





# What's that Scary Sound? Ambient Sound in Gothic Fiction

Svenja Guhr<sup>1</sup>   
Mark Algee-Hewitt<sup>2</sup> 

1. Institute of Linguistics and Literary Studies, Technical University of Darmstadt , Darmstadt, Germany.
2. English Department, Literary Lab, University of Stanford , Palo Alto, U.S.A..

## Citation

Svenja Guhr and Mark Algee-Hewitt (2023). "What's that Scary Sound? Ambient Sound in Gothic Fiction". In: *Journal of Computational Literary Studies* 2 (1). [10.48694/jcls.3583](https://doi.org/10.48694/jcls.3583)

**Date published** 2024-03-24

**Date accepted** 2024-02-02

**Date received** 2023-01-31

## Keywords

sound studies, ambient sound, Gothic fiction, 19<sup>th</sup> century, literary prose, English

## License

CC BY 4.0 

## Reviewers

Julian Schröter, Simon Kroll

## Note

This paper has passed through the conference track of JCLS. In addition to being peer reviewed, it was presented and discussed at the 2nd Annual Conference of Computational Literary Studies at Würzburg University in June 2023.

**Abstract.** This paper presents an approach to operationalizing ambient sound as a literary phenomenon. To illustrate the importance of the ambient soundscape in literary studies, we both manually and automatically detect ambient sound markers and use these annotations to analyze a sample of nineteenth-century English novels and short stories. Our hypothesis is that descriptions of a story's ambient soundscape can be associated with specific genres, and is, for example, a hallmark of Gothic novels. We use a classification approach based on a state-of-the-art transfer learning algorithm and a domain-dependent fine-tuned BERT model for English to automatically detect word-level sound indicators and compare their occurrence over the course of the fiction and with a comparative view on our corpus texts.

## 1. Introduction

"The rising **blast sighed** through the towering pines, which rose loftily above Matilda's head: the distant **thunder, hoarse** as the **murmurs** of the grove, in indistinct **echoes** mingled with the hollow **breeze**; the **scintillating lightning flashed** incessantly across her path, as Matilda, heeding not the **storm**, advanced along the trackless forest. [...] The **battling elements paused**: an **uninterrupted silence, deep, dreadful** as the **silence of the tomb**, succeeded. Matilda **heard a noise – footsteps** were distinguishable, and looking up, a flash of vivid lightning disclosed to her view the towering form of Zastrozzi." (Shelley, *P. Zastrozzi*; emphasis by authors)

Thunder, lightning, breezes, and echoes: Percy Shelley's narrator describes the contesting elements of nature that surround the protagonist Matilda in his Gothic novel *Zastrozzi*. Through these representations of sensation, the reader vicariously shares Matilda's sensory experiences: seeing the lightning, hearing the thunder, feeling the breeze. Many of these descriptions provide the basis for the soundscape of the scene.

In this paper, we explore the function of sound in literary fiction. Applying the field of sound studies to literary analysis, we investigate the representation of ambient sound in fiction with a focus on its use in 19<sup>th</sup> century British Gothic novels and short stories.

In this paper, we adopt a capacious understanding of the Gothic: not just including the central Gothic novels of the late 18<sup>th</sup> or early 19<sup>th</sup> century, but broadening our reach to

include Gothic inflected fiction of the later 19<sup>th</sup> century that make use of the tropes of the genre. Our choice of corpus texts was, in part, guided by previous studies of the Gothic, including the “writers of Gothic” mentioned in *The Handbook of the Gothic* edited by Mulvey-Roberts (2009).

Our work proceeds from the hypothesis that detailed descriptions of the story’s ambient soundscape (e.g., the growl of a wild animal, the creaking of a wooden floor) is a trope of the Gothic and we pay particular attention to sounds at either end of the loudness spectrum – from deep silence to loud screams and clashing thunder. This article offers a new approach to operationalizing sound on the word level and introduces methods to manually and automatically detect ambient sound markers in English literary prose.

Our approach is as follows: First, we offer insight into the subfield of literary sound studies (see [subsection 2.1](#)), and discuss its utility in analyzing Gothic literature (see [subsection 2.2](#)). We then describe our methods for operationalizing sound and detail the method we used to analyze it, drawing from manually and automatically generated annotations of a selected corpus of 19<sup>th</sup> century English novels and short stories. The methods we adopt for the analysis of sound include a dictionary approach to define a baseline for our analysis (see [subsubsection 3.3.1](#)) as well as a transfer learning approach using the *NEISS TEI Entity Enricher* (Zöllner et al. 2021) (see [subsection 3.3](#)). Our ability to accurately detect sound words automatically is evaluated through comparisons against manually annotated data. In [section 4](#) and [section 5](#), we examine sound references across a corpus of novels and short stories with particular interest in passages with a high density of sound words that contain particularly loud or low sound indications. Overall, our study contributes to the burgeoning research field of literary sound studies, connecting it to a rising interest in the operationalization of sensual experiences as we analyze sound in fictional prose using current distant reading methods.

## 2. Theoretical Background

### 2.1 Sound and Literary Studies

There is a great diversity of approaches towards sound in literary studies. On the one hand, sound can be analyzed as it is actually produced during the oral recitation of literary texts, where pronunciation, stress, pauses, speech rhythm can be studied phonologically (Blohm et al. 2021). In their book chapter “Sound Shape and Sound Effects of Literary Texts” as part of the *Handbook of Empirical Literary Studies*, Blohm et al. (2021) describe such an empirical approach to sound in literature. They claim that “reading a word automatically activates its abstract sound representation [...] referred to as phonological recoding [...] of written text” and “experience[d] as an ‘inner voice’” (Blohm et al. 2021, 11). On the other hand, onomatopoeia, alliterations and rhymes within the text can be analyzed stylistically for the ways in which they form the sound design of literary texts, shaping the flow of reading whether aloud or silent (Hinton et al. 1995, 1–10).

Both of these approaches are reception-oriented and define sound as physical vibrations that transmit information. In literary texts, however, a third type of analysis is made possible by studying the representation of sounds, and their associated soundscapes,

within the fictional world itself. We see this, for example, in Schafer (1994) who analyzes diachronically changing soundscapes (composite of ‘sound’ and ‘landscape’). In addition to analyzing real world sonic environments as he did in his *World Soundscape Project*, Schafer (1994) also reads soundscapes through literary texts, claiming that “writers of fiction were reliable ‘earwitnesses’ whose writings ‘constitute the best guide available in the reconstruction of soundscapes past’ ” (Schafer 1994, 9 in Snaith 2020, 20).

There have been numerous recent attempts in Literary Studies to analyze these fictional soundscapes. These often focus on descriptions of sound which aid the reader’s imaginative capacity for experiencing the fictional world. For audionarratologists, the sound imagination is related to the reader’s own memories of described sounds that “create the details of a storyworld in our heads [as] a ‘theatre of the mind’ ” (Verma 2012, Mildorf 2019, 297). Similarly, Picker (2003) analyzes sounds in English literature through Dickens’s soundscape descriptions of his fictionalized London. More recent examples can be found in the essay collection *Literature and Sound* (Snaith 2020).

The most apparent and easily accessible sounds in fiction, according to the soundscape analysis of Döblin’s *Alexanderplatz* by Bernhart (2008, 61), are those that occur in dialogue. Character utterances and their descriptors are both the most frequent and most explicit depictions of sound. Ambient sounds (e.g., those related to machinery or nature) are less often explicitly mentioned (Bernhart 2008, 62). Depictions of sensory experience evidence a hierarchy of sensation: descriptions of settings in literary texts focus almost exclusively on sight, e.g., describing a landscape or a living-room. The depiction of sound beyond dialogue is rare. Narrative seems focused on directing the imaginative gaze of the reader’s eyes rather than directing the reader’s mental ear with imaginative sounds (Smith 2015, 27–37). Nevertheless, the representation of sound can play a crucial role in aiding the reader’s understanding and intuitive experience of a narrative.

Narrative draws on a reader’s prior experience to supplement its descriptive techniques. For example, in describing a train entering the station, the moving train is the focus; however, the narrator’s description of the world is incomplete, relying on the reader’s world knowledge to fill in the missing information about the furnishings of the station building or the color and model of the arriving train. The description of the station’s soundscape is often similarly omitted. When sounds are described, it is because they differ from the expected soundscape. As an example, Bernhart (2008, 61–62) emphasizes the sound of bells ringing in Döblin’s *Alexanderplatz* by giving an explicit description of sounds that are not typically part of the reader’s default experience of the setting.

## 2.2 Sound in Gothic Fiction

In Gothic fiction, readers are often placed in unfamiliar, or uncanny, settings that deviate from the default. In its “decaying Gothic castles, ruined chapels, underground passages, dark forests and ghostly groaning” and for its “[s]hocks, supernatural incidents and superstitious beliefs”, all of which “promote a sense of sublime awe and wonder [...] entwined with fear and elevated imaginations” (Botting 1996, 29, 46), the Gothic sets itself apart from the realistic literature of the 19<sup>th</sup> century in favor of the sensational, fantastic, and the uncanny (Bacon 2018, 1, Hurley 2002, 191).

Stylistically, Gothic novels deploy a vocabulary of mystery, uncertainty, terror, horror, fear, and “hyperbolic language [...] [which] attempts to create a brooding, suspenseful atmosphere” (Hurley 2002, 191). In addition to its ‘ghosts’, ‘phantoms’, and ‘wretches’, depictions of non-human behavior, supernatural forces, or inexplicable events as in quotation (1) work to create the sensation of mystery:

(1) “As I said this I suddenly beheld the figure of a man, at some distance, advancing towards me with **superhuman speed**. He bounded over the crevices in the ice, among which I had walked with caution; his stature, also, as he approached, **seemed to exceed that of man.**” (Shelley *Frankenstein*; emphasis by authors)

Such uncertainty is enhanced by the conditional phrases and questions (rhetorical or actual) voiced by characters (“I figured to myself”, “I wondered if”, “he might”). Characters, and by extension, readers, often seek to alleviate this uncertainty by paying attention to details in their environment, as in quotation (2).

(2) “I preternaturally listened; I figured to myself **what might** portentously be; I **wondered if** his bed were also empty and he too were **secretly** at watch. It was a deep, soundless minute, at the end of which my impulse failed. He was quiet; he **might** be innocent; the risk was hideous; I turned away.” (James *The Turn of the Screw*; emphasis by authors)

In addition to their environmental details, Gothic texts also build their fictional worlds through the depiction of “emotions [...] by detailing the protagonist’s thoughts and feelings” (Ellis 2000, 9) as well as through sensory experiences described by the narrator (sight, smell, taste, touch, hearing) as in quotation (3).

(3) “In a few minutes after, I **heard** the creaking of my door, as if some one endeavoured to open it softly. I **trembled** from head to foot; I **felt** a presentiment of who it was and wished to rouse one of the peasants who dwelt in a cottage not far from mine; but I was overcome by the sensation of helplessness, so often **felt** in frightful dreams, when you in vain endeavour to fly from an impending danger, and was rooted to the spot. Presently I **heard** the sound of footsteps along the passage; the door opened, and the wretch whom I dreaded appeared.” (Shelley *Frankenstein*; emphasis by authors)

Quotation (3) gives just one instance of the importance of auditory descriptions in the Gothic. In his monograph on *Gothic Voices: The Vococentric Soundworld of Gothic Writing*, Foley (2023) discusses how the Gothic atmosphere relies on the soundscape to generate horror or suspense: “creaking floorboards, howling winds and thunder rolling are just some of the acoustic motifs that alert us to a Gothic atmosphere” (Foley 2023, 1).

Mysterious sounds and deep silences frequently occur within the Gothic soundscape (Glotova 2021, 1), and are represented through hearing events<sup>1</sup> that can be signaled by the verbs ‘listen’ or ‘hear’ as in quotation (3) or by sound words as ‘scream’, ‘burst’, ‘cry’, ‘yell’, in quotation (4).

1. By using the term ‘event’ for narratological segments, we refer to the *event I* definition by Hühn (2013), Gius and Vauth (2022, 3): “event I is any change of state and thus a general type of event without further requirements”.



(4) “A terrible **scream** – a prolonged **yell** of horror and anguish – **burst** out of the **silence** of the moor. That frightful **cry** turned the blood to ice in my veins.” (Doyle *The Hound of the Baskervilles*; emphasis by authors)

In quotation (4), we can also observe an oscillation between especially loud and quiet sound descriptions that echoes Hurley (2002, 9)’s observation of Gothic fiction’s continuous “confrontations between the low and the high [...] [or] other opposed conditions – including life/death, natural/supernatural, ancient/modern, realistic/artificial, and unconscious/conscious”.

Later in this paper, we will show that Gothic texts often depict extended periods of silence which are interrupted by sudden, often loud, sounds (see [subsection 5.1](#)). Quotation (5) offers an overview of these Gothic representational tropes functioning together:

(5) “It’s eleven o’clock **striking** by the **bell** of Saint Paul’s. **Listen** and you’ll **hear** all the **bells** in the city **jangling**. Both sit **silent**, **listening** to the metal **voices**, near and distant, **resounding** from towers of various heights, in **tones** more various than their situations. When these at length **cease**, all **seems** more **mysterious** and **quiet** than before. One disagreeable result of **whispering** is that it **seems** to evoke an atmosphere of **silence**, **haunted** by the **ghosts** of **sound**, **strange cracks** and **tickings**, the **rustling** of garments that have no substance in them, and the tread of **dreadful** feet that would leave no mark on the sea-sand or the winter snow. So **sensitive** the two friends happen to be that the air is full of these **phantoms**, and the two **look** over their shoulders by one consent to **see** that the door is shut.” (Dickens *Bleak House*; emphasis by authors)

The sensory experiences of the two characters are described (“listen/ing”, “see”, “look”). Known (“bell of Saint Paul’s”) and unknown (“strange cracks”) sounds interrupt the silent ambiance of the city scene (“atmosphere of silence”). A vocabulary of mystery (“mysterious”, “ghosts”, “phantoms”) triggers an atmosphere of uncertainty (“seems”) and fear, prompting the reader’s desire to know what happens next.

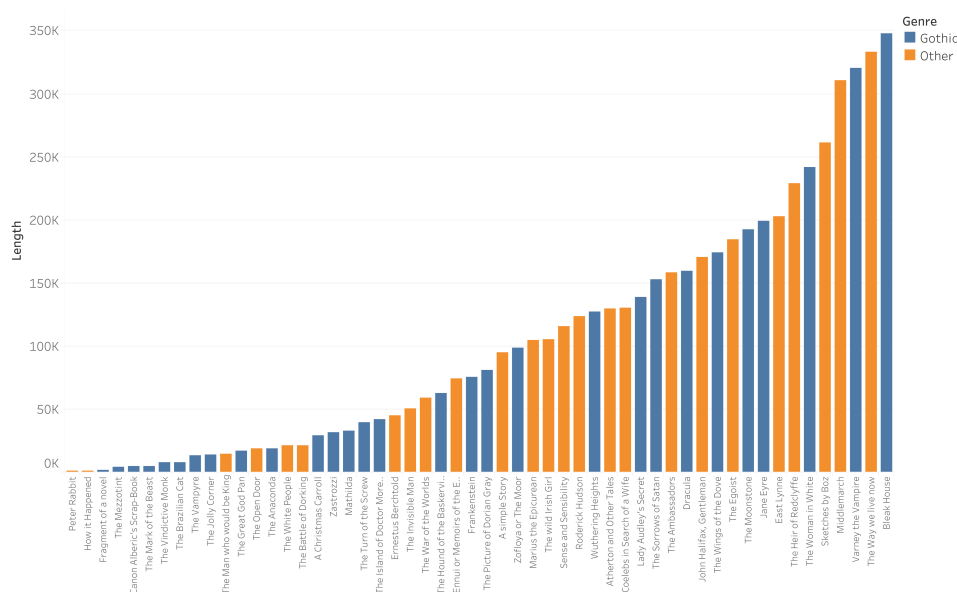
Our analysis will focus on these represented sounds as we show how their operationalization, through the systematic annotation and automated detection of sound indicators, can reveal new facets of the soundscape of the Gothic.

### 3. Method

Traditionally, research on sound in fiction has relied on close-reading. In our article, we present a distant reading approach as an alternative that can access disparate elements of a soundscape that are invisible to even the most careful reader. Our analysis of ambient sound is based on a corpus of 19<sup>th</sup> century literary texts. We analyzed this corpus through a combination of standard computational literary studies methods, namely manual and semi-manual annotation (Horstmann 2020) as well as automated annotation using a Transfer Learning Named Entity Recognition approach (Zöllner et al. 2021) evaluated on manually annotated data.

Property	number
Number of texts	55
Number of texts labeled as Gothic/Gothic-themed	30
Other 19 <sup>th</sup> century British fictional prose texts	25
Number of texts written by female authors	20
Number of texts written by male authors	35
Shortest text (in words)	988
Longest text (in words)	348,079
Texts manually sound annotated	14
Texts dictionary-based sound annotated (corrected false positives)	7
Texts automatically annotated for sound	36

**Table 1:** Corpus Description.



**Figure 1:** Length of the corpus texts with different colors for Gothic or other genre texts.

### 3.1 The Research Corpus

For our corpus, we selected 55 texts of different length (short stories, novellas and novels) based on a selection of 30 English novels and short stories that were, i.a., mentioned in *The Handbook of the Gothic* (Mulvey-Roberts 2009) and supplemented by 25 canonical works of 19<sup>th</sup> century English fiction.

The corpus texts (original texts as well as in-line sound annotated texts) are accessible as plain TXT files (in UTF-8) and XML files with TEI annotations (TEI Consortium 2022). Additionally, we provide a metadata table containing information on, e.g., text name, author name, author gender, publication year, text length in words, file names, annotation status, and more). For some statistical information about the corpus, see [Table 1](#).

### 3.2 Operationalizing Ambient Sound

To operationalize the phenomenon of ambient sound in literature, we adopted the proven procedure of reflective text analysis of Pichler and Reiter (2020). Through an

iterative manual approach, we systematically annotated ambient sound at the word level, as a lexical unit.

We distinguished between implicit and explicit sound indicators. Explicit sound descriptions are concrete: there is detailed information about the sound present in a scene which is represented through a semantically loaded sound word. Implicit sound description relies on the reader's interpretation of a description of an event. For example:

**a) Implicit Sound Description:**

(6) "The train is entering the station."

As experienced readers who have heard a train entering a station, we know that the arrival of the train is noisy. Nevertheless, in this sentence, no sound is annotated because it is not explicitly described with a lexical unit.

**b) Explicit Sound Description:**

The lexical unit becomes an explicit sound description when, for example, the action verb 'enter' is exchanged for the sound-indicating verb 'rattle', as in the sentence:

(7) "The train rattles into the station."

The rattling sound of the arriving train is explicitly indicated through the text's vocabulary such that the sound can be attached to a lexical unit – here the verb 'rattle'. We consider this word an annotation unit. An annotation in TEI would be:

(8) "The train <sound>rattles</sound> into the station."

**Particular Annotation Cases**

Human sounds can also be part of the ambient soundscape. This is the case, for example, when the scream of a woman is depicted, or when the sound of a crowd singing or rumbling is mentioned. In these cases, the human-made sound is not explicitly communicating information through language: It does not convey a verbal message of an identifiable speaker to a specific addressee.

Some ambient sound depictions are not annotated. For example, sound descriptions referring to iterative events, generalizations and regularities, or references to sounds realized in the past are not included into the annotation (e.g., "the bells always ring at noon time" (generalization, regularity), "they often sang the Requiem at funeral services" (regularity, past). Here 'ring' and 'sang' do not indicate a represented sound in a particular scene). Consequently, only sounds that can be diegetically related to events in the fiction were tagged. This also excludes negated sounds (e.g., "the bell did not ring today"), as well as the articulation of wishes and conditional statements (e.g., "Oh, that some encouraging voice would answer in the affirmative!" (Shelley *Frankenstein*)). The situation is different, however, for often explicitly described silence (e.g., "there was a peaceful silence over the misty morning landscape"). Generally, silence is treated as absence of loud sounds resulting in a calm soundscape. Our decision to tag these passages does not mean that we simply felt that there were no sounds: rather the language of text flags the complete absence of sounds by referring explicitly to silence or to sounds that occur at an imperceptible volume.

**Manual Annotation**

On the basis of this method for operationalizing ambient sound, we formulated annotation guidelines for the manual annotation of ca. 25% of our corpus. Three annotators

author	year	title	words	sw	swd	man/dic
Brontë	1847	<i>Jane Eyre</i>	188,598	604	0.32	dic
Brontë	1847	<i>Wuthering Heights</i>	119,475	351	0.29	dic
Byron	1819	<i>Fragment of a Novel</i>	1,977	1	0.05	man
Doyle	1898	<i>The Brazilian Cat</i>	8,148	65	0.80	man
Dickens	1848	<i>A Christmas Carol</i>	29,243	194	0.66	man
Gaskell	1852	<i>The Old Nurse's Story</i>	9,805	66	0.67	man
M.R. James	1895	<i>Canon Alberic's Scrap-Book</i>	4,716	20	0.42	man
M.R. James	1904	<i>The Mezzotint</i>	4,682	0	0.0	man
Kipling	1890	<i>The Mark of the Beast</i>	5,109	37	0.72	dic
Lewis	1808	<i>The Anaconda</i>	18,996	75	0.39	man
Oliphant	1881	<i>The Open Door</i>	18,763	161	0.86	dic
Potter	1902	<i>The Tale of Peter Rabbit</i>	981	10	1.02	man
Shelley, P.	1818	<i>Zastrozzi</i>	30,971	229	0.74	dic
Trollope	1875	<i>The Way we live now</i> (Ch.1–10)	35,895	12	0.03	man
Wells	1897	<i>The Invisible Man</i>	49,808	385	0.77	dic
Wilde	1891	<i>The Picture of Dorian Gray</i>	80,396	288	0.36	dic
Yonge	1853	<i>The Heir of Redclyffe</i> (Ch.1–10)	59,774	61	0.10	man
<b>total</b>			<b>1,323,574</b>	<b>4,139</b>	<b>0.31</b>	

**Table 2:** Texts of the different training sets: Manually annotated or dictionary-based annotated with manual false positive correction. Indicated are the total number of words, the number of sound words (sw), the calculated swd, and how it was annotated.

author	year	title	words	sw	swd
Crookenden	1802	<i>The Vindictive Monk</i>	7,672	16	0.21
Doyle	1913	<i>How it happened</i>	1,429	11	0.77
Doyle	1902	<i>The Hound of the Baskervilles</i>	59,931	125	0.21
Shelley, M.	1818	<i>Frankenstein</i>	75,235	254	0.34
<b>total</b>			<b>144,267</b>	<b>406</b>	<b>0.28</b>

**Table 3:** Test Set : Manually annotated corpus texts. Indicated are the total number of words, the number of sound words (sw), and the calculated swd.

(trained in both literary studies and annotations) manually annotated a total of 14 texts of varying length from the corpus following the *Guidelines for Ambient Sound Annotation* (Guhr 2023). The annotation guidelines were developed following an iterative process according to Reiter (2020). In a manual annotation of Lewis' *The Anaconda* by two annotators, we obtained an inter-annotator-agreement of 0.80 Cohen's *kappa* (Scikit-learn 2022), which is considered to be a decent agreement for a manual annotation task of literary phenomena but also indicates the complexity of this task for human readers. Our test set contains four of the 14 manually annotated texts.

### 3.3 Approaches for Automatizing the Annotation of Ambient Sound

In order to automate the annotation of ambient sound descriptors, we compared two approaches: a simple dictionary approach that consequently served as the baseline for automated annotation (see subsection 3.3.1), and a classification approach based on a state-of-the-art transfer learning algorithm and a BERT language model (see subsection 3.3.2).

### 3.3.1 Dictionary Approach

To determine a baseline for automated ambient sound annotation, we adopted a simple dictionary approach. After lemmatizing the manually annotated training texts (see [Table 2](#)) using the NLTK (Loper and Bird 2002), we extracted the unique sound word lemmas. We then took these lemmas and found matches to them in a lemmatized set of texts resulting in a dictionary with a key-value pair {'lemma' : 'sound annotation'}. After each of three rounds of annotations, the sound word list was refined based on discussions among the annotators and the guidelines were updated. In each new round, a smaller list of sound words were extracted. Starting from 289 sound words in the first round, only 258 sound word lemmas were left in the second round, and only 228 in the third round.

#### Error Analysis of the Dictionary Approach

We used this dictionary approach to automatically annotate Doyle's short story *How it Happened*, comparing the results with our manual annotation of a total of 11 sound words. The evaluation results can be found in [Table 5](#). Comparing the first round (19 false positives, 2 false negatives) to the second round (12 false positives, 2 false negatives) and the third round (2 false positives, 9 false negatives), we see a decrease in false positives as the sound word list is revised; however we also see a rise in false negatives. The dictionary approach therefore has high accuracy and recall due to the generalization of the annotation process that tends to be oversensitive to words that could lexically be sound words but do not indicate actual sounds in the diegesis or are homographs of non-sound words (see *Particular Annotation Cases* in [subsection 3.2](#)).<sup>2</sup> Looking at the false negatives across all rounds, which consequently were not part of the sound word lists, we can see the genre and time span dependence of our dictionary approach. For example, in all rounds, "whir" was left out in the annotation, which is not treated explicitly in any other of the 19<sup>th</sup> century training texts.

In summary, the dictionary approach is useful to detect mentions of sound words on the lexical level; however, it does not take context into account, resulting in a high number of false positives.

### 3.3.2 Classification with NEISS NTEE

To get context sensitive annotations of ambient sound words in our corpus, we adopted a classification approach based on a state-of-the-art transfer learning algorithm and a BERT language model from *NEISS TEI Entity Enricher* by Zöllner et al. (2021).

In our approach, we followed the findings from earlier studies on generalized named entity recognition for the detection of abstract entities such as places and spaces or character gender in German language novels (Flüh et al. 2022; Schumacher 2022). Both approaches employed the open access and open source software *Stanford Named Entity Recognizer* (*StanfordNER*) (Manning et al. 2014), that was originally trained to detect named entities in a narrow sense, namely, names of people, organizations or places,

2. To evaluate our results, we use the following metrics commonly used in computational literary studies: accuracy, precision, recall. Furthermore, in line with Zöllner et al. (2021), we calculate F1-score as well as E-F1-score to distinguish between entity-wise F1-score (E-F1) based on the *seqeval* Python framework's Hironsan (2018), which "evaluates a complete entity as true positive only if all tokens belonging to the entity are correct", and the more commonly used token-wise F1-score (T-F1, here just 'F1').

sound word (lemma)	absolute frequency in training set
sound	44
silence	18
cry	17
voice	15
wept/weep	19
silent	14
loud	13
thunder	10
step	10
scream	10
groan	8
calm	8
stillness	6

**Table 4:** The 13 most frequent sound words appearing in the training set.

	1 <sup>st</sup> round	2 <sup>nd</sup> round	3 <sup>rd</sup> round
sw lemmas	289	258	228
accuracy	0.98	0.99	0.99
precision	0.32	0.43	0.5
recall	0.81	0.81	0.82
F1-score	0.46	0.56	0.62

**Table 5:** Table with evaluation results of the dictionary approach (baseline).

using a conditional random field algorithm (Finkel et al. 2005). They then fine-tuned the model on manually annotated data (Schumacher 2022, 79–93).

Recent approaches to entity detection use the software *NEISS TEI Entity Enricher* (Zöllner et al. 2021) to fine-tune pre-trained models with manual annotations (Schumacher et al. 2022). Based on a transfer learning (Kamath et al. 2019) approach, the tool provides access to large-scale language models like the *Bidirectional Encoder Representations from Transformers (BERT)* architecture by *Hugging Face* (Devlin et al. 2018). Using this method, Flüh and Lemke (2022) were able to recognize named entities in German language letters from the 19<sup>th</sup> and 20<sup>th</sup> century.

In our study, we used the pre-trained BERT model provided in the software *NEISS NTEE* (originally the English language model (*bert-base-cased*) by *Hugging Face*; Devlin et al. 2018) and fine-tuned it with our manual annotations of ambient sound words, see subsection 3.2. For the prediction step, the software takes a non-labeled literary text in XML-format as input and automatically annotates it with the XML-tag <sound>TOKEN</sound>.

After each of five training round evaluations, we adapted the training data and tried different combinations of manually sound-annotated texts. As part of the ground truth building in *NTEE*, the training data was split into training, validation, and test data. Furthermore, one advantage of *NTEE* is the “Shuffle By Sentence” option to train and evaluate the performance of the trained entity tagger independently of the text type (novel or short story), using sentence-by-sentence training and evaluation.



Presumably, a set of meaningful segments (like events or scenes) would be more useful for shuffling over the data than shuffling sentences (Sperfeld and Lemke 2022); however, such automatic segmentation is not yet advanced enough to be integrated into the software training process (e.g., Zehe et al. 2021).

### Combining the Dictionary Approach with the Automatic Prediction

As we saw in the training step, the prediction performance (entity-wise F1-score, E-F1 for short) rose with the amount of training data; however, manual human annotation is a costly task (Gühr and Gius 2023). We therefore integrated our dictionary approach to aid in generating new training data: we first calculated the frequency of sound words for each text in the corpus and then we selected seven additional texts with a high incidence of sound words for annotation. We next used our dictionary approach to annotate these texts (see [subsection 3.3.1](#)) and manually corrected the resulting annotations by removing false positives.

Through an error analysis of these semi-automatically predicted sounds, we discovered the following groups of false positives:

1. Human communication that is not explicitly non-verbal or verbal:  
“crying to get free”.
2. Sound related to human communication: “louder voice”, or “chattering”.
3. Sounds related to human communication that were edge cases:  
“But while she hesitated what to do, she heard a <sound>voice</sound> at the door requesting admission”. In this sample, it is uncertain whether the ‘voice’ should be annotated as communication given the ambiguity of what is said or who says it.
4. Acceptable annotations missed by human annotators or edge cases:  
“trampled” is comparable to “stamping”.
5. Adjectives and adverbs that indicate properties of sounds:  
“A <sound>piercing</sound> <sound>shriek</sound> of horror <sound>burst</sound> from me!” In the sample “piercing” is a false positive, but could be annotated as sound-indicating property of “shriek”. One has to distinguish between properties relating to descriptive sound properties (namely, “loud”, “calm”) and judgmental properties without direct relation to the property of a sound like “beautiful”, “charming”, “violent” that indicates the perception of a given narrative perspective. In a revision of the annotation guidelines, we explicitly excluded non-sound-indicating properties.
6. Negated sounds: “she heaved not one sigh”.
7. Hypothetical sound, subjunctives, wishes:  
“I might have cried, but I didn’t.”, “I guess, she will cry.”, “I wish I could cry.”
8. Sounds in the past: “Last night I heard a woman screaming.”
9. Polysemy: ‘ring’ that can either be the sound of a bell or jewelry.

By focusing on the correction of false positives in a dictionary-annotated set, we were

training sets	texts	words	unl. words	sw	tr. ep.	b.ep.	E-F1
1	6	73,027	72,656	371	4	4	0.6753
2	11	285,745	284,088	1,657	11	8	<b>0.7403</b>
3	13	640,533	637,919	2,614	13	9	0.7007
4	13	640,531	638,103	2,428	12	10	<b>0.7157</b>
5	15	656,077	653,583	2,494	8	6	<b>0.6589</b>

**Table 6:** Table with evaluation results of the training rounds partly with added manually corrected dictionary-based annotated data. The abbreviations stand for: unl. = unlabeled, tr.ep. = training epoch, b.ep. = best epoch.

training sets	Precision	Recall	F1-score	E-F1
1	0.6226	0.75	0.6804	0.6804
2	0.4384	0.7272	0.5470	0.5470
3	0.5741	0.7045	0.6327	0.6327
4	<b>0.6538</b>	<b>0.7727</b>	<b>0.7083</b>	<b>0.7083</b>
5	0.5454	0.6818	0.6060	0.6061

**Table 7:** Table with test set evaluation results of the training partly with added manually corrected dictionary-based annotated data.

able to reduce the labor of manual annotation while still creating a more robust training set of true positives.

Adding these semi-annotated texts to the training set resulted in higher E-F1-scores on the split evaluation set and on the test set (see Table 6, training set 2 and 3). Training set 4 with 13 manually and semi-automatically annotated training texts received the best results on the split evaluation set (0.7157 E-F1) as well as on the test set (0.7083 F1) (see Table 7). However, the addition of two additional short stories to that training set (set 5) did not improve the evaluation results and so we elected to stop adding training data to our corpus.

**Evaluation of the Automatic Prediction** To evaluate the predictions, we used the provided evaluation option of the *NEISS NTEE* software, namely the sequence labeling evaluation entity-wise F1-score (E-F1) based on the *seqeval* Python framework by Hironson (2018), which “evaluates a complete entity as true positive only if all tokens belonging to the entity are correct”, in comparison to the more commonly used token-wise F1-score (T-F1, here just ‘F1’) (Zöllner et al. 2021, 6, 14). Examining the E-F1-score is especially useful when looking at entities that contain more than one token, as in the sound event “she <sound>screamed out</sound>”.

Using this framework, we calculated the E-F1-score first on the basis of the validation set (part of the split training set) and second on the chosen test set, additionally calculating precision, recall, and token-wise F1-score for further comparability (see Table 7). In comparison to the dictionary approach (see subsection 3.3.1), F1-score performance improved by 0.1 using this method. Looking at the F1-score and the E-F1-score calculations of training set 2 and 4, it is interesting to note that when given more annotated data, training set 4 received a lower E-F1-score on the split evaluation set and a higher F1-score on the independent test set. This contradiction, between the better performance of the model overall and its lower performance on the split training set, may be explained by the semi-manually annotated texts, which still contain false negatives (see Table 7).

### Error Analysis of the Automatic Prediction

Looking at the false positives and false negatives, the following five error groups were identified, with the errors highlighted in bold in the sample passages:

1. References to an event that happened in the past, especially in analepses as in this sample passage from *The Great God Pan* by Arthur Machen: “You must remember, Villiers, that I have seen this woman, in the ordinary adventure of London society, talking and **<sound>laughing</sound>**, and sipping her coffee in a commonplace drawing-room with commonplace people.” In this case, the sound described is not part of the actual scene, but a reported event from the past. Even if it adds to the soundscape of the novel overall, it is not an actual sound of the scene in which the event is reported.
2. Generalizations and sound descriptions in non-scenic narration like the false positives in the following sample passage from *Bleak House* by Charles Dickens: “One disagreeable result of whispering is that it seems to evoke an atmosphere of **<sound>silence</sound>**, haunted by the ghosts of **<sound>sound</sound>** – strange cracks and tickings, the **<sound>rustling</sound>** of garments that have no substance in them, and the tread of dreadful feet that would leave no mark on the sea-sand or the winter snow.” These annotations show the sensitivity of the model to sound descriptions, even when these sounds are not realized in their context, much like the generalizations in the non-scenic narration of the sample passage.
3. Negation of sound that is not taken into account: According to the guidelines, negated sound should not be annotated as “the wind did not blast”. Often, however, negation can be interpreted as an implicit indication of silence through the absence of sound, as in this sample passage from Mary Elizabeth Braddon’s *Lady Audley’s Secret*: “There was no **<sound>sound</sound>** but the **<sound>flapping</sound>** of the ivy-leaves against the glass, the occasional falling of a cinder, and the steady **<sound>ticking</sound>** of the clock.”
4. Rhetorical use of a sound event as a means of comparison, such as the comparison of the quiet atmosphere outside to the “quiet after an earthquake or storm” in the following sample passage from *The Great God Pan* by Arthur Machen. Here, “quiet” and “storm” are false positives according to the annotation guidelines: “The **<sound>noise</sound>** and **<sound>clamour</sound>** of the street had **<sound>died</sound>** away, though now and then the **<sound>sound</sound>** of **<sound>shouting</sound>** still came from the distance, and the dull, leaden **<sound>silence</sound>** seemed like the **<sound>quiet</sound>** after an earthquake or a **<sound>storm</sound>**. Villiers turned from the window and began speaking.”
5. Rhetorical use of sound words as metaphors as in the following sample passage from *The Wings Of The Dove* by Henry James: “Mrs. Stringham was now on the ground of thrilled recognitions, small **<sound>sharp</sound>** **<sound>echoes</sound>** of a past which she kept in a well-thumbed case, but which, on pressure of a spring and exposure to the air, still showed itself ticking as hard as an honest old watch.” It is important to observe that common metaphors, e.g., those referring to

verbs like “ticking” in the example, were correctly annotated as not a sound event, while less common metaphors like the “sharp echoes of a past” were detected as a sound event, ignoring their metaphorical nature. Future work will investigate whether the false-positive annotation of metaphors decreases with more training data.

Importantly, the errors in automatic prediction correspond to the difficulties encountered in manual annotation by human readers. We can conclude that the *NTEE* approach is more sensitive than manual annotation and produces many false positives that are still related to sounds, such as described sounds from the past or in non-scenic narration that are simply not part of the fictional soundscape of a particular scene in the fiction, but generalizations or reports of a soundscape of an unspecified moment in the fiction.

After completing the training and selecting the model with the best predictive performance according to the evaluation, we used the model (best epoch) to automatically annotate the remaining texts in the corpus.

### 3.3.3 Measurement of Sound Word Density

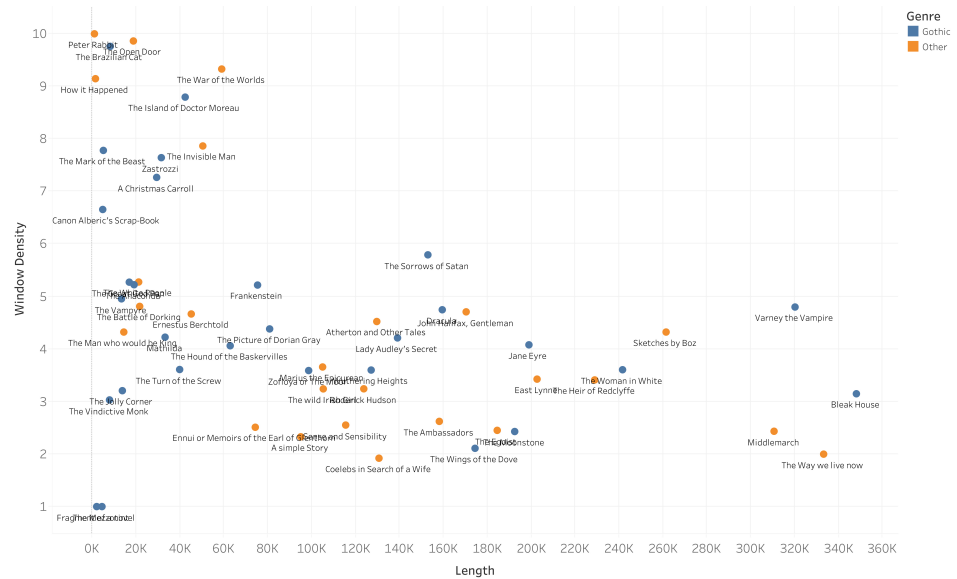
To compare the use of sound words across several texts, we adapted a method to measure a text’s sound word density. This measurement was developed in Guhr’s dissertation work to focus on character sounds and was based on comparable analyses by Schumacher (2022, 127).

The calculation normalizes occurrences over text length: the number of annotated sound words *sw* is divided by the total number of tokens *t*, and multiplied by 100. Sound word density (*swd*) scores can be found in Table 2 and Table 3 and visualized in Figure 2 and Figure 3.

$$\text{swd} = \frac{n_{sw}}{n_t} \cdot 100 \quad (1)$$

Using *swd* measurements, we can compare the tendency to represent ambient sound between, e.g., texts of different periods, genres, or authors. At finer scales, this measurement can also be used to compare individual segments of text.

To address a possible bias due to different text lengths, we additionally used a sliding window approach with windows of 1,000 tokens. For each window, we counted the number of sound words and then took the average of all windows with at least one sound word. Thus, we excluded the 1,000 token windows without any sound words in order to make novel-length texts and short stories comparable (working with the hypothesis that novels contain longer passages without sound words than short stories, biasing our *swd* calculation). The results of this approach, however, (Figure 2), show almost no difference from the results we obtained from the calculation without using a sliding window to address the different text lengths (see Figure 3 and Github repository). Based on our analysis, sound word density is not comparable to measurements like Type-Token-Ratio (which can be calculated by the sliding window approach) as we do



**Figure 2:** The dots show the texts and their average sound word density calculated for each window of 1,000 tokens.

not measure the diversity of sound words but the raw quantity of their appearance in the literary text.

As all of our analyses indicated that the sound word density is smaller in longer texts, in the following comparative analyses, we divide the texts into two length groups of either more or less than 100,000 words (the median text length of the corpus is about 102,000 words). Our results echo the findings of Algee-Hewitt et al. (2023, 354–355) on short stories, in that longer texts (novels) are fundamentally differently constructed than shorter fiction and are not merely reducible to a series of concatenated short story-like segments.

### 3.3.4 Loudness Level Labeling

In addition to measuring the density of the occurrence of sound words, our descriptors can also be used to approximate loudness levels within a given text. To generate these levels, two annotators manually annotated the 228 word dictionary of sound words with loudness levels as follows: 1 for non-audible sounds, e.g. *silence*; 2 for low sounds, e.g. *rustling*, 3 indicating normal indoor volume, e.g. *snoring*, 4 for loud sounds, e.g. *thunder*, *screaming*, *explosions*. Based on the inter-annotator-agreement of this set (0.71 Cohen's *kappa*; Scikit-learn 2022), it is clear that even for human readers, labeling the loudness of context-free sound words is a non-trivial problem. For the purposes of this study, the two annotators were able to compromise on an agreed-upon set. We applied these loudness levels to our automatically tagged texts by mapping detected sound words to the annotators' loudness values.

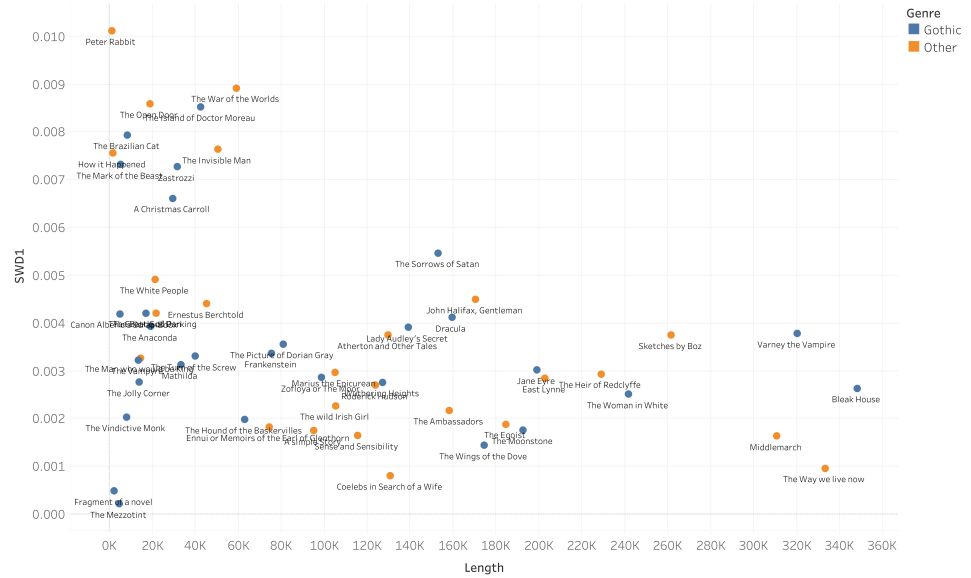
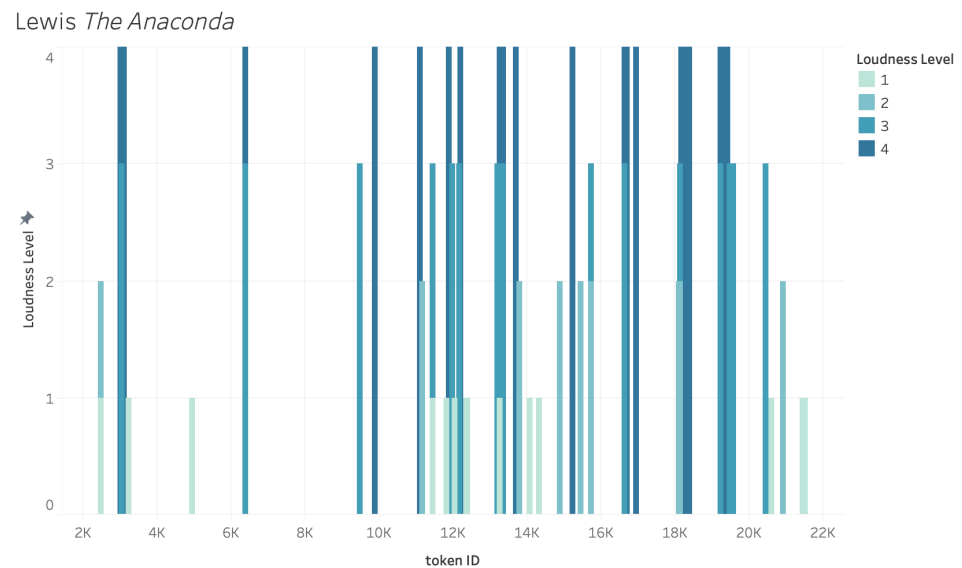


Figure 3: The dots show the texts and their average sound word density calculated for the whole text.

loudness level	example
0 no annotation	—
1 non-audible sounds	<i>silence</i>
2 low sounds	<i>rustling</i>
3 normal indoor volume	<i>snoring</i>
4 loud sound	<i>thunder</i>

Table 8: Loudness levels.





**Figure 4:** The columns indicate the loudness level of each sound word in Matthew Lewis' *The Anaconda* (from 1 (silence) to 4 (loud sound)).

## 4. Analysis: Loudness in the Gothic

### 4.1 Mapping the Manual Annotations

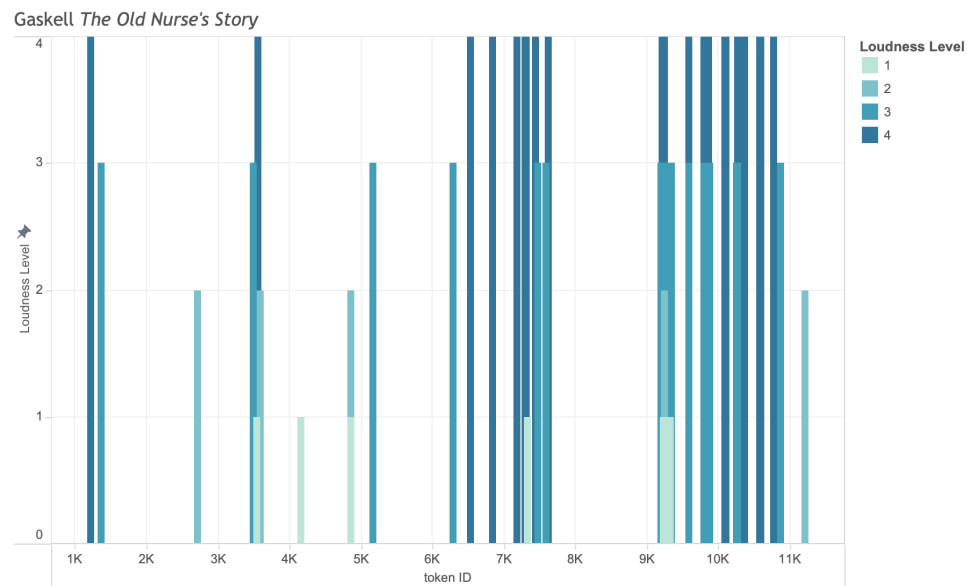
For each text, we created visualizations of which tokens were annotated, as well as the enriched loudness levels for the annotated tokens.

Figure 4 is the visualization of the manual annotations and loudness levels in Matthew Lewis' Gothic short story *The Anaconda*. Of note are the clusters of loud sounds (represented by the highest bars), which tend to occur together, interspersed with periods of quiet or explicit references to absolute silence. The majority of the story describes an encounter with an anaconda, which the characters attack at discrete moments of the text. These conflicts are captured by the two clusters of loud sounds towards the end of the text and represent the final attempt to kill the creature:

“But on a sudden a loud and rattling rush was heard among the palms, and with a single spring the snake darted down like a thunder-lap and twisted herself with her whole body round her devoted victim. [...] We all at once attacked her, and she soon expired under a thousand blows.” (Lewis *The Anaconda*)

In Figure 5, the distribution of the sound words over the course of Elizabeth Gaskell's *The Old Nurses' Story* is even more clustered. Interestingly, representations of loud sounds are particularly frequent in the second half of the story and correlate with scenes<sup>3</sup> that are particularly suspenseful.

3. By using the term 'scene' for narratological segments, we refer to the scene definition by Zehe et al. (2021): “From a narratological point of view, a scene can be defined by reference to a set of four dimensions: *time, space, action* and *character constellation*. Using these dimensions, a scene is a segment of the *discours* (presentation) of a narrative which presents a part of the *histoire* (connected events in the narrated world) such that (1) time is equal in *discours* and *histoire*, (2) place stays the same, (3) it centers around a particular action, and (4) the character constellation is equal”.

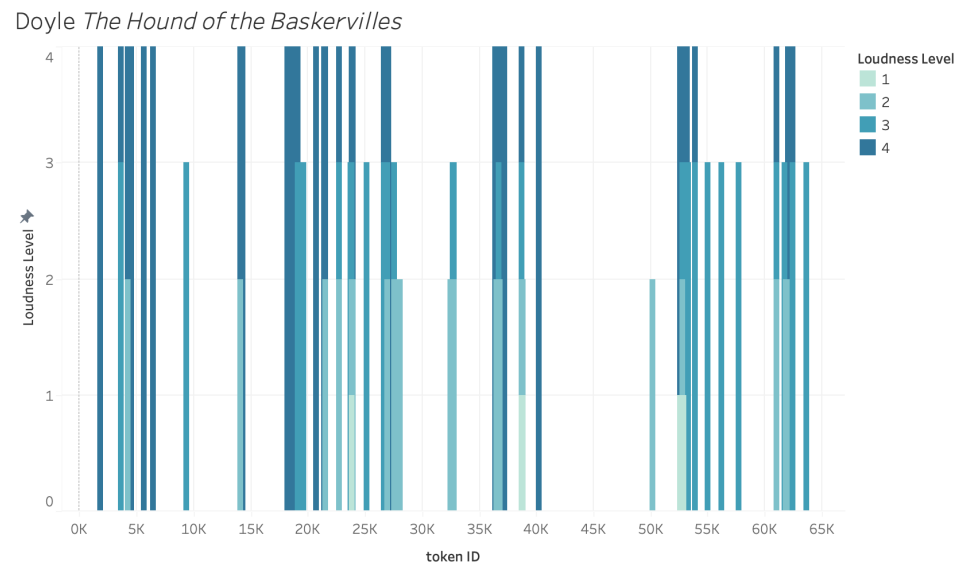


**Figure 5:** The columns indicate the loudness level of each sound word in Elizabeth Gaskell's *The Old Nurse's Story* (from 1 (silence) to 4 (loud sound)).

(9) “One fearful night, just after the New Year had come in, when the snow was lying thick and deep; and the flakes were still falling – fast enough to blind any one who might be out and abroad – there was a great and violent **noise** heard, and the old lord’s **voice** above all, **cursing** and **swearing** awfully, and the **cry** of a little child, and the proud defiance of a fierce woman, and the **sound** of a **blow**, and a dead **stillness**, and **moan** and **wailing dying** away on the hill-side!” (Gaskell *The Old Nurse's Story*; emphasis by authors)

Quotation (9) is the beginning of the cluster (around tokens 9,000– 11,000) that we can identify in Figure 5 indicating a detailed description of the fictional soundscape. The implied silence of the snowfall is interrupted by “a great and violent noise”, “cursing”, “swearing”, and a “cry”. The eruption of these loud sounds within a short interval increases the suspense of the scene and brings the story’s plot to a climax. The characters, and reader, however, have already encountered these loud sounds earlier in the text (around tokens 6,500– 7,500), where they may foreshadow the catastrophe.

Finally, the visualization of Doyle’s detective novel *The Hound of the Baskervilles* in Figure 6 shows relatively few sound words over the course of the plot. When they do occur, however, they appear clustered around passages that also make heavy use of Gothic tropes, including a vocabulary of mystery and fear. This is particularly the case in the passages with the highest density of sound words: (a) swd 3.13: 23 sw on 734 words in the penultimate cluster; (b) swd: 2.9: 18 sw on 621 words in the final cluster. These are both significantly higher than the average swd of the entire text ( $\emptyset$ -swd: 0.21). The actual passages reveal that the clusters of sound words coincide with two scenes that bracket the climax of the story. The last passage, containing a high density of sound words, contains the novel’s denouement: the appearance and killing of the mysterious hound.



**Figure 6:** The columns indicate the loudness level of each sound word in Doyle *The Hound of the Baskervilles* (from 1 (silence) to 4 (loud sound)).

## 4.2 Comparing Sound Word Density

We compared the sound word density (see subsection 3.3.3) in the automatically annotated corpus texts to determine if Gothic texts had a significantly higher swd than texts labeled as “other”, such as city or romance novels.

Despite the normalization of the measure, there was still a strong negative correlation (-0.4685) between swd and length at the extremes of the corpus size (an increase of 10 sound words has a much larger effect on a text of 1,000 words than one of 1,000,000 words). To counteract this bias, we divided the corpus into two subcorpora based on whether the texts were longer or shorter than 100,000 words, which lowered the correlation per subcorpus to between -0.17 and -0.19.

The plots of swd in the corpus texts (see Figure 9) demonstrate the categorical difference between the density of sound words in longer and short texts, with a mean swd of 0.471 in short texts and a mean swd of 0.26 for longer texts.

In the plot of longer texts, texts labeled as “other” have an observably lower sound word density than the Gothic novels. Romance and city novels, like Jane Austen’s *Sense and Sensibility* or Frances Trollope’s *The Way we live now*, have an especially low sound word density with Hannah More’s *Coelebs in Search of a Wife*, whose subtitle promises “Observations on Domestic Habits and Manners, Religion and Morals”, having the lowest value. The two long texts with the highest sound word density are Charles Dickens’s *Bleak House* – originally serialized in magazines, and Marie Corelli’s *The Sorrows of Satan* – a late 19<sup>th</sup> century horror novel: both are Gothic texts. Interestingly, Mary Russell Mitford’s *Atherton and Other Tales* has a high sound word density, although this could be because it contains a series of shorter tales and consequently may be more comparable to the short texts, where its sound word density of 0.35 is more in line with (although still less than) the mean swd in short stories.

The short texts do not show such a clear difference in swd between Gothic and “other”

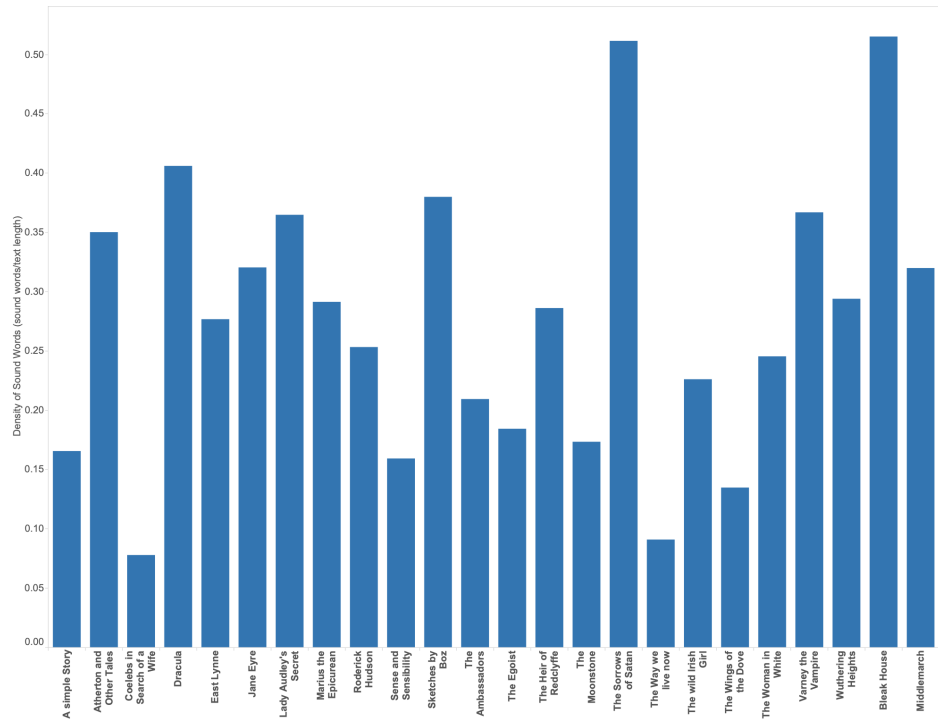


Figure 7: Swd of the long corpus texts (> 100,000 word tokens), ordered by date of publication. The scale differs from the one in Figure 8.

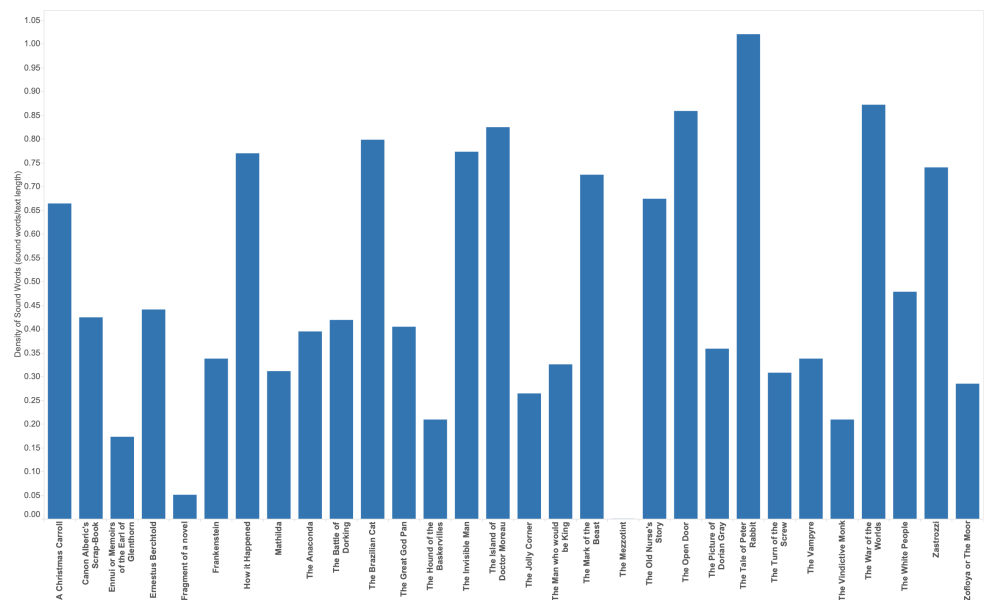
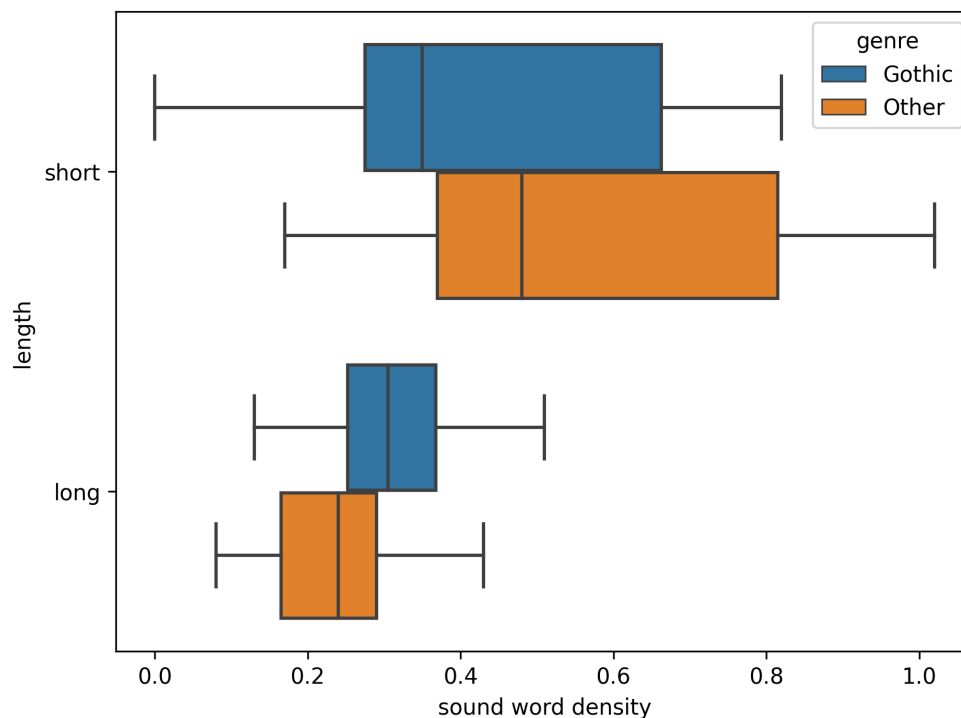


Figure 8: Swd of the short corpus texts (< 100,000 word tokens), ordered by date of publication. The scale differs from the one in Figure 7, because shorter texts have the tendency to have a higher sound word density, overall, than longer texts.



**Figure 9:** Swd of the corpus texts by length (short vs. long; < 100,000 word tokens) and genre (Gothic vs. “other”).

texts. Nevertheless, there are outliers with a particularly high sound word density such as Margaret Oliphant’s *The Open Door*. However, also H. G. Well’s *The War of the Worlds* that is labeled as “other” (Science Fiction) shows a high sound word density (we may hypothesize that *genre* fiction, rather than just Gothic fiction, may evidence higher swd scores overall, but this remains to be tested). There are, however, also Gothic texts with only one detected sound word like Byron’s vampyre story *Fragment of a Novel* or even no sound word at all as in M.R. James’ *The Mezzotint*). In contrast, Beatrix Potter’s children’s story *The Tale of Peter Rabbit* shows the highest sound word density of all corpus texts and simultaneously is also the shortest corpus text followed by Doyle’s *How it happened* – the second shortest text that has a slightly lower swd than the other short texts.

## 5. Discussion

As we stated in the introduction (see [section 1](#)), we are particularly interested in whether Gothic texts have more detailed ambient sound descriptions than texts labeled as “other”, particularly city or romance novels. As our sound word density analysis in [subsection 4.2](#) shows, there does seem to be a relationship between the Gothic genre and a high density of ambient sound words. Similarly, we found that passages that have a higher sound word density indicate important passages for the plot as we could see in the sample scenes mentioned in [subsection 4.1](#). In these cases, we could detect the climax of the plot by looking at the sound word distribution. In the following section, we will discuss these findings with an eye towards the distribution of loud sounds, as well as the role that silence plays in Gothic fiction.

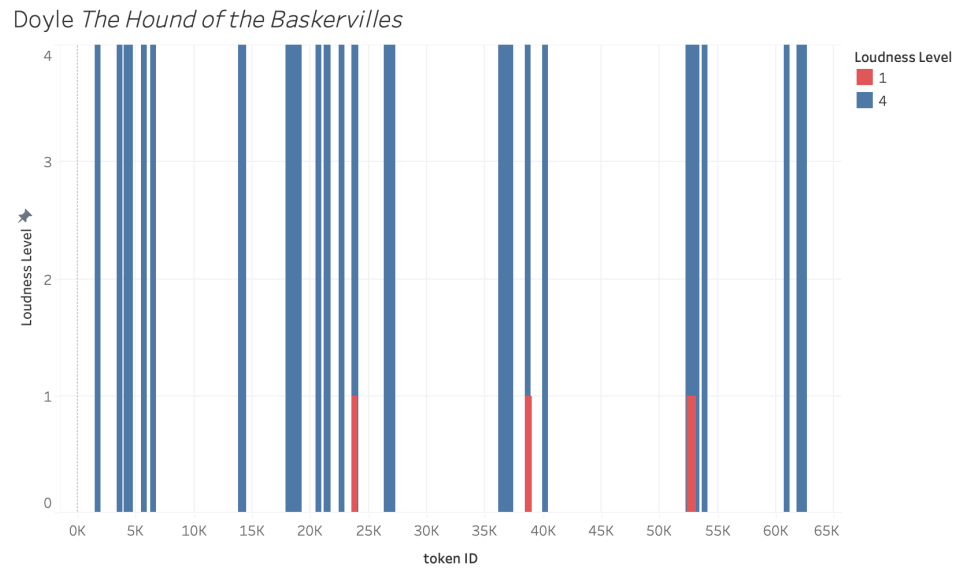


Figure 10: Doyle *The Hound of the Baskervilles* – silence and loud sounds.

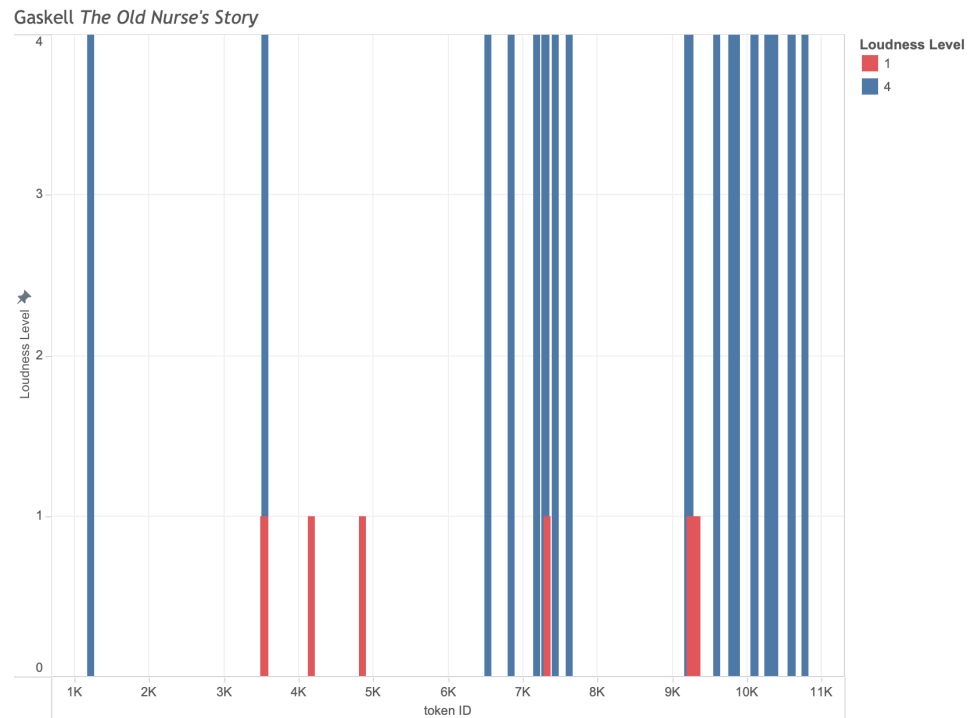
### 5.1 Silence versus Loud Sounds

As we explained in the discussion on the operationalization of sound in literary texts (see subsection 3.2), *silence* is a particular sub-phenomenon of the ambient soundscape. We introduced the differentiation between the explicit indication of absolute silence and the absence of any sound indication because the latter does not indicate the absence of sound in the fiction. Rather, it plays with the imagination of the reader, offering gaps in the narration that trigger the reader to fill them with world knowledge of an expected soundscape according to the given scene setting. Consequently, only explicitly indicated silence is tagged as *silence*: the lack of the representation of sound does not imply the lack of sound in the same way that explicit references to silence does.

From these results, we might conclude that silence often sets a scene, as it involves a sustained period featuring the explicit absence of sounds perceptible to humans. Loud sounds, by contrast, flag events that occur spontaneously and irregularly, often interrupting this state of silence. The contrast between silence and loudness amplifies the effects of sound, thereby increasing its effects on the reader. An effect of this pattern is that several events together convey a loud ambiance, while the silent initial state is often mentioned only once and therefore also has only a small effect on a passage's mean loudness (see Figure 10). Doyle's *The Hound of the Baskervilles* offers an important example of this scene-setting process. The denouement of the novel starts at token 52,548 with "A terrible scream – a prolonged yell of horror and anguish – burst out of the **silence** of the moor" (emphasis by authors). Here, the silence of the moor is interrupted by a "scream" that is described with words typical for the Gothic vocabulary (like "terrible", "horror", "anguish"). The interruptions oscillate quickly between loudness levels, disorienting both the characters and the reader, and mixing the uncertainty of the passage with moments of surprise in a constant play of tension and release.

Gaskell's *The Old Nurse's Story* has few moments of silence, but in the two times that it does occur, it sets up a sequence of suspenseful scenes that begin with general silence or at least a quiet ambiance. Instead of explicit representations of silence, Gaskell's





**Figure 11:** Gaskell *The Old Nurse's Story* – silence and loud sounds.

text works through absence and negation, describing the missing sounds from an environment. These do not indicate absolute silence, but have a similar effect on the soundscape of a scene. For example, in a key scene, it is the *absence* of the expected sounds that should occur when a character is crying and battering her hands against the window-panes that creates the uncertainty and tension:

(10) “[A]ll of a sudden, she cried out, ‘Look, Hester! look! there is my poor little girl out in the snow!’ I turned towards the long narrow windows, and there, sure enough, I saw a little girl [...] crying, and beating against the window-panes, as if she wanted to be let in. [...] all of a sudden, and close upon us, the great organ pealed out so **loud** and **thundering**, it fairly made me tremble; and all the more, when I remembered me that, even in the stillness of that dead-cold weather, I had heard **no sound** of little battering hands upon the windowglass, although the phantom child had seemed to put forth all its force; and, although I had seen it wail and cry, **no faintest touch of sound** had fallen upon my ears.” (Gaskell *The Old Nurse's Story*; emphasis by authors)

Although in most examples, it is loud sounds that punctuate a silent atmosphere, the reverse is also possible. These occasions can be more unsettling given that the expected sound is replaced by an unexpected silence. In Rymer’s *Varney the Vampire*, we have a low scratching noise. However, rather than a crash of an invader, we have a sudden silence and only then does the vampire appear. The low sound is interrupted by the silence rather than the silence interrupted by the sound.

“Mrs. Bannerworth [...] heartily regretted she had not rung the bell, for, before, another word could be spoken, there came too perceptibly upon their

ears for there to be any mistake at all about it, a strange scratching noise upon the window outside. A faint cry came from Flora's lips [...]. **The scratching noise continued for a few seconds, and then altogether ceased.** [...] When the scratching noise ceased, Flora spoke in a low, anxious whisper, as she said, – 'Mother, you heard it then?'" (Rymer *Varney the Vampire*; emphasis by authors)

## 5.2 Sound and Suspense

With regard to the sound analysis results of our 19<sup>th</sup> century English fiction corpus, we suggest that sound plays an important role in the Gothic as well as in other suspenseful genres of the period, such as science fiction, or detective stories. However, these different genres contain different types of ambient sound that lie outside of the scope of this article to distinguish. For example, sounds that are indicative of a character's uncertainty about a current state of affairs, e.g., 'rustling', 'crackling', or even as in quotation (11) the 'sound of wheels', serve a different purpose than, for example, sounds that result from dangerous events (e.g., 'thunder', 'explosive blasts'). The distinction between further subcategories of ambient sound and their effects should be investigated in a further study. Still, we could recognize that sounds appear to be much less frequent in romance or city novels in which even the background soundscape is rarely referenced:

(11) "At that moment the **sound of wheels** was heard and Charlotte flew off to her private post of observation." (Yonge *The Heir of Redclyffe*; emphasis by authors)

When representations of sound do intrude into these novels, they signal a deviation from the default soundscape and suggest action, but they do not offer enough information for the reader to interpret it, creating questions and uncertainty. The sudden interruption by explicit sound references in an implicit soundscape serves as an unexpected event, and has a surprising and suspenseful effect on the reader by interrupting the expected experience of the scene.

With regard to Gothic texts, especially mysterious, inexplicable sounds amplify the affect of suspense, creating uncertainty and driving the reader's desire to resolve a given mystery:

(12) "'That is the story. **Whatever the sound is, it is a worrying sound,**' says Mrs. Rouncewell, [...] 'and what is to be noticed in it is that it **MUST BE HEARD.** My Lady, who is afraid of nothing, admits that when it is there [...]'" (Dickens *Bleak House*; emphasis by authors)

For the analysis of the fictional soundscape however, it is not sufficient for a sound word to simply lexically appear in the text as in quotation (12). As we argue in [subsection 3.2](#), the word must represent the presence of the sound itself in the scene. Conditional statements about possible sounds, descriptions of eagerly awaited sounds, comparisons to known sounds, or reports of sounds that happened in the past do not effect the fictional soundscape in our analysis. Consider, for example, the atmosphere created by the conversation on the mysterious sobbing of a woman that the characters in Doyle's text have heard the night before:

(13) “‘And yet it was not entirely a question of imagination,’ I answered. ‘Did you, for example, happen to hear someone, a woman I think, sobbing in the night?’ ‘That is curious, for I did when I was half asleep fancy that I heard something of the sort’.” (Doyle *The Hound of the Baskervilles*)

The mystery of the sound is undercut by the rational conversation that contextualizes it: it is not experiential but recollected and so does not trigger suspense for either the reader or character.

## 6. Conclusion and Outlook

In this article, we have demonstrated the great potential of sound studies for literary analysis. Our analyses, which combined distant reading methods with close readings, offer evidence for our hypothesis that Gothic texts contain more detailed descriptions of the story’s ambient soundscape than our corpus texts labeled as “other”. Our operationalization of ambient sound and the prediction model enabled us to explore sound from a computational perspective to reveal new facets of the soundscape of fiction. The distinction between represented and implicit or hypothetical sounds, however, presented a continuing challenge to our model. Consequently, context is crucial for understanding the role that sound words play as demonstrated by the difference in success of our dictionary model versus the transfer-learning classifier. Despite the relatively high number of false positive predictions, the model trained on the manual and semi-automatic annotations performed surprisingly well at detecting ambient sounds.

Our results argue for increased scholarly attention to sound in fiction, and, in particular, for the ways in which such automated approaches to the analysis of sound could be harnessed to provide a deeper understanding of the role sound plays in narrative across a much broader period. For example, the systematic analysis of the relation between sound and suspense could be an important direction for future work. Similarly, as we close read the passages surfaced by our study, we also found evidence of other sensory descriptors. In a future project, the analysis of olfactory or haptic sensations could extend our study, as well as open up new affective representations for analysis.

## 7. Data Availability

Data can be found here: [https://github.com/SvenjaGuhr/Sound\\_and\\_Suspense](https://github.com/SvenjaGuhr/Sound_and_Suspense).

## 8. Software Availability

Used Software can be found here: [https://github.com/NEISSproject/tei\\_entity\\_enricher](https://github.com/NEISSproject/tei_entity_enricher) and [https://github.com/SvenjaGuhr/Sound\\_and\\_Suspense](https://github.com/SvenjaGuhr/Sound_and_Suspense).

## 9. Acknowledgements

We thank the reviewers for their detailed comments on our manuscript and constructive feedback that helped to refine and focus this article.

This article is the result of Svenja Guhr’s research stay at the Stanford Literary Lab as a visiting student researcher in Autumn Term 2022, which was financially supported by the following three parties:

- Department of Digital Philology, Institute of Linguistics and Literary Studies at the Technical University of Darmstadt,
- Support of Women in Academia, program of the Equal Opportunity Commission of the Department of History and Social Sciences at the Technical University of Darmstadt,
- Fellowship of the German Academic Exchange Service (DAAD Program for visiting doctoral students).

Special thanks go to our student assistant Alina Klein who supported the iterative process of annotation guideline creation and the annotation of the training data as a third annotator.

## 10. Author Contributions

**Svenja Guhr:** Conceptualization, Data Curation, Methodology, Visualization, Writing – original draft

**Mark Algee-Hewitt:** Supervision, Visualization, Writing – review & editing

## References

- Algee-Hewitt, Mark, Anna Mukamal, and J. D. Porter (2023). “The Affordances of Mere Length: Computational Approaches to Short Story Analysis”. In: *The Cambridge Companion to the American Short Story*. Ed. by Michael J. Collins and Gavin Jones. Cambridge University Press, 341–357. [10.1017/9781009292863.028](https://doi.org/10.1017/9781009292863.028).
- Bacon, Simon, ed. (2018). *The Gothic: a Reader*. Peter Lang.
- Bernhart, Toni (2008). “Stadt hören: Auditive Wahrnehmung in Berlin Alexanderplatz von Alfred Döblin”. In: *Zeitschrift für Literaturwissenschaft und Linguistik* 38.1, 51–67. [10.1007/BF03379955](https://doi.org/10.1007/BF03379955).
- Blohm, Stefan, Maria Kraxenberger, Christine A. Knoop, and Mathias Scharinger (2021). *Sound Shape and Sound Effects of Literary Texts*. De Gruyter. [10.1515/9783110645958-002](https://doi.org/10.1515/9783110645958-002).
- Botting, Fred (1996). *Gothic*. The New Critical Idiom. Routledge.
- Devlin, Jacob, Ming-Wei Chang, Kenton Lee, and Kristina Toutanova (2018). “BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding”. In: *CoRR abs/1810.04805*. [10.48550/arXiv.1810.04805](https://arxiv.org/abs/1810.04805).
- Ellis, Markman (2000). *The History of Gothic Fiction*. Edinburgh University Press.
- Finkel, Jenny Rose, Trond Grenager, and Christopher Manning (2005). “Incorporating Non-local Information into Information Extraction Systems by Gibbs Sampling”. In: *Proceedings of the 43rd Annual Meeting of the Association for Computational Linguistics (ACL’05)*. Association for Computational Linguistics, 363–370. [10.3115/1219840.1219885](https://doi.org/10.3115/1219840.1219885).

- Flüh, Marie, Jan Horstmann, and Mareike Schumacher (2022). "Genderaspekte in Fantasy-Jugendromanen von 2008 bis 2020: Distant Gender Reading". In: *Gender in der deutschsprachigen Kinder- und Jugendliteratur*. Ed. by Weertje Willms. De Gruyter, 457–482. [10.1515/9783110726404-025](https://doi.org/10.1515/9783110726404-025).
- Flüh, Marie and Marc Lemke (2022). "An Experimental Attempt to Use Transfer Learning for Named Entity Recognition in Letters from the 19th and 20th Century". In: *Book of Abstracts*. <https://dh2022.dhii.asia/dh2022bookofabsts.pdf> (visited on 12/28/2022).
- Foley, Matt (2023). *Gothic Voices: The Vococentric Soundworld of Gothic Writing*. 1st ed. Cambridge University Press. [10.1017/9781009162579](https://doi.org/10.1017/9781009162579).
- Gius, Evelyn and Michael Vauth (2022). "Towards an Event Based Plot Model. A Computational Narratology Approach". In: *Journal of Computational Literary Studies* 1.1. [10.48694/jcls.110](https://doi.org/10.48694/jcls.110).
- Glotova, Elena (2021). *Soundscapes in nineteenth-century Gothic short stories*. Umeå University. <https://urn.kb.se/resolve?urn=urn:nbn:se:umu:diva-183080> (visited on 01/27/2024).
- Guhr, Svenja (2023). *Sound and Suspense*. GitHub Repository. [https://github.com/SvenjaGuhr/Sound\\_and\\_Suspense](https://github.com/SvenjaGuhr/Sound_and_Suspense).
- Guhr, Svenja and Evelyn Gius (2023). "Maschinen als Erzähltheoretiker". In: *Kongressakten IVG 2020. Internationales Jahrbuch für Germanistik*. Peter Lang.
- Hinton, Leanne, Johanna Nichols, and John Ohala (1995). "Introduction: Sound-symbolic processes". In: *Sound Symbolism*. Ed. by Leanne Hinton, Johanna Nichols, and John J. Ohala. 1st ed. Cambridge University Press, 1–12. [10.1017/CB09780511751806.001](https://doi.org/10.1017/CB09780511751806.001).
- Hironsan, Hiroki Nakayama (2018). *seqeval: A Python framework for sequence labeling evaluation*. <https://github.com/chakki-works/seqeval>.
- Horstmann, Jan (2020). "Undogmatic Literary Annotation with CATMA". In: *Annotations in Scholarly Editions and Research*. Ed. by Julia Nantke, Frederik Schlupkothén, and Jan Horstmann. De Gruyter, 157–176. [10.1515/9783110689112-008](https://doi.org/10.1515/9783110689112-008).
- Hühn, Peter (2013). "Event and Eventfulness". In: ed. by Peter Hühn, John Pier, Wolf Schmid, and Jörg Schönert. <https://www-archiv.fdm.uni-hamburg.de/lhn/node/39.html> (visited on 10/18/2022).
- Hurley, Kelly (2002). "British Gothic Fiction, 1885–1930". In: *The Cambridge Companion to Gothic Fiction*. Ed. by Jerrold E. Hogle. Cambridge Companions to Literature. Cambridge University Press, 189–207.
- Kamath, Uday, John Liu, and James Whitaker (2019). "Transfer Learning: Domain Adaptation". In: *Deep Learning for NLP and Speech Recognition*. Springer International Publishing, 495–535. [10.1007/978-3-030-14596-5\\_11](https://doi.org/10.1007/978-3-030-14596-5_11).
- Loper, Edward and Steven Bird (2002). "NLTK: The Natural Language Toolkit". In: *arXiv preprint*. [10.48550/ARXIV.CS/0205028](https://arxiv.org/abs/10.48550/ARXIV.CS/0205028).
- Manning, Christopher, Mihai Surdeanu, John Bauer, Jenny Finkel, Steven Bethard, and David McClosky (2014). "The Stanford CoreNLP Natural Language Processing Toolkit". In: *Proceedings of 52nd Annual Meeting of the Association for Computational Linguistics: System Demonstrations*. Association for Computational Linguistics, 55–60. [10.3115/v1/P14-5010](https://doi.org/10.3115/v1/P14-5010).
- Mildorf, Jarmila (2019). "Can Sounds Narrate? Prosody in Sound Poetry Performance". In: *CounterText* 5.3, 294–311. [10.3366/count.2019.0167](https://doi.org/10.3366/count.2019.0167).

- Mulvey-Roberts, Marie, ed. (2009). *The Handbook of the Gothic*. 2nd ed. New York University Press.
- Pichler, Axel and Nils Reiter (2020). "Reflektierte Textanalyse". In: *Reflektierte algorithmische Textanalyse*. Ed. by Nils Reiter, Axel Pichler, and Jonas Kuhn. De Gruyter, 43–60. [10.1515/9783110693973-003](https://doi.org/10.1515/9783110693973-003).
- Picker, John M. (2003). *Victorian soundscapes*. Oxford University Press.
- Reiter, Nils (2020). "Anleitung zur Erstellung von Annotationsrichtlinien". In: *Reflektierte algorithmische Textanalyse*. Ed. by Nils Reiter, Axel Pichler, and Jonas Kuhn. De Gruyter, 193–202. [10.1515/9783110693973-009](https://doi.org/10.1515/9783110693973-009).
- Schafer, R. Murray (1994). *The Soundscape: our Sonic Environment and the Tuning of The world*. Destiny Books.
- Schumacher, Mareike (2022). *Orte und Räume im Roman*. Digitale Literaturwissenschaft. J.B. Metzler. [10.1007/978-3-662-66035-5](https://doi.org/10.1007/978-3-662-66035-5).
- Schumacher, Mareike, Marie Flüh, and Marc Lemke (2022). "The Model of Choice. Using Pure CRF- and BERT-Based Classifiers for Gender Annotation in German Fantasy Fiction". In: *Book of Abstracts*. <https://dh2022.dhii.asia/dh2022bookofabsts.pdf> (visited on 12/28/2022).
- Scikit-learn, Developers (2022). 3.3. *Metrics and Scoring: Quantifying the Quality of Predictions*. [https://scikit-learn/stable/modules/model\\_evaluation.html](https://scikit-learn/stable/modules/model_evaluation.html) (visited on 12/17/2022).
- Smith, Mark (2015). *Listening to Nineteenth-Century America*. The University of North Carolina Press.
- Snaith, Anna, ed. (2020). *Sound and literature*. Cambridge University Press.
- Sperfeld, Konrad and Marc Lemke (2022). *NEISS NTEE. User Interface. Documentation*. [https://github.com/NEISSproject/tei\\_entity\\_enricher/wiki/user-interface](https://github.com/NEISSproject/tei_entity_enricher/wiki/user-interface).
- TEI Consortium (2022). *TEI P5: Guidelines for Electronic Text Encoding and Interchange*. Version v4.5.0. [10.5281/zenodo.7382490](https://doi.org/10.5281/zenodo.7382490).
- Verma, Neil (2012). *Theater of the Mind: Imagination, Aesthetics, and American Radio Drama*. University of Chicago Press. [10.7208/9780226853529](https://doi.org/10.7208/9780226853529).
- Zehe, Albin, Leonard Konle, Svenja Guhr, Lea Dümpelmann, Evelyn Gius, Andreas Hotho, Fotis Jannidis, Lucas Kaufmann, Marcus Krug, Frank Puppe, Nils Reiter, and Annekea Schreiber (2021). "Shared Task on Scene Segmentation (STSS). Task Description Paper". In: *Proceedings of the 17th Conference on Natural Language Processing (KONVENS)*. <http://lsx-events.informatik.uni-wuerzburg.de/files/stss2021/proceedings/stss.pdf> (visited on 12/21/2022).
- Zöllner, Jochen, Konrad Sperfeld, Christoph Wick, and Roger Labahn (2021). "Optimizing Small BERTs Trained for German NER". In: *Information* 12.11, 443. [10.3390/info12110443](https://doi.org/10.3390/info12110443).