertung von Literatur: Systematik – Geschichte – Legitimation*. Paderborn: Schöningh.

von Schelling, F.W. J. (1989 [original 1802]). *The Philosophy of Art*. (D.W. Stott, Trans.) Minneapolis: Minnesota University Press.

Suspense: conceptualizations, theoretical analyses, and empirical explorations. In P. Vorderer, H. J. Wulff, & M. Friedrichsen (Eds.). London: Routledge.

Wassiliwizky, E., Wagner, V., Jacobsen, T., Menninghaus, W., et al. (2015). Art-elicited chills indicate states of being moved. *Psychology of Aesthetics, Creativity, and the Arts*, *9*(4), 405–416. http://dx.doi.org/10.1037/aca0000023.

Weninger, R. (1994). *Literarische Konventionen: Theoretische Modelle, Historische Anwendung*. Tübingen: Stauffenburg.

Weyrauch, W. (1989). Nachwort, In W. Weyrauch (Ed.), *Tausend Gramm: Ein deutsches Bekenntnis in dreißig Geschichten aus dem Jahr 1949* (pp. 175–183). Revised ed. Reinbek: Rowohlt.

Wortschatzprojekt Universität Leipzig (2015). *Wortschatzportal*. Retrieved from: http://wortschatz.uni-leipzig.de.

Wulff, H. J. (1996). Suspense and the influence of cataphora on viewers' expectations. In P. Vorderer, H. J. Wulff, & M. Friedrichsen (Eds.), *Suspense: conceptualizations, theoretical analyses, and empirical explorations* (pp. 1–17).London: Routledge.

Yanal, R. J. (1996). The paradox of suspense. *British Journal of Aesthetics*, *36*(2), 146–158.

Zentner, M., Grandjean, D., & Scherer, K. R. (2008). Emotions evoked by the sound of music: characterization, classification, and measurement. *Emotion*, *8*(4), 494–521. http://dx.doi.org/10.1037/1528-3542.8.4.494.

Zillmann, D. (1980). Anatomy of suspense, In: P. H. Tannenbaum (ed.). *The Entertainment Functions of Television*. Hillsdale, NJ: Erlbaum, 133–163.

Zillmann, D. (1983). Disparagement humor. In P. E. McGhee, & J. H. Goldstein (Eds.), *Handbook of humor research, 1: basic issues* (pp. 85–107).New York: Springer.

Zwaan, R. A. (1994). Effect of genre expectations on text comprehension. *Journal of Experimental Psychology—Learning Memory and Cognition*, *20*, 920–933.

Christine A. Knoop is a Senior Research Fellow at the Max Planck Institute for Empirical Aesthetics (MPIAE) in Frankfurt. She holds a Ph.D. in Comparative Literature from University College London, and taught and conducted research at UCL and Freie Universität Berlin prior to joining the MPIAE. Her main research interests include experimental approaches to literary aesthetics, aesthetic emotion, and authorship theories.

Valentin Wagner is a Senior Research Fellow at the Max-Planck-Institute for Empirical Aesthetics in Frankfurt. He holds a Ph.D. in psychology (University of Leipzig, Germany) and has been working in the fields of language psychology, emotion psychology, and empirical aesthetics.

Thomas Jacobsen is Professor of Experimental and Biological Psychology at Helmut-Schmidt-University/UniBw Hamburg. He received his degree in Psychology (Diplom-Psychologe) from Freie Universität Berlin in 1994. He was visiting scholar in Cognitive Neuroscience at UCSD and pre-doctoral researcher at the MPI of Cognitive Neuroscience. He obtained his PhD in Psychology in 2000 from the University of Leipzig. There he became Assistant Professor, and, after his Habilitation in 2004, Associate Professor. He was a Visiting Professor at the University of Vienna and the Freie Universität Berlin. In 2009, he took his current position as Professor of Psychology in Hamburg.

Winfried Menninghaus is director of the Department of Language and Literature of the Max Planck Institute for Empirical Aesthetics (Frankfurt am Main). He is a member of the Berlin-Brandenburgische Akademie der Wissenschaften, and has worked as a visiting Professor at the universities of Jerusalem, Berkeley, Yale, Princeton, Rice, and the EHESS Paris. *Fields of research*: classical rhetoric and poetics; philosophical, evolutionary and empirical aesthetics; literature from 1750 until present. *Most recent empirical study*: Menninghaus W., Bohrn I., Knoop C., Kotz S., Schlotz W., Jacobs A. (2015). Rhetorical features facilitate prosodic processing while handicapping ease of semantic comprehension. *Cognition*. doi.org/10.1016/j.cognition.2015.05.026.