



Jan Kochanowski University Press

This is a contribution from *Token: A Journal of English Linguistics*
Volume 13/2021.

Edited by John G. Newman, Marina Dossena and Sylwester Łodej.

© 2021 Jan Kochanowski University Press.

Corpus stylistics, classic children's literature and the lexical field of laughter

John Corbett* & Li Li**

* *BNU-HKBU United International College*

** *Macao Polytechnic University*

ABSTRACT

This article draws upon data from the *Historical Thesaurus of English (HTE)* to explore the lexical field of laughter in a corpus of thirteen children's novels. The thirteen novels are all from the first 'Golden Age' of children's literature in English, namely the late 19th and early 20th century. The study takes items in the *HTE's* lexical domain of laughter and identifies the frequency, distribution and collocation of those items as they appear in the corpus. The results are discussed with reference to the ways in which different lexicalisations of laughter indicate the authorial stance towards childhood and also towards members of the communities represented in the novels. By indicating, through the representation of different types of laughter, the author's preferred moral stance towards particular individuals and groups, the novels prompt young readers to accept or challenge modes of behaviour that exemplify or threaten communal values and good citizenship. The study thus demonstrates how readers of children's literature from the 'Golden Age' are linguistically conditioned to reject negative forms of laughter and instead embrace positive forms, as they move from the undisciplined laughter of childhood to the relative restraint of adulthood.

Keywords: *Historical Thesaurus of English*, corpus stylistics, laughter, children's literature.

1. Introduction

This study combines both established and innovative lexicographical resources with text analysis software to explore the lexical domain of laughter in a small corpus of 13 novels from the first 'Golden Age' of children's literature in the 19th and early 20th centuries. The thirteen novels

are selected from a rich period in which, critics generally agree, children's authors balanced their moral and educational aspirations with a desire to entertain children in their own terms (see Carpenter 1985; Darton 1982 [1932]; Sorby 2011). We were therefore interested in exploring, stylistically, in the novels the frequency and distribution of a set of lexical items that indexed both moral stance and pleasure. Given that laughter can represent a range of expressive and relational stances, from the spontaneous outburst of pleasure to a sneering indication of superiority, we have investigated the frequency and distribution of a set of items in the lexical field of laughter, as they appear in a number of 'Golden Age' novels. We address the following questions:

- Which items in the lexical field of laughter appear in the novels?
- What is the distribution of these items across the novels?
- What do the collocations of these items tell us about the conceptual domain of laughter, as it is represented in the novels?
- How do the choices made from the lexical domain of laughter manage the readers' stance towards the characters in the novels, and the characters' relations with each other?

To answer these questions, we began by consulting the *Historical Thesaurus of English* (HTE; see Kay et al. 2009) in order to identify the expressions related to laughter that were available during the first 'Golden Age' of children's literature. We then used the text analysis software, AntConc (Anthony 2019), to determine which lexical items from that set of expressions were actually used in the novels, to calculate their frequency, distribution, and statistical significance, and to explore how they were modified. The corpus-informed stylistic approach that we take to the study of laughter in a corpus of children's novels is quantitative, and it can be understood as a type of 'distant reading' (Moretti 2005).

2. The lexical field of laughter

The availability over the past decade of the *Historical Thesaurus of English* has made possible certain types of investigation into semantics and stylistics that were hitherto impossible (see Anderson et al. 2016; Busse 2012). We consulted the HTE to identify a set of words within the lexical field of laughter that were current during the first 'Golden Age' of writing for children. Then we used text analysis software to find the frequency and distribution of the

various expressions within the corpus. Finally, collocations of *laugh/laughter* were analysed to discover the ways in which acts of laughing and laughter are conceptualised in the novels. The collocations show that certain concepts related to laughter, such as spontaneity, control, and sincerity, are salient in the corpus. The analysis suggests that, in these novels, unrepressed, communal laughter is an index of unconstrained youth, and that learning to identify the 'proper' kinds of laughter is a rite of passage from childhood to adulthood.

The methods parallel other corpus stylistics investigations, such as Oster's (2010) use of corpora to explore the linguistic expression of emotions. Based on the historical dictionaries of English, principally the *Oxford English Dictionary* (OED), the HTE classifies the lexical resources of English into 250,000 discrete concepts, divided first into The World, The Mind and Society. These concepts include the verb and noun *laugh*, the noun *laughter*, and all the expressions in English that are semantically related to them, such as *giggle*, *titter*, *sneer*, *chuckle*, *chortle*, and so on. The HTE also indicates the chronology of the development of the lexical domain. Both *laugh* and *laughter* are recorded as far back as the early Old English period (from OE *ahliehhan*, *hleahter*, etc.). The OED suggests that the etymology is imitative, and that the word is Indo-European in origin. Although the form of the present-day word *laughter* dates from Old English, its meaning has changed. While in OE texts, *laughter* could signify a single instance of laughter, and appear as a plural (*hleahrtres*), from at least the 16th century, the verbal form *laugh* began also to be used as a countable noun indicating an instance of laughter, and the form *laughter* was increasingly reserved for use as a mass noun with a more generalized meaning.

Specific ways of laughing in English are more recent. The OED suggests that *giggle* is echoic, but its use as a verb is not recorded until 1509 and a nominal usage is not recorded until 1611. The expression *chortle* – apparently a blend of *chuckle* and *snort* – is not recorded until 1871, when Lewis Carroll coined it in the poem 'Jabberwocky.' Again, the verbal use precedes the nominal use, which is not recorded in the OED until 1903. The expression, now reasonably common, appears only once in our corpus, in the second of the 'Alice' books, *Through the Looking Glass*, in which the poem, 'Jabberwocky,' first appeared.

While the lexical domain of laughter contains other forms of words – the adverb *laughingly* is recorded from 1475 and *sniggeringly* from 1886, we focus here, for reasons of space, largely on verbal and nominal forms and senses that indicate types of laugh/laughter that are current during the

period of our corpus (1865-1911). From the evidence of the *HTE*, in this lexical domain, verbal forms and senses usually predate the nominal ones. Certain expressions arose in English and died out before the period of the corpus (e.g. *unlaugh*, 'to reverse the laughing process,' is sporadically recorded only in 1532 and 1637), while others appear later (e.g. *hoot* in the sense of *laugh* is not recorded until 1926, and *laugh-in*, an event characterised by laughter, is not recorded until 1968). As we have already noted, although some words endure, their senses change; thus from 1598-1823, the word *chuckle* is recorded with the sense of laughing convulsively or immoderately, but thereafter its meaning weakens or narrows to the sense of laughing quietly and with contentment, a sense first recorded in 1803.

The selection of search items for this study was motivated by a concern to understand how laughing and laughter are deployed stylistically in the corpus. We are thus less concerned with whether *laugh* is used in the texts as a verb or a noun, and more with whether, when and why characters (give a) *laugh*, *giggle*, *chuckle*, *snort*, or, indeed, *chortle*. We are also concerned with how these expressions are modified, that is, whether, when, and why characters might give, for example, a brave, honest laugh, or a cowardly, hollow laugh. We are concerned, in short, with how laughter relates to the moral universe of the texts.

The intransitive verb, *laugh*, and the related noun, *laughter*, are both coded 02.04.10.11 in the *HTE*. This code indicates that the lexical domain of *laugh/laughter* has been categorized under 02 The Mind > 04 Emotions > 10 Pleasure > 11 Laughter. This coding is evidently based on the sense of *laugh/laughter* as a vocal outburst that is expressive of pleasure; there are other senses and homonyms that do not concern us here, such as the northern English and Scots use of *a laughter* referring to the total number of eggs laid by a hen or another bird, e.g. 'A hen lays her laughter, that is, all the eggs she will lay that time.'¹ These unrelated senses are excluded from our analysis.

Further and finer *HTE* codings indicate relevant subcategories of *laugh/laughter*. The search items selected for inclusion in our analysis are based on the lists in Appendices 1 and 2. We have excluded from the lists in the Appendices those items that were current in the period, according to the *OED/HTE*, but which do not appear in our corpus, such as *cachinnate* 1824- 'to laugh loudly/coarsely'. The potential search items in Appendix 1 are based solely on the intransitive uses of the verb – the transitive list in

¹ "laughter, n.2" (*OED Online*).

the *HTE* brings up only one further possible item, namely *guff* (1865-) in the sense of 'utter/express with loud/coarse laughter' and this item, like the similar *guffaw*, is absent from our corpus. The nominal forms in Appendix 2 largely repeat exactly the word forms of the verbs, with a few exceptions, e.g. the verb *convulse* corresponds to *convulsion(s)* in the sense of 'a fit or fits of laughter.'

There are a few points worth noting at the outset about this lexical domain. Since the action of laughing involves, to quote part of the first *OED* definition of *laugh*, 'the spontaneous sounds and movements of the face and body usual in expressing joy, mirth, amusement, or (sometimes) derision,'² it is not surprising that many of the items in the lexical domain of laughter (including *laugh* itself) are imitative or echoic, for instance, *roar*, *snort*, *haw-haw*, *tee-hee*, *giggle*, *titter*, *cackle*. Although laughter is a spontaneous expression of pleasure, arguably only *chuckle* is now used to express moderate contentment. The animal nature of laughter is salient in expressions such as *roar*, *snort*, *whinny*, *cackle*, and *horse-laugh*. The idea of a spontaneous outburst that cannot be controlled is evident in expressions like *die with laughter*, *split the sides*, *laugh oneself sick/silly*, *break up*, *convulsion(s)*, etc. The lack of control associated with certain forms of laughter may be associated with foolish laughter, for example, *giggle*, or *titter*. Furthermore, although laughter is categorised in the *HTE* primarily in relation to pleasure and mirth, an element of derision is highlighted in numerous expressions, such as *snicker*, *snigger*, *cackle*, *laugh in one's sleeve*. Finally, there is evidently some overlap and fuzziness in the use of these expressions. For example, the word *sneer* is categorised within 'foolish laughter,' but it can also mean 'derisory laughter'; and different senses of *chuckle* mean that it can be considered either as a type of 'snigger' or a subcategory by itself.

Despite these instances of fuzziness and overlap, the lexical domain of *laugh/laughter*, as shown in the *HTE*, gives a reasonably clear indication of the extent and boundaries of the concept in English: laughter is a physical, usually vocal, outburst of mirth that might indicate pleasure or derision; it can be uncontrolled and immoderate, and may be suggestive of an animal nature or foolishness. These are the conceptual traits of laughter as described in reference works like dictionaries and thesauri. However, *laugh/laughter* can also be modified in context to suggest a broader range of characteristics, and to analyse this phenomenon further, we need to turn from lexicographical works of reference to a corpus.

² "laugh, v." (*OED Online*).

3. The corpus of 'Golden Age' novels

The present study explores the uses of laughter in a small corpus of thirteen children's novels in what is generally considered the first 'Golden Age' of children's literature in English. Harvey Darton and Humphrey Carpenter suggest that this 'Golden Age' begins with Lewis Carroll's *Alice's Adventures in Wonderland* (1865) and draws to a close with A.A. Milne's stories of Winnie-the-Pooh, the last of which was *The House at Pooh Corner* (1928). The 'Golden Age' canon consists of texts written specifically for children that are characterised by an emphasis on entertainment as much as or rather than moral instruction (Sorby 2011), and a new willingness to view the world from the perspective of a child, like Alice, or Peter Pan, or a child-surrogate, such as Beauty in *Black Beauty*, or Buck in *The Call of the Wild*.

Certainly, the texts of the period acknowledge the complexities a child faces when negotiating and eventually entering the world of adults. The specific question that the present study raises is the role of laughter in making that transition. The thirteen novels in the corpus were downloaded in digital form from the Project Gutenberg website³ and edited to remove extraneous matter. The novels are listed in Table 1. They were not selected as necessarily being representative of all writing for children in this period; we wished simply to be able to compare the uses of laughter in any single novel of the period with the uses of laughter in a reasonable number of other novels of the same period. Lewis Carroll's *Through the Looking Glass* was added to the corpus when we realised that there are no occurrences whatsoever of expressions relating to *laugh/laughter* in *Alice's Adventures in Wonderland*, a fact that is interesting, in itself, and which is discussed below. The titles, authors, dates of publication and provenance (UK/USA) are given, with the word count for each novel. The provenance is given as there might be a preference for a particular form (e.g. *snicker/snigger*) in American or British English.

The lexical field of laughter affords the authors of these novels a range of possible expressions with more or less specific senses by which they can portray characters and their actions, and thus indicate the stance of the characters towards themselves, others, and events in the world. The expressions within the domain of laughter can themselves be modified, particularly the less specific, superordinate term, *laugh*, which is general in meaning, no matter whether it is used as a noun or a verb. The action of

³ *Books in Children's Literature* (Project Gutenberg).

laughing can be expressed and modified either by using a verb, i.e. *she laughed*, or by using a delexicalized or 'light verb' (that is, verb whose meaning is reduced and the nature of the action is conveyed by the grammatical object, as in *she gave a laugh/she had a laugh*, etc).

Table 1. The Corpus

Text No.	Code	Title	Author	Date	Number of words (tokens)	Provenance
1	AW	<i>Alice's Adventures in Wonderland</i>	Lewis Carroll	1865	10,021	UK
2	LG	<i>Through the Looking Glass</i>		1871	30,618	
3	LW	<i>Little Women</i>	Louisa May Alcott	1868	191,196	USA
4	WKD	<i>What Katy Did</i>	Susan Coolidge	1872	51,129	USA
5	BB	<i>Black Beauty</i>	Anna Sewell	1877	60,848	UK
6	TI	<i>Treasure Island</i>	Robert Louis Stevenson	1883	70,425	UK
7	WVO	<i>The Wonderful Wizard of Oz</i>	L. Frank Baum	1900	39,888	USA
8	CW	<i>The Call of the Wild</i>	Jack London	1903	32,365	USA
9	RSF	<i>Rebecca of Sunnybrook Farm</i>	Kate Douglas Wiggin	1903	76,090	USA
10	WW	<i>The Wind in the Willows</i>	Kenneth Grahame	1908	60,754	UK
11	PPW	<i>Peter Pan and Wendy</i>	James M. Barrie	1911	48,178	UK
12	SG	<i>The Secret Garden</i>	Frances Hodgson Burnett	1911	83,164	UK
13	DD	<i>The Story of Doctor Dolittle</i>	Hugh Lofting	1920	27,570	UK
Total number of tokens in the corpus					782,246	

The use of ‘light’ verbs allows for a wide range of modifications, e.g. *she gave a hearty/delicate/sinister/hollow laugh*, etc. The generality of *laugh* is evident when one attempts to substitute it with one of its subcategories; one can hardly give a **hearty giggle*, **delicate roar*, **sinister guffaw* or **hollow titter*.⁴ The more restricted senses of the expressions in the subcategories constrain the ways in which they are modified, e.g. *hearty roar*, *foolish giggle*. While the description of a character as *chuckling*, *sneering* or *giggling* may be taken to be indicative of a character’s personality, then, the act of laughing is, by itself, less revealing. It is therefore necessary to look at the immediate contexts in which the terms *laugh*, *laughter* and their related forms, occur, and consider how they are modified.

From the lexical domain of laughter, as categorised by the *HTE*, we have selected as search items a set of lemmas that occur in our corpus (Table 2). The lemma *laugh** searches for all forms of the verb, *laughs/laughed/etc.* plus the nouns *laugh* and *laughter*. A number of relatively common expressions, such as *guffaw*, *shake with laughter*, *tee-hee* do not occur in any of the novels in our corpus, and so they have been omitted from the table. The expression *split my sides* is uttered as an oath by a pirate in *Treasure Island*, but it does not seem to be related to laughing. Table 2 summarises the selected search items plus their senses. Further details about these expressions are given in Appendices 1 and 2.

Our exploration of the lexical domain of laughter in the corpus of thirteen novels consisted of a series of searches using the text analysis tool, AntConc version 3.5.8 (Anthony 2019). The first set of searches addressed the frequency and distribution of the *HTE* search items in each of the novels, and measured the ‘keyness’ of the items, that is, whether their frequency of occurrence in any given novel is statistically higher or lower than might be expected from an analysis of the corpus as a whole. The findings of these searches indicate the degree to which each novel draws upon the lexical domain of laughter – and, if so, where. *Alice’s Adventures in Wonderland* is remarkably devoid of laughter, as we have already noted, while *Little Women* and *The Secret Garden* overflow with it. The occurrences in the other novels fall in between.

⁴ As is conventional, we indicate an unacceptable usage by placing an asterisk before it, e.g. **hearty snigger*. An asterisk at the end of an item, such as *laugh**, indicates that it is a lemma, or ‘wild card’ used in corpus searches to identify instances not only of *laugh* but also of *laughs*, *laughed*, *laughing*, *laughter*, etc.

Table 2. Search items from the lexical domain of *laughter*

Search item	Meaning
laugh*	the expression of pleasure or derision through sounds, bodily movement and/or facial expressions; an instance of such expression; to express pleasure or derision in this way.
roar* snort* ha ha*	(give) an outburst of loud laughter
die* with/of laughing convulsion*	(give) an outburst of immoderate or wild laughter
giggle* titter*	(give) a giggle
sneer*	(give) a foolish or mocking laugh
whinny*	(give) a laugh in the manner of a horse
snicker* snigger* cackle*	(give) a snigger
chuckle* chortle*	(give) a chuckle

The second set of searches focuses on the lemma *laugh** and considers its lexical contexts in the corpus. The lexical company kept by the forms of *laugh** in each of the texts indicates whether laughter and laughing are semantically positive or negative concepts in the novels. Finally, the discussion reviews the findings and considers how each of the novels draws upon the conceptual field of laughter as part of the moral universe of these 'Golden Age' narratives.

4. Frequencies, distributions and keyness

Appendix 3 shows the raw frequency of each search item in the corpus, that is the number of times the lemma appears in each text, plus the normalised frequency, which refers to the number of occurrences per 1,000,000 words. Normalising the frequencies allows for comparison between texts that differ

in length. The shortest text contains 10,021 word tokens (*Alice's Adventures in Wonderland*) and the longest contains 191,196 word tokens (*Little Women*). Normalised frequencies are used as the basis for the charts shown in the present section of this article; figures have been rounded to the nearest whole number, for clarity of presentation.

Some broad preliminary observations can be made on the basis of this data. Table 3 shows a chart of the normalised frequency of *laugh** in the individual texts in the corpus. As we have already observed, it is a curious fact that no items from the lexical domain of laughter appear in *Alice's Adventures in Wonderland*. Only when she goes through the looking glass in the second novel does Alice, along with several other characters, laugh. By contrast, in *The Secret Garden* and *Little Women*, one or another of the forms of *laugh/laughter* appears over 1200 times per million tokens. As the chart shows, the other novels contain a varying frequency of occurrences of the forms.

Table 3. Normalised frequency of *laugh** in the novels (per million words)

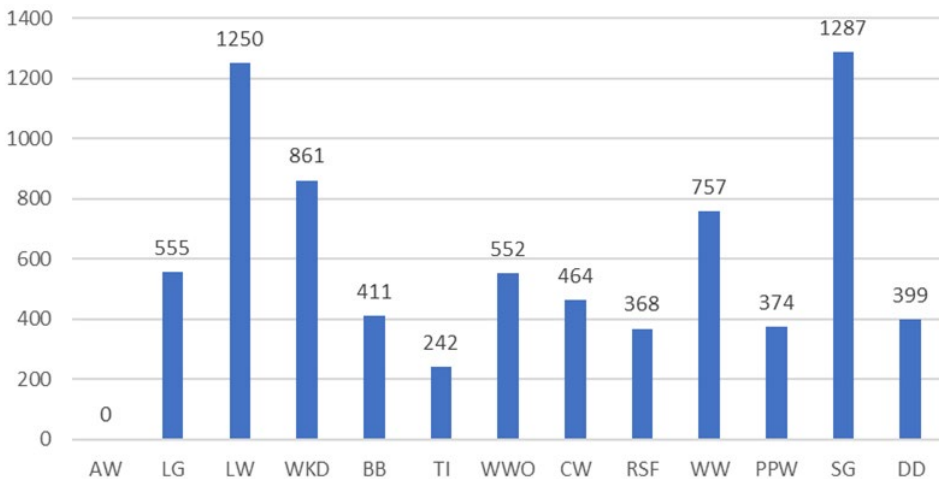


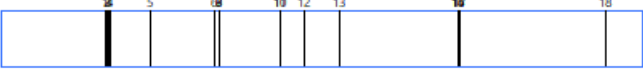
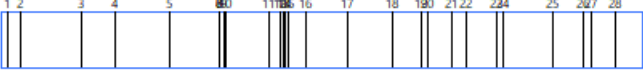
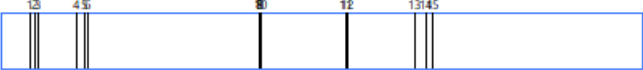

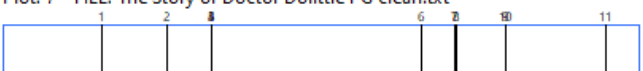
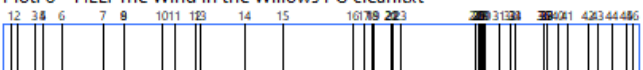
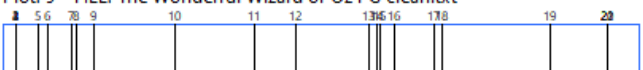
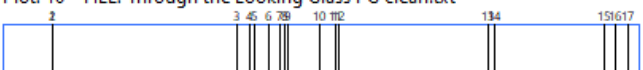
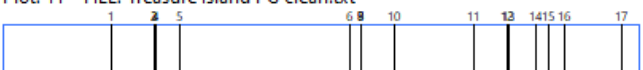



Table 4 shows what AntConc labels the 'concordance plot' of the lemma *laugh** in the twelve novels in which forms of *laugh/laughter* appear (*Alice in Wonderland* is not included, as there are no occurrences). This plot shows the distribution of the expressions of laughter in the novels. This table adds some further detail to the frequency data; for example, while *The Secret Garden* has more occurrences than *Little Women*, the distributions differ to some extent.

Table 4. Concordance plots of *laugh** in the corpus

<p>Plot 1 FILE: Black Beauty The Autobiography of a Horse PG clean.txt</p> 	<p>Hits: 25 Chars: 303978</p>
<p>Plot 2 FILE: Little Women PG clean.txt</p> 	<p>Hits: 239 Chars: 1013162</p>
<p>Plot 3 FILE: Peter Pan and Wendy PG clean.txt</p> 	<p>Hits: 18 Chars: 256966</p>
<p>Plot 4 FILE: Rebecca of Sunnybrook Farm PG clean.txt</p> 	<p>Hits: 28 Chars: 411265</p>
<p>Plot 5 FILE: The Call of the Wild PG clean.txt</p> 	<p>Hits: 15 Chars: 175566</p>
<p>Plot 6 FILE: The Secret Garden PG clean.txt</p> 	<p>Hits: 107 Chars: 428628</p>
<p>Plot 7 FILE: The Story of Doctor Dolittle PG clean.txt</p> 	<p>Hits: 11 Chars: 142745</p>
<p>Plot 8 FILE: The Wind in the Willows PG clean.txt</p> 	<p>Hits: 46 Chars: 325654</p>
<p>Plot 9 FILE: The Wonderful Wizard of Oz PG clean.txt</p> 	<p>Hits: 22 Chars: 208370</p>
<p>Plot 10 FILE: Through the Looking Glass PG clean.txt</p> 	<p>Hits: 17 Chars: 162174</p>
<p>Plot 11 FILE: Treasure Island PG clean.txt</p> 	<p>Hits: 17 Chars: 364306</p>
<p>Plot 12 FILE: What Katy Did PG clean.txt</p> 	<p>Hits: 44 Chars: 267892</p>

The apparent degree of difference in the distributions is perhaps exaggerated by the fact that the concordance plots are normalised in size, and *Little Women* contains more than double the number of word tokens that *The Secret Garden* does. Even so, in the latter novel, the occurrences of laughter come in bunches or waves, while in the former, laughter saturates much of the novel, tailing off towards the conclusion.

A final measure of the frequency of *laugh/laughter* in the corpus is 'keyness,' which is a measure of the extent to which the frequency of occurrence of the forms in one text or set of texts is statistically more or less what might be expected when considered with reference to a larger corpus. The measure of 'keyness' of *laugh/laughter* thus indicates the extent to which it is being used *unusually* frequently or infrequently in a particular text, compared to that in other texts. Obviously, much depends on the reference corpus to which the individual novel, in our case, is being compared. For the purposes of this study, we have taken the thirteen novels as a whole to be the reference corpus against which the individual novels are compared. Since AntConc does not lemmatise words in a key word search, we have focused on measurements of the keyness of three main word-forms, namely *laugh*, *laughed*, and *laughing*. There are different statistical ways of calculating keyness; in this study we have used a log likelihood measure. With this measurement, a value above +6.63 is generally considered to be statistically significant; that is, if the keyness exceeds that value, we can confidently assert that the word-form is used unusually frequently in the given text. A negative value suggests that, where it is used, the form is unusually infrequent. Again, *Alice's Adventures in Wonderland* has not been given a score as no forms of the word appear in the text.

Table 5. Keyness values for *laugh** in each of the novels

	AW	LG	LW	WKD	BB	TI	WWO	CW	RSF	WW	PPW	SG	DD
Laugh	-	-0.11	+33.24	-0.02	-11.54	-5.75	-1.04	-5.18	-16.58	-5.06	-1.44	+3.65	-1.96
Laughed	-	-1.16	+11.88	+0.24	-3.12	-12.53	-0.51	+0.31	-0.63	-1.42	-7.39	+7.35	-2.93
Laughing	-	-0.06	+1.71	+1.18	+0.06	-13.52	-0.78	-1.7	-11.84	+5.24	-7.38	+8.7	0

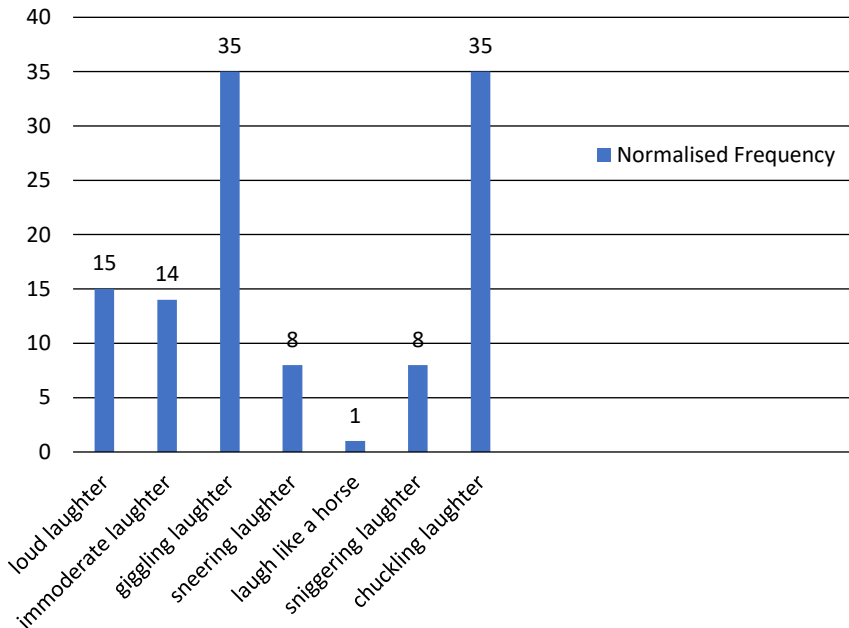
The main finding in the key word calculation is that the main lexical forms associated with *laugh* are very unusually frequent in *Little Women*, their keyness scores being considerably higher even than those of *The Secret Garden*. By comparison, there is a relatively infrequent use of the forms of *laugh* in *Rebecca of Sunnybrook Farm*, and *Treasure Island*, where keyness

values are negative. In most of the other novels, the values suggest neither a significantly frequent nor significantly infrequent use of the forms with respect to the corpus as a whole.

The frequency of occurrence of the forms of the lemma *laugh** by far exceeds that of any of the other items in the lexical domain of laughter. As can be seen in Appendix 3, there are only sporadic uses of the other items in the lexical domain. *Little Women* again has the widest range of expressions within the lexical domain (*ha ha*; *die of laughing*, *convulse/convulsion*, *giggle* and *chuckle*) but even the most common of these (*convulse/convulsion*) has a normalised frequency of only 36.6 per million words, compared to 1250 per million for *laugh**.

A close study of Appendix 3 prompts a number of noteworthy observations: within our corpus, *Treasure Island*, *Peter Pan and Wendy* and *The Call of the Wild* have a near monopoly on *sneer**; while *The Wind in the Willows* and *The Secret Garden* have a high normalised frequency of *chuckle**. However, the frequencies are too small for an analyst to make much of. It is more illuminating to group the lexical items in the corpus as a whole according to the semantic categories suggested by the *HTE* (Table 6).

Table 6. Types of *laughter* in the corpus – normalised frequency (per million words)



Admittedly, when categorising some of the items used in the calculations in Table 6, there is a degree of subjectivity on the part of the analyst which affects the results when the occurrences are so low. Some of the items in the lexical domain of laughter are associated with animal or bird sounds (*roar, snort, whinny, cackle*) and it is a staple of children's fiction that animals can be major characters. Thus, in *The Wizard of Oz*, the Cowardly Lion roars (but never with laughter), Beauty in *Black Beauty* snorts and whinnies (but only once with joy), and Polynesia, the parrot in *The Story of Doctor Dolittle*, cackles. Concordance analysis has been used to disambiguate some of these usages, and, unless there is an explicit indication that the roaring, snorting, or whinnying was an expression of pleasure or derision, then the item was excluded from the figures shown.

Overall, what the present section indicates is how the general conceptual domain of laughter is lexicalised in English and how the members of the lexical domain are realised in a corpus of children's novels. So far, we have seen that the lexical domain of laughter can be completely ignored (as in *Alice's Adventures in Wonderland*) or it can be so frequently utilised (as in *Little Women*) as to be an obvious thematic element. The children's novels also show a relatively high frequency of loud and immoderate laughter, giggling and chuckling – there is clearly a concern with laughter as something irrepressible and uncontrollable, or as a signal of quiet contentment. However, expressions indicating other forms of laughter – sneering, sniggering or horse-like – are relatively less in evidence.

The analysis so far has indicated the relative presence and absence of members of the lexical domain of laughter in the English language in thirteen novels from the first 'Golden Age' of children's literature; we now turn to a more detailed examination of the uses of laughter in the texts.

5. Collocates and concordances

The collocation of a word is, to quote John Firth, "the company it keeps" (Firth 1957: 11; see also Bartsch 2004; Sinclair 1991). Those who study semantic preference suggest that the habitual presence of a frequently occurring collocate can impact on the meaning of a word; for example, they argue that since *regime* frequently collocates with modifiers such as *brutal* and *repressive*, the concept of a *regime*, in contrast to, say, that of a *government*, begins to acquire negative connotations. The implications drawn from this claim can be contentious (cf. Hunston 2007; Stewart 2010), but the fact

remains that habitual co-occurrence among words may make salient certain aspects of their potential meaning. In a study of the meanings and uses of the expression *Irish* in a corpus of British parliamentary discourse, Corbett (2021) argues that diachronic corpus stylistics can illustrate the dynamics by which the association of a particular term with a positive or negative attitude can be reinforced or challenged.

We have seen in the foregoing sections that the members of the lexical domain of *laugh/laughter* enable English speakers to make salient the particular, embodied, and relational aspects of laughing and laughter: its volume, its spontaneous and uncontrollable nature, its animal-like qualities, its potentially derisive import, and whether or not it is concealed. We would expect habitual collocates of *laugh/laughter* to express similar meanings, and to extend them. That is, collocations such as *burst out laughing* express the spontaneous, uncontrolled nature of the behaviour, while *laugh at/with* express the relational meanings of mockery or solidarity and empathy. Further collocations such as *frank/hollow laughter* extend the meanings of laughter to encompass concepts like sincere/insincere laughter, for which no single word is available in English.

There are two main ways of analysing co-occurrence using corpora. The researcher can look at a concordance line, that is, a stretch of text that spans a given number of words on either side of the search item or node. This technique is used when the analyst wishes to identify and interpret patterns manually, and we do this with some selected data below. However, when dealing with a large amount of textual data, programs like AntConc also search automatically for frequently occurring items on either side of a given node. The results of such searches depend on the span selected, and they usually show two types of result: the *frequency* of co-occurrence, which is self-explanatory, and the *strength* of the collocation. The latter value is a statistical measure of the likelihood or probability of co-occurrence, which can be calculated in a number of ways; however, no matter how it is calculated, the probability value depends on the overall collocational behaviour of items in the corpus. Thus, for example, the words *hollow* and *honest* might both modify *laughter* only once in the corpus. Their frequency of co-occurrence would obviously be equal. However, if *honest* modifies a number of other words in the corpus, while *hollow* is largely reserved for *laughter*, the strength of collocation between *laughter* and *hollow* will be greater than that between *laughter* and *honest*. The same statistical calculation often means that a less frequently occurring collocate has a higher collocational

strength. The analyst thus needs to pay attention to the values of both frequency and strength of collocation, since an understanding of both is necessary in order to understand how collocates are behaving in a corpus. The measure of collocational strength used in the analyses below is Mutual Information (MI) and it is conventionally assumed that an MI score of 3 or above indicates that there is a statistically strong bond between the items in question.

Tables 7 to 10 show the results of several collocation searches for members of the lemma *laugh** in the corpus as a whole. Table 7 is given largely to illustrate the principle of collocation by showing ways of describing instances of laughter in the corpus, through a search for expressions that occur one item to the left of the phrase *of laughter*. The collocates listed indicate, again, the nature of laughter in the corpus, namely that it is sudden, involuntary and noisy. It can be violent (*explosions, convulsions*) but can have pleasant associations with, say, the pealing of bells. Of the 10 collocates listed, the first 5 have a statistically strong association (MI of 3 or above) with *of laughter*. However, of these 5, only *peals* occurs more than once in the corpus – and both times in *Little Women*, which is also the source of the singular occurrence, *peal of laughter*.

Tables 8 to 10 show a number of further searches that explore other collocational aspects of the lemma *laugh**. In these searches, a stop-word list has been used to exclude common grammatical items from the results, and the spans, though narrow, are intended to identify salient modifiers, mainly adjectives and adverbs. Table 8 shows collocates occurring one word before *laughter*, Table 9 shows adjective collocates occurring one word before *laugh*, and Table 10 shows adverbial collocates occurring one word after *laughed*. All are ranked according to MI, and only those MI values above 3 are shown in the tables.

As expected, the collocates of *laugh/laughter* confirm what we know about the basic nature of the conceptual field from a study of members of the lexical domain as represented in the corpus of children's writing (e.g. the volume of laughter can be low or uproarious, and the quality of laughter can be pleasant or disagreeable), but they also add further detail.

The collocates describing *laugh/laughter* in these tables fall into certain positive or negative thematic groups, as shown in Table 11, which groups the collocates that are listed in Tables 8-10 according to their semantic themes.

Table 7. Collocates immediately to the left of of *laughter*, sorted by MI

Rank	Frequency	MI	Collocate
1	2	5.7	peals
2	1	4.7	explosions
3	1	4.7	convulsions
4	1	4.1	shouts
5	1	4.1	explosion
6	1	2.9	peal
7	1	2.5	shrieks
8	3	1.2	burst
9	1	0.8	scream
10	1	-2.9	full

Table 8. Collocates immediately to the left of *laughter*, sorted by MI

Rank	Frequency	MI	Collocate
1	1	11.0	uncontrollable
2	1	11.0	convulsive
3	1	10.7	boisterous
4	1	10.3	mocking
5	1	9.9	suppressed
6	1	8.7	childish
7	1	8.4	boyish
8	1	7.7	hollow
9	1	7.5	careless
10	1	7.4	honest
11	1	5.4	secret
12	1	5.0	low

Table 9. Adjective collocates immediately to the left of *laugh*, sorted by MI

Rank	Frequency	MI	Collocate
1	1	10.9	sardonic
2	1	10.9	irresistible
3	1	10.9	heartier
4	1	9.3	sonorous
5	1	8.6	reckless
6	1	8.1	mocking
7	4	7.9	haughty
8	1	7.7	mellow
9	1	7.7	suggestive
10	1	7.7	hearty
11	1	6.7	hoarse
12	1	6.4	disagreeable
13	3	6.2	merry
14	1	6.0	frank
15	1	5.9	winged
16	2	5.3	comfortable
17	3	4.9	short
18	1	4.8	jolly
19	1	3.4	pleasant

Table 10. Adverb collocates immediately to the right of *laughed*, sorted by MI

Rank	Frequency	MI	Collocate
1	1	9.7	uproariously
2	1	9.7	harshly
3	4	8.4	outright
4	1	9.1	hysterically
5	1	7.5	noiselessly
6	1	7.2	scornfully
7	4	7.0	aloud
8	3	6.7	heartily

Table 11. Modifying collocates of *laugh** sorted thematically

Theme	Collocate	
	positive	negative
<i>Control</i>		hysterically reckless convulsive uncontrollable boisterous irresistible careless
<i>Honesty/sincerity/ transparency</i>	outright honest aloud frank	suppressed hollow secret suggestive
<i>Engagement</i>	hearty/heartier/heartily	haughty
<i>Derision</i>	sardonic mocking scornfully	
<i>Maturity</i>		childish boyish
<i>Volume</i>	uproariously sonorous	low hoarse noiselessly
<i>Contentment</i>	comfortable mellow	
<i>Pleasure</i>	pleasant jolly merry	disagreeable harshly
<i>Duration</i>		short

5. From distant to close reading

Up to this point, then, we have selected a group of expressions related to laughter, drawing upon the lexicographical resources of the *HTE*, and we have used this group as the basis for a number of corpus searches. The findings reported in the sections above show the frequency and distribution

of members of the lexical set throughout our corpus, and further searches of the collocates of *laugh(ed)* and *laughter* have shown how the items are modified in the novels. In this section, we offer a number of interpretations that these findings would support, with extracts from the novels. Effectively, this section, then, marks a shift from distant to closer reading.

Insofar as patterns of lexical frequency and distribution and the choice of collocates suggest themes, the findings set out in the previous sections are highly suggestive. The collocates of *laugh** in Table 11 accord with the earlier tables in showing laughter in the novels as indicative of a lack of control and care. This attitude towards laughter is not necessarily negative: if children's literature of the 'Golden Age' is characterised in part by adult nostalgia for the carefree, unrepressed period of youth (when laughter might be 'childish' or 'boyish'), then convulsive, even hysterical laughter can be indicative of a stage in life when pleasure may be unconstrained and undisciplined. This view of laughter is made explicit in *The Secret Garden*:

- (1) It seemed actually like the **laughter** of young things, the uncontrollable **laughter** of children who were trying not to be heard but who in a moment or so—as their excitement mounted—would burst forth.

A nostalgic yearning for a time of unconstrained pleasure is most evident in *Little Women*, where laughter is frequently a marker of communal pleasure and solidarity, and it can often send the person who laughs into fits. Laughter is, for example, one of the phenomena that mark holidays such as the annual apple-picking and Christmas:

- (2) Everybody **laughed** and sang, climbed up and tumbled down. Everybody declared that there never had been such a perfect day or such a jolly set to enjoy it, and everyone gave themselves up to the simple pleasures of the hour as freely as if there were no such things as care or sorrow in the world.
- (3) As Christmas approached, the usual mysteries began to haunt the house, and Jo frequently **convulsed** the family by proposing utterly impossible or magnificently absurd ceremonies, in honor of this unusually merry Christmas.

The thematic elements of honesty and engagement are more evenly balanced between positive and negative connotations, and they suggest disciplined laughter as a moral good: *suppressed* laughter is counterbalanced by *outright*

laughter, *hollow* laughter by *honest* laughter, *haughty* laughter by *heartly* laughter and so on. The social values attached to the positive connotations (e.g. *honesty*, *heartiness*) can be seen as educative for youthful readers as they mature from childhood into adulthood: since laughter is spontaneous and not subject to control, it functions to reveal aspects of moral character, good and bad.

The theme of derision involves the collocates *sardonic*, *mocking*, *scornfully* as well as lexical items such as *sneer*. These terms represent a relational aspect to laughter, the fact that people can laugh derisively or contemptuously at others and at events. These acts of laughter signify the superior attitude that derives from the person who is laughing having power or moral authority. The expression of such power and authority through laughter may index flawed character or villainy (unsurprisingly, the pirates in *Peter Pan and Wendy* and *Treasure Island* sneer), although it is equally possible for a character, such as the doctor in *Treasure Island*, to sneer at a pirate, or a parent to laugh scornfully at a child's perceived foolishness. Peter Pan can also sneer at the laws of nature and, in the same novel, the mocking laughter of mermaids can be directed at the inadequacies of those confined to land.

Other positive and negative collocates of laughter indicate character through engagement (*heartly* versus *haughty* laughter), volume of laughter (*uproarious* versus *low*), degree of contentment (*mellow* laughter) and whether or not it signifies or causes pleasure (*disagreeable/harsh* versus *pleasant/merry* laughter). Where binary choices are available, youthful readers are socialised through their reading to recognise positive and negative character traits by the description of styles of laughter produced in the texts. While, in *Peter Pan and Wendy*, laughter is not particularly frequent, its moral nature is emphasised. One of the tasks Wendy gives to her younger brothers to remind them of home is to describe their parents' laughter. Moreover, when Hook devises a plan to lure the Lost Boys to their doom by using a cake as bait, his true nature is revealed through his laughter:

- (4) 'They will find the cake and they will gobble it up, because, having no mother, they don't know how dangerous 'tis to eat rich damp cake.' He burst into **laughter**, not hollow **laughter** now, but honest **laughter**. 'Aha, they will die.'

The sinister import of this exchange is intensified by the fact that Hook's shift from insincere to sincere laughter coincides with his acknowledgement that the plan will lead to the children's death.

The major counter to our argument that the presence of laughter in children's literature socialises youthful readers into a recognition and adoption of adult social values is, of course, *Alice's Adventures in Wonderland*, in which there is no lexical trace whatsoever of any kind of laughter at all. Given that the novel is comic, this fact is perhaps surprising. However, it could be argued that laughter in this novel is displaced from the text to the reader, who is invited to observe Alice and her curious encounters in Wonderland, and to be amused by the eccentricities on display. By the time Alice goes through the looking glass in the sequel, however, she is allowed to laugh, no fewer than ten times. Arguably, the reader is now invited to share Alice's laughter and thus have a more empathic relationship with Alice than in the preceding novel. As Kramer (2012: 289) observes: "The contagious factor of laughter is relevant to a discussion on empathy and its role in intersubjectivity as a mechanism to bring people together in shared experience".

In *Through the Looking Glass* Alice's laughter functions in part to model the reader's response to the situations she is encountering. For example, when the king is unable to mount his horse without falling over, she recommends that he acquire a wooden one with wheels. He then asks her if such a horse goes smoothly.

- (5) 'Much more smoothly than a live horse,' said Alice, with a little scream of **laughter**, in spite of all she could do to prevent it.

The laughter in this episode indicates to young readers that the slapstick tumbling of the figure of authority from his horse is funny, but also that the uncontrollable scream of laughter that they might share with Alice is something that should ideally be disciplined or even prevented. In *Through the Looking Glass*, as in most of the other novels in the corpus, the controlling of laughter and the expression of proper kinds of laughter are associated with the child's conformity with the social norms of the adult world.

6. Concluding comments

In the spirit of distant reading, then, leavened with a necessary scepticism about the very categories our data produces, we offer the following comments on the significance of the findings outlined in the sections above. The categories of laughter that we have explored are based, first of all, on two

extensive lexicographical works of reference, each of which took generations of intellectual labour to produce: the *OED* and *HTE*. Despite their justifiable reputation as authoritative works of reference, neither is completely consistent or entirely comprehensive. Even so, the lexicographical references offer, we argue, a marvellously rich point of departure for corpus stylistics.

Once the search items were selected, drawing on the lexical domain of *laughter* as delineated by the *HTE*, the text analyses produced a set of results that required interpretation. No interpretation can simply be ‘read off’ the tables and graphs we have produced (cf. Fish 1980); each individual reading demands some previous knowledge of what the items might mean and assumptions about how and why they might be used in the texts. Even so, we were surprised by some of the findings and consequently revised our understanding of the uses of laughter in the texts. We were puzzled by the total absence of any lexicalisations of laughter in *Alice’s Adventures in Wonderland*; we also assumed that there would be higher numbers of words expressing specific types of laughter across the novels analysed. We expected more *sneering*, *cackling* and *giggling* than we found. What we did discover was evidence of the ways in which different novels draw upon the lexical domain of laughter in their narratives, and how the expressions contribute to readers’ sense of community and morality. Our findings ended up affirming our initial hypothesis that laughter in children’s literature is the echo-chamber of the soul.

REFERENCES

Sources

- Alcott, L.M.
1868 *Little Women*. <https://www.gutenberg.org/ebooks/514>, accessed September 2020
- Barrie, J.M.
1911 *Peter Pan and Wendy*. <https://www.gutenberg.org/ebooks/26654>, accessed September 2020
- Baum, L.F.
1900 *The Wonderful Wizard of Oz*. <https://www.gutenberg.org/ebooks/55>, accessed September 2020
- Books in Children’s Literature*
<https://www.gutenberg.org/ebooks/bookshelf/20>, accessed September 2020

- Burnett, F.H.
 1911 *The Secret Garden*. <https://www.gutenberg.org/ebooks/113>, accessed September 2020
- Carroll, L.
 1865 *Alice's Adventures in Wonderland*. <https://www.gutenberg.org/ebooks/928>, accessed September 2020
 1871 *Through the Looking Glass*. <https://www.gutenberg.org/ebooks/12>, accessed September 2020
- Coolidge, S.
 1872 *What Katy Did*. <https://www.gutenberg.org/ebooks/8994>, accessed September 2020
- Grahame, K.
 1908 *The Wind in the Willows*. <https://www.gutenberg.org/ebooks/289>, accessed September 2020
Historical Thesaurus of English
<http://historicalthesaurus.arts.gla.ac.uk>, accessed September 2020
- Kay, C. et al. (eds.)
 2009 *Historical Thesaurus of the Oxford English Dictionary*. Oxford: Oxford University Press.
- Lofting, H.
 1920 *The Story of Doctor Dolittle*. <https://www.gutenberg.org/ebooks/501>, accessed September 2020
- London, J.
 1903 *The Call of the Wild*. <https://www.gutenberg.org/ebooks/215>, accessed September 2020
Oxford English Dictionary (OED online)
<http://www.oed.com>, accessed September 2020
- Sewell, A.
 1877 *Black Beauty*. <https://www.gutenberg.org/ebooks/271>, accessed September 2020
- Stevenson, R.L.
 1883 *Treasure Island*. <https://www.gutenberg.org/ebooks/120>, accessed September 2020
- Wiggin, K.D.
 1903 *Rebecca of Sunnybrook Farm*. <https://www.gutenberg.org/ebooks/498>, accessed September 2020

Special studies

- Anderson, W. – E. Bramwell – C. Hough
 2016 *Mapping English Metaphor through Time*. Oxford: Oxford University Press.
- Anthony, L.
 2019 *AntConc* (Version 3.5.8). Tokyo: Waseda University. Available from <https://www.laurenceanthony.net/software>, accessed September 2020

- Bartsch, S.
2004 *Structural and Functional Properties of Collocations in English: A Corpus Study of Lexical and Pragmatic Constraints on Lexical Co-occurrence*. Tübingen: Gunter Narr Verlag.
- Busse, B.
2012 "A celebration of words and ideas: The stylistic potential of the *Historical Thesaurus of the Oxford English Dictionary*", *Language and Literature* 21 (1), 84-92.
- Carpenter, H.
1985 *Secret Gardens: A Study of the Golden Age of Children's Literature*. London: Faber & Faber.
- Corbett, J.
2021 "Terminology and the evolution of linguistic prejudice: The conceptual domain of 'Irishness' in the *Historical Thesaurus of English* and the *Hansard Corpus of British Parliamentary Speeches*", *TradTerm* 37 (2), 515-537.
- Darton, F.J.H.
1982 *Children's Books in England: Five Centuries of Social Life* (3rd edn.). [1932] Cambridge: Cambridge University Press.
- Firth, J.R.
1957 *Papers in Linguistics, 1934-1951*. Oxford: Oxford University Press.
- Fish, S.
1980 *Is There a Text in this Class? The Authority of Interpretive Communities*. Cambridge, Mass.: Harvard University Press.
- Hunston, S.
2007 "Semantic prosody revisited", *International Journal of Corpus Linguistics* 12 (2), 249-268.
- Kramer, C.A.
2012 "As if: Connecting phenomenology, mirror neurons, empathy, and laughter", *PhaenEx* 7 (1), 275-308.
- Moretti, F.
2005 *Graphs Maps Trees: Abstract Models for Literary History*. London: Verso.
- Oster, U.
2010 "Using corpus methodology for semantic and pragmatic analyses: What can corpora tell us about the linguistic expression of emotions?", *Cognitive Linguistics* 21 (4), 727-763.
- Sinclair, J.
1991 *Corpus, Concordance, Collocation*. Oxford: Oxford University Press.
- Sorby, A.
2011 "Golden Age." In: P. Nel – L. Paul (eds.) *Keywords for Children's Literature*. New York: New York University Press, 96-99.
- Stewart, D.
2010 *Semantic Prosody: A Critical Evaluation*. London; New York: Routledge.

APPENDIX 1

Laugh (intransitive verb) in a specific manner (*HTE* 02.04.10.11.01vi)

Lexical item	Currency	Meaning	<i>HTE</i> code
roar	1815-	Laugh loudly/ coarsely	02.04.10.11.01.01vi
snort	1825-		
ha-ha	1320-		
die with/ of laughing	1596-	Laugh convulsively, immoderately	02.04.10.11.01.02vi
giggle	1609-	Giggle	02.04.10.11.01.03vi
titter	1619-		
sneer	1683-	Laugh foolishly	02.04.10.11.01.05vi
whinny	1825-	Laugh in the manner of a horse	02.04.10.11.01.08vi
snicker	1694-	Snigger	02.04.10.11.01.09vi
snigger	1706-		
cackle	1712-		
chuckle	1803-	Chuckle	02.04.10.11.01.10vi
chortle	1871-		

APPENDIX 2

Types of laughter (HTE code 02.04.10.11.01n)

Lexical item	Currency	Meaning	HTE code
roar	1778-	instance/outburst of loud/coarse/ immoderate laughter	02.04.10.11.01.01.01n
ha ha	1806-		
convulsion(s)	1735-	outburst of vehement/ convulsive/ wild laughter	02.04.10.11.01.02.01n
giggling	a1510 +1786-	giggling/tittering	02.04.10.11.01.01.05n
tittering	1657-		
giggle	a1677-		
titter	1728-		
snickering	1775-	sniggering	02.04.10.11.01.01.07n
sniggering	1775-		
snigger	1823-	instance of sniggering	02.04.10.11.01.01.07.01n
chuckling	1820-	chuckling	02.04.10.11.01.01.08n
chuckle	1837-		
chuckle	a1754	instance of chuckling	02.04.10.11.01.01.08.01n
cackle	1856-		

APPENDIX 3

Raw and normalised frequencies of search items in the novels

Text Code	AW		LG		LW		WKD		BB		TI		WVO		CW		RSF		WW		PPW		SG		DD	
	R	N	R	N	R	N	R	N	R	N	R	N	R	N	R	N	R	N	R	N	R	N	R	N	R	N
Raw/Normalised frequencies	0	17	555.2	239	1250	44	860.6	25	410.9	17	242.0	22	551.5	15	463.5	28	368	46	757.1	18	373.6	107	1286.6	11	399	
laugh*	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	16.5	0	0	0	0	0	0	
roar*	0	0	0	0	0	0	0	1	16.4	0	0	0	0	1	30.9	0	0	1	16.5	0	0	0	0	0	0	
snort*	0	0	0	4	20.9	0	0	1	16.4	0	0	0	0	0	0	0	0	0	0	0	0	0	0	3	108.8	
ha ha*	0	0	0	1	5.2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
die* with/of laughing	0	0	0	7	36.6	3	58.7	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
convuls*	0	0	0	4	20.9	9	176.0	0	0	0	0	0	0	0	0	1	13.1	6	98.8	0	0	5	60.1	1	36.3	
giggl*	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	13.1	0	0	0	0	0	0	0	0	
titter*	0	0	0	0	0	0	0	0	0	4	56.9	0	0	2	61.8	0	0	0	0	4	83	1	12	1	36.3	
sneer*	0	0	0	0	0	0	0	1	16.4	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
whinn*	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
snicker*	0	0	0	0	0	1	19.6	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
snigger*	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	4	65.8	0	0	0	0	0	0	
cackl*	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	13.1	0	0	0	0	0	0	0	0	
chuckl*	0	0	0	3	15.7	1	19.6	1	16.4	2	28.5	0	0	1	30.9	1	13.1	5	82.3	1	20.8	12	144.3	0	0	
chortl*	0	0	1	32.7	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	

Raw frequency = number of occurrences in the text

Normalised frequency = number of occurrences per 1,000,000 words

Address: JOHN CORBETT, BNU-HKBU United International College, English Language & Literature Studies, Zhuhai, Guangdong, China.

ORCID code: <https://orcid.org/0000-0002-5805-1607>

Address: LI LI, Macao Polytechnic University, Faculty of Languages and Translation, Macao, Macao SAR.

ORCID code: <https://orcid.org/0000-0003-3652-3498>