Cognitive Linguistic methods for literature
A usage-based approach to metanarration and metalepsis

Mette Sjölin
Dylan Glynn
Lund University

Abstract
Can contemporary empirical linguistic techniques be applied to literary analysis? This study takes the theoretical framework of Cognitive Linguistics and applies corpus-driven multivariate analysis to the study of narratorial commentary. The study shows that new quantified techniques, when combined with the theoretical tenets of Cognitive Linguistics, can offer insights into poetics research. More specifically, two theoretical constructs of Cognitive Linguistics, construal and encyclopaedic semantics, make possible the application of multivariate statistics to the study of narratorial stylistics of different authors. The case study focuses on four English novels, two from the 19th century and two from the turn of the 20th century. The novels, G. Eliot’s Silas Marner, J. Austen’s Northanger Abbey, J. M. Barrie’s Peter Pan, and E. M. Forster’s Howards End, are chosen because they all have overt extradiegetic - heterodiegetic narrators who use narratorial commentary in an interesting way but differ in other aspects, such as style and period. The stylistic devices under investigation are metanarration and metalepsis. These literary devices can be realised in many different ways, but the stylistic differences are not easily distinguishable.

1. Introduction
1.1 Narratorial commentary. Metanarration and metalepsis
Narratorial commentary is constituted by those passages of a narrative where the narrator pauses from the actual events of the story or the character’s thoughts and gives the reader his/her comments on what he/she is narrating, making his/her own voice more directly audible to the reader. The narrator can be an extradiegetic - heterodiegetic or an intradiegetic - homodiegetic narrator. In either case, the narrator should be distinguished from the author when, for example, the narrative voice of Northanger Abbey, refers to an ‘I’, that ‘I’ should not be read as ‘Jane Austen’ but as ‘the narrator of Northanger Abbey’, about whom we should be careful of mak-
ing any assumptions on the basis of what we may know of the author. Genette’s term **extradiegetic - heterodiegetic narrator** denotes a narrator who tells the story from outside the story (as opposed to an **intradiegetic narrator** who tells the story from inside the story) and who tells a story which is not about him- or herself (as opposed to a **homodiegetic narrator**, who tells a story which is about him- or herself) (Genette 1972: 251-259). An extradiegetic-heterodiegetic narrator can be either **overt** or **covert**; an overt narrator is, in the words of Keen (2003: 40), a narrator who ‘announces his or her presence through self-reference’ and who ‘can offer commentary, including interpretation of the action, judgments about characters or events, generalizations, and even self-conscious remarks about the narration’.

A concept which is central for understanding the conception of narratorial commentary is **narrative levels**. Every narrative has at least two narrative levels: the **discourse level** (where the narrator resides) and the **story level** (where the story takes place). Narratorial commentary is the level on which narrator and reader meet; narratorial commentary is directed to the reader and contains the narrator’s comments on the story.

In addition to these literary terms, **metanarrative commentary** and **metalepsis** will be key terms in the study. **Metafiction** is fiction that draws attention to its own fictionality, and **metanarration** draws attention to the act of narrating; metanarrative commentary, then, is narratorial commentary which draws attention to the fictionality of the text that it is part of and to the narrator’s act of narrating that text (Keen 2003: 110; Håkansson 2009: 61). More than one definition of ‘metanarrative commentary’ is possible, but this is the definition that will be used in this study. Metalepsis occurs when a narrator or character moves from one narrative level to another in a way that, according to Genette, defies logic – for example when a narrator who has previously been established as extradiegetic enters the story world (Genette 1972: 243-251, Keen 2003: 110-111; Håkansson 2009: 25). Metanarration and metalepsis can be manifested in other forms than narratorial commentary, but narratorial commentary is the only form that will be dealt with here. For reasons of practicality, no distinction will be made between **reader** and **narratée** in this study.
1.2. Cognitive Linguistics and Corpus Linguistics

Cognitive Linguistics (Talmy 1985, Fillmore 1985, Lakoff 1987, Langacker 1987) is a theoretical model of language that places creativity at the centre of language structure. Creative cognitive processes such as construal, subjectification, metaphor, metonymy, and blending are argued to be basic to grammar. From this perspective, language is seen as a web of semantically motivated structures for painting the world as the speaker sees it. It follows that the analytical framework of Cognitive Linguistics has been successfully applied to the study of literary stylistics and poetics (Turner 1996, 2006, Semino & Culpeper 2002, Stockwell 2002, Gavins & Steen 2003). However, in recent years, Cognitive Linguistics has been experiencing a methodological revolution (Geeraerts 2005, 2006, 2010, Croft 2009, Stefanowitsch 2010). Current tendencies within the discipline are moving toward empirical and quantified methods such as corpus-driven and experimental analysis. Extending these new linguistic methods to the study of literature poses problems. Corpus-driven techniques seek to make generalisations based on large numbers of naturally occurring examples and experimental techniques examine the responses of speakers to input stimuli. At first glance, the usefulness of such methods in literary analysis is far from evident. This study attempts to apply recently developed corpus-driven techniques to literary analysis.

Quantitative corpus-driven Cognitive Linguistic research uses either advanced collocation analysis or profile-based usage-feature analysis. The first approach examines formal patterns of co-occurrence and bases generalisations about usage patterns on those formal co-occurrences. In other words, it identifies how often a given form occurs in the corpus, relative to another form (Szmrecsanyi 2006, Hilpert 2008, Gilquin 2009, Wulff 2009). The second method, extracts sets of examples and manually analyses them for semantico-pragmatic, formal, and sociolinguistic usage features. This second method, examines exactly how a given form is used without knowing how common or uncommon it is in a corpus or its frequency of occurrence with other forms (Heylen 2005, Gries 2006, Grondelaers et al. 2007, Glynn 2009, 2010, forthc., Janda & Solovyev 2009, Divjak 2010). The manual analysis, or annotation, of the dataset in is a detailed and laborious process covering sometimes hundreds of usage features concerning differ-

---

1 Testimony to this revolution is the growing number of anthologies devoted to the application of these methods. A small sample of recent anthologies would include Gries & Stefanowitsch (2006), Stefanowitsch & Gries (2006), Geeraerts et al. (2010). Glynn & Fischer (2010), and Glynn & Robinson (in press).
ent grammatical and semantic categories. The manual analysis results in two limitations for the application of this method. Firstly, due to practical constraints, the number of examples analysed is relatively small, limiting both the representativity of the sample and the possibility for the application of statistical treatment of the results. Secondly, the analysis is subjective and is therefore open to the vagaries of different interpretation as well as simple human error. Despite these inherent shortcomings, the method is excellent for the analysis of subtle variation in use, typical of both creative and natural language use. For these reasons, this method is employed in the current study.

The principle of usage-feature analysis is simple – identify patterns of usage through examining hundreds or thousands of usage events. For example, let assume that, sometimes, speakers use a given word in a given tense and aspect to profile a certain semantic relation, but in another tense to indicate something slightly different, if this occurs often enough, by analysing hundreds of examples, this tendency will be visible as a pattern in the data. Patterns can be understood as non-random re-occurrences of correlating usage-features. This approach enjoys growing popularity in the study of semantics from a cognitive perspective.

However, this method, like all corpus-driven methods, relies on frequency and not salience in order to identify structure. That is, the method does not distinguish between the linguistic impact of conceptually or discursively important occurrences versus a quotidian example. It assumes that every language event is of equal significance for the language system. Obviously, this is a shortcoming when analysing natural language, but when examining carefully chosen creative language, this limitation is all the more important. Nevertheless, that does not mean that the application of a such a method to creative language cannot offer important insights.

In order to understand how these quantitative techniques can inform literary analysis, we need to evoke two theoretical constructs basic to the Cognitive Linguistic enterprise. Firstly, research within the paradigm assumes that all linguistic form is semantically motivated and therefore an expression of meaning. However, it is important to note that this meaning is not understood as traditional ‘linguistic’ meaning, but is understood to be encyclopaedic in nature (Fillmore 1985, Lakoff 1987). Linguistic forms, rather than encoding denotational information, are actually used to represent the world-knowledge of speakers. This means information such as context of use (the gender, age, social identity of speakers as well as situation context) is part of the ‘meaning’ of the language event. Assuming this,
contextual information in language utterances can be indicators of the meaning of the utterance, and variation in language form is believed to be there to allow speakers to construe world events differently (Talmy 1985, Langacker 1987). Each linguistic form, and the various combinations of those forms, are complex tools for depicting the same real-world event in different ways, seeing the event from different perspectives and highlighting different dimensions of that event. Assuming that different forms in language are used in subtly different ways to represent the world, is a theoretical position that permits the linguistic analysis of literary texts. Applying this to a usage-feature analysis, means that we can look for indices of how a speaker or writer construes the world by identifying re-occurring patterns in his or her language use. These indices can be semantic, pragmatic, or ‘extra’-linguistic and will be detected as distinctively re-occurring usage-features.

2. Stylistic features and literary data
2.1 Data
The data were taken from four novels with overt third-person extradiesgetic-heterodiegetic narrators: Jane Austen’s Northanger Abbey (finished in 1803, published posthumously in 1818), George Eliot’s Silas Marner (1861), E. M. Forster’s Howards End (1910) and J. M. Barrie’s Peter Pan (1911). From these texts, 125 occurrences of narratorial commentary were extracted manually. This was done by reading the novels and randomly selecting instances of narratorial commentary. The examples were then analysed for a range of linguistic and stylistic features: metanarration, metaalepsis, story tense, world tense, narrator’s tense, second person personal pronoun, first person personal pronoun, reference to the reader, generalisation, interrogation, grammatical mood, reference to the reader’s reaction, irony and the narrator’s referring to themselves as author. We consider each of the usage-features in turn, beginning with the stylistic features and moving to the formal, or grammatical, features.

2.2 Analysis
The first two stylistic features considered are metanarration and metaalepsis. In metanarrative commentary the narrator draws attention to the act of narrating. Example (1) shows a typical metanarrative comment in the narrator’s overt reference to telling the reader the story.
(1) Monday, Tuesday, Wednesday, Thursday, Friday, and Saturday have now passed in review before the reader; the events of each day, its hopes and fears, mortifications and pleasures, have been separately stated, and the pangs of Sunday only now remain to be described, and close the week. (Northanger Abbey)

‘Metalepsis’ denotes a jump between different narrative levels within a single passage. Example (2) represents an instance of metalepsis, as the narrator and reader (‘we’) move into the story and are able to watch the characters.

(2) We will not follow them now; for may there not be some others in this departing congregation whom we should like to see again- some of those who are not likely to be handsomely clad, and whom we may not recognize so easily as the master and mistress of the Red House? (Silas Marner)

The stylistic usage-feature ‘reference to the reader’ indicates a mention of the word ‘reader’ in reference to the reader of the novel, such as in example (3):

(3) It is now expedient to give some description of Mrs. Allen, that the reader may be able to judge in what manner her actions will hereafter tend to promote the general distress of the work (Northanger Abbey)

An interesting feature that was analysed is ‘generalisation’. These instances refer the narrator generalising about the world outside the narrative. This is illustrated by example (4), where the narrator generalises about ‘all men’, not just the characters in the novel:

(4) Favourable Chance, I fancy, is the god of all men who follow their own devices instead of obeying a law they believe in. (Silas Marner)

Another stylistic usage-feature is the reference to ‘reference to literary conventions’. In the following example (5), there is a reference to the literary convention of not introducing a new character near the end of a novel, as well as the narrator’s referring to the main character as the ‘heroine’ and to
the events of the narrative as her ‘adventures’:

(5) Concerning the one in question, therefore, I have only to add – aware that the rules of composition forbid the introduction of a character not connected with my fable – that this was the very gentleman whose negligent servant left behind him that collection of washing-bills […] by which my heroine was involved in one of her most alarming adventures. (Northanger Abbey)

The use of ‘irony was also considered. This stylistic feature is generally considered important in the description of narratorial commentary. Identifying ironic tone can be difficult but since the author typically wanted such references to be read as ironic, close reading of the text makes the analysis reasonably reliable. Example (6) represents what is understood as ironic narratorial commentary:

(6) Her father was a clergyman, without being neglected, or poor, and a very respectable man, though his name was Richard – and he had never been handsome. He had a considerable independence besides two good livings – and he was not in the least addicted to locking up his daughters. Her mother was a woman of useful plain sense, with a good temper, and, what is more remarkable, with a good constitution. (Northanger Abbey)

The ‘narrator’s referring to himself or herself as the author of the work’ was found to be an important stylistic feature. The following example is typical:

(7) One thing I should like to do immensely, and that is to tell her, in the way authors have, that the children are coming back. (Peter Pan)

In the case of third-person extradiegetic-heterodiegetic narrators it can be difficult to distinguish between the narrator and the author, but in narratology this distinction is made even when narrators refer to themselves as authors.

A final stylistic feature concerns the narrator’s reference to or anticipation of the reader’s reaction to the narrative. Example (8) is an instance of reference to the reader’s reaction:
(8) To Margaret – I hope that it will not set the reader against her
– the station of King’s Cross had always suggested Infinity. (How-
ards End)

The remaining features are less stylistically orientated, being of a more
grammatical nature. Firstly, the feature ‘interrogative’ indicates one or
more occurrences of interrogation within the example:

(9) Was Mrs. Wilcox one of the unsatisfactory people – there are
many of them – who dangle intimacy and then withdraw it? They
 evoke our interests and affections, and keep the life of the spirit
dawdling round them. Then they withdraw. When physical passion
is involved, there is a definite name for such behaviour – flirting –
and if carried far enough it is punishable by law. But no law--not
public opinion even – punishes those who coquette with friendship,
though the dull ache that they inflict, the sense of misdirected effort
and exhaustion, may be as intolerable. Was she one of these?
(Howards End)

Secondly, we turn to temporal reference. We analyse three levels of te-
mporal reference – ‘story tense’, ‘world tense’, and ‘narration tense’. This is
because one single example of narratorial commentary may concern more
than one of these levels and use a different tense for each level.

The first level of temporal reference was termed ‘story tense’. This
feature refers to the tense used to describe temporal reference at the story
level - this can be either the past (as in example (10)) or the present. The
future tense is also possible, but the data did not contain sufficient exa-
mples to include its analysis. When no story tense is found, the narrator is
most likely not talking about the story, and is either describing the dis-
course level or the world outside the novel).

(10) There she fell miserably short of the true heroic height.
(Northanger Abbey)

The second level of temporality, the ‘world tense’, designates time refer-
ce outside the narrative, and can be past, present or future. In example
(11), the world tense is the present:
The third level, ‘narration tense’, is the tense used when the narrator talks about the discourse level, about him- or herself and about the reader. This can be past, present or future. Example (12) is an instance of the present narrator’s tense:

(12) Ought the Wilcoxes to have offered their home to Margaret? I think not. (Howards End)

Two kinds of pronominal reference are indicated. The first pronominal usage-feature is termed the ‘second person personal pronoun’ and refers to an occurrence of the second person in reference to the reader, as in the following example (13):

(13) I don’t know whether you have ever seen a map of a person's mind. (Peter Pan)

A second feature is termed ‘first person personal pronoun, singular’ and refers to the occurrence within the example of one or several of the pronouns I, me and my in reference to the narrator, and ‘first person pronoun, plural’ refers to the occurrence of we, us or our in reference to the narrator and the reader. An example of this is the following observation (14):

(14) Now we know why she was prejudiced against the redskins. (Peter Pan)

Grammatical mood, the indicative, the imperative, and conditional, were also annotated. Examples were deemed conditional or imperative if a single occurrence of the mood was found in an example.

3. Results. Quantitative insights into the usage-feature analysis

This section examines the results of the above analysis. Two multivariate statistical techniques are used to identify usage patterns in the dataset extracted and annotated from the four novel; one technique is a purely explor-
atory technique, the another more detailed and confirmatory. It must be remembered that the advantage such techniques pose over both qualitative and non-multivariate quantitative approaches is that they allow us to identify the ‘interaction’ of the different dimensions of usage simultaneously. Moreover, the confirmatory techniques offer important options of result verification and benchmarks for explanatory power.

3.1 Exploratory techniques. Identifying patterns in the texts
The first usage pattern we will investigate is that of the interaction of the use of metanarration, reference to the narrator as author, and reference to literary convention. These usage factors were correlated against the different novels by submitting them to a multiple correspondence analysis. This statistical technique visualises ‘correspondences’ in the data and, due to its visual output that captures tendencies rather than discrete, sometimes, reductionist descriptions, it has been shown to be well adapted to research employing the usage-based model (Szelid & Geeraerts 2008, Glynn in press). The principle behind the technique is simple - all the frequencies of all possible co-occurrences of different usage-features are calculated. These frequencies are then converted to relative distances. In the analyses below, three or four factors are considered simultaneously, this means the distances calculated between the different usage-features should be placed in three and four-dimensional spaces. Obviously, we cannot visualise four-dimensional space and three-dimensional space is impractical for the current purposes, so these n-way spaces are conflated to the two-dimensional plots presented below. Since the plots a two-dimensional representations of n-dimensional spaces, one must take due caution when interpreting them as the dimensional compression can result in misleading visual correlations. To interpret the plots, one considers the relative distance between data points. Each data point represents a usage-feature: proximity indicates high correlation in usage where distance indicates low correlation. Importantly, the points can ‘push each other away’. Thus, a correlated point may seem far from a given point but this is because it is distinctly disassociated with some other point. We will encounter one such instance below in Figure 3.
In Figure 1, we see stylistic correlations that confirm what one would expect of the novels. Two clear clusters of usage-features emerge. The first cluster, labelled (i), consists of the features ‘no reference to narrator as author’ <NoRef_to_Nar_as_Author>, ‘no reference to literary conventions’ <NoRef_toLit-Conv>, and ‘no metanarration’ <NoMetanarration>. This usage-feature cluster encircles the data point for the novel *Silas Marner*. This is in contrast to cluster (ii), where the novel *Northanger Abbey* is clearly correlated with the use of metanarration <Metanarration>, ‘reference to narrator as author’ <Ref_to_Nar_as_Author>, and ‘reference to literary conventions’ <Ref_to_Lit-Conv>. Although the features of ‘reference to the narrator as author’ and ‘reference to literary conventions’ are not necessarily associated with metanarration, their co-occurrence does entail metanarration. The combination of these feature can be seen as an excellent operationalisation of metanarrative commentary and their clustering with the presence of metanarration is to be expected. Most importantly, they distinguish stylistically two novels – *Northanger Abbey* and *Silas Marner*, where the former is more associated with metanarration than the latter. In other words, relative to the other three novels, *Northanger Abbey* uses much more metanarration in its narratorial commentary, especially through the reference to the narrator as author and through references to
literary conventions. This is in contrast to *Silas Marner* where such techniques are less common. This matches accepted readings of both the novels. The other two novels, *Peter Pan* and *Howards End* are not seen to cluster with any of the features, suggesting, that relative to *Northanger Abbey* and *Silas Marner*, they not distinctly metanarrative but neither are they distinctly un-metanarrative. Example (15), from *Northanger Abbey*, is typical of the kind of narratorial commentary in question.

(15) *Concerning the one in question, therefore, I have only to add—aware that the rules of composition forbid the introduction of a character not connected with my fable—that this was the very gentleman whose negligent servant left behind him that collection of washing-bills, resulting from a long visit at Northanger, by which my heroine was involved in one of her most alarming adventures.* (*Northanger Abbey*, p. 138).

The correspondence analysis presented in Figure 1 confirms what is a largely predictable reading. Let us now turn to a second analysis that reveals stylistic patterns less evident using traditional techniques of literary analysis. Figure 2 presents the results of a correspondence analysis of the metalepsis and grammatical mood across the four novels.

Just as for the previous analysis, two clear clusters emerge. Cluster (i), at the top of the plot, shows three novels, *Howards End* and *Northanger Abbey* distinctly correlates with a lack of metalepsis and the indicative mood. This is in stark contrast to the novel *Peter Pan*, which correlates with metalepsis and the imperative mood (ii). Although, *Peter Pan* is known to employ a great deal of metalepsis, that this is expressed through the use of imperatives, relative to the other novels, is less evident, even in close reading. The correspondence analysis brings out this stylistic feature and when one returns to the text with this knowledge, the imperative mood is, indeed, found to be part of the author’s use of metalepsis. By employing the imperative, the narrator addresses the reader and asks him or her to participate in the action. Example (16) captures the kind of narratorial commentary in question.

(16) *Look at the four of them, Wendy and Michael over there, John here, and Mrs. Darling by the fire. There should have been a fourth night-light.* (*Peter Pan*)
The interaction of the grammatical feature of mood and metalepsis showed how one author expresses metalepsis, but we can turn the same analysis to different grammatical features. Let us consider the interaction of metalepsis and the narration tense. Figure 3, below, presents a third correspondence analysis, this time examining this interaction – novel, metalepsis and tense of narration.

In Figure 3, on the left, we have cluster (i), which is the ‘neutral’ clustering of lack of narration tense or metalepsis. Two novels, *Howards End* and *Northanger Abbey* are clearly grouped here. However, to the right of the plot we see *Silas Marner* and *Peter Pan*, which, although not associated in use with regards to both metalepsis and tense, are both associated with metalepsis. Although interesting, traditional close reading would reveal the importance of metalepsis in these novels. What this approach has revealed is how these two novels are differentiated through grammatical differences in the use of metalepsis. *Peter Pan*, although on the metalepsis ‘side’ of the plot with *Silas Marner*, is distinctly associated with the present narration tense in metalepsis where *Silas Marner* is correlated with future narration tense in metalepsis passages. In this interpretation of the plot, it must be remembered that proximity is relative. This means that in Figure 3, *Silas*
Marner is, in fact, more highly associated with future narration than Peter Pan is with present tense. This is because, relative to the other data points, future narration tense is distinctly closer to Silas Marner, the novel being between the tense data point and all the other novels. The correlation between the future narration tense and metalepsis can be seen in example (17):

(17) We will not follow them now; for may there not be some others in this departing congregation whom we should like to see again – some of those who are not likely to be handsomely clad, and whom we may not recognize so easily as the master and mistress of the Red House? (Silas Marner)

![Figure 3. Multiple Correspondence Analysis (Euclidean matrix) Novel, Metalepsis, & Narrator’s Tense](image)

Here, the narrator and the reader are brought into the story world, into the presence of the characters, and are able to observe them directly. The most common narration tense in Silas Marner is the present and the most common story tense is the past. In metalpsis these two levels merge and the same tense is used for the story as for the discourse level. The effect of such a metalepsis is that of being transported into the story, here and now,
which makes the past tense ill-suited. The future tense expresses a certain hypotheticality, indicating what will or will not happen, but also what may happen – as the events that the narrator talks about have not yet taken place, he/she has slightly less control over them than otherwise.

In order to better understand how these grammatical features interact with the stylistic features, and how they profile differences in narratorial commentary, we need to return to the texts and close reading, employing traditional analytical techniques. However, the purposes of this study are to demonstrate the usefulness of the method in revealing the stylistic variation. The interpretation of these findings must, unfortunately, wait for further research.

3.2 Confirmatory techniques. Identifying significance in the patterns
Correspondence analysis is a powerful tool for identifying patterns, but it offers no information on the statistical significance of the findings or the predictive power of the correlations observed. Let us consider a method that will offer confirmatory information on observed patterns. Logistic regression is more complicated to perform than correspondence analysis but the principle is straightforward. The method is used to contrast two possible types, for example, 19C novels from turn of the century novels. For each example, the method examines the different factors analysed and the interaction of those factors and from this, attempts to predict the kind of example. Based on these factors, if the analysis can correctly predict the example, then we can assume the factors in question accurately capture the differences in style.

Several models were run on the data and none of them were able to accurately predict metalepsis or metanarration. There are three non-mutually exclusive reasons for this. Firstly, our grammatical and stylistic analysis was not sufficiently accurate to capture the differences. Secondly, there are no stylistic or grammatical differences associated with the types of narratorial commentary. Thirdly, we had insufficient data for the tendencies observed to be confirmed in terms of statistic probability. It is very likely that a combination of all three reasons are involved. Extending the study to include more data and a wider range of usage-features may offer further insights. However, distinct stylistic differences were found between the two older novels and the two novels from the turn of the century.
Binary Logistic Regression Analysis – Period: 19C vs. fin de siécle

Deviance Residuals:

<table>
<thead>
<tr>
<th></th>
<th>Min</th>
<th>1Q</th>
<th>Median</th>
<th>3Q</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>-2.2506</td>
<td>-0.8663</td>
<td>0.4069</td>
<td>0.7514</td>
<td>2.2174</td>
</tr>
</tbody>
</table>

Coefficients:

| Predicting features | Estimate | Std. Error | z-value | Pr(>|z|) |
|---------------------|----------|------------|---------|---------|
| Metalepsis          | 0.6118   | 1.4547     | 0.421   | 0.67407 |
| Story tense future  | 1.1773   | 1.3431     | 0.877   | 0.38074 |
| Story tense past    | 1.8209   | 0.5771     | 3.155   | 0.00160 ** |
| Story tense present | 0.3604   | 0.7799     | 0.462   | 0.64398 |
| No 1st Pers. Pronoun| 0.3554   | 0.7094     | 0.501   | 0.61634 |
| 1st Pers. Pronoun /us/ /our/ | 1.7560 | 0.9788 | 1.794 | 0.07281 . |
| 2nd Pers. Pronoun /you/     | 3.1787   | 1.2037     | 2.641   | 0.00827 ** |
| Generalisation             | -1.4156  | 0.5560     | -2.546  | 0.01090 * |
| Irony                    | -1.3293  | 0.5323     | -2.498  | 0.01251 * |
| Reference lit. conventions| -2.2918 | 0.7797 | -2.939 | 0.00329 ** |
| Metalepsis interaction with No 1st Personal Pronoun | -3.6096 | 2.1080 | -1.712 | 0.08684 . |
| Metalepsis interaction with 1st Pers. Pronoun /us/ /our/ | -2.3749 | 1.9225 | -1.235 | 0.21671 |

---

Null deviance: 198.44 on 145 degrees of freedom
Residual deviance: 139.20 on 133 degrees of freedom

Frequencies of Responses

d.f.  P  C  Dxy  R2
19C FdS 12 0 0.842 0.684 0.449
71 75

The model presented above is capable of predicting whether an example of narratorial commentary originates from the 19th century or the turn of the 20th century, based on the set of usage features listed. The column on the left lists the relevant factors and features. The column on the far right (Pr(>|z|), lists the p-values, or statistical significance. All lines with stars are statistically significant; their role in predicting the era of an example is, therefore, not chance. Other factors might be important in predicting such an outcome, but based on this dataset, we cannot prove that their behaviour is not simply a coincidence. It must be remembered that this dataset is small and obtaining statistically significant results with small samples is difficult since the larger the sample, the surer one can be that one’s findings are not chance.

Firstly, we identify which features are significant in distinguishing the two periods. In this dataset, the table of coefficients presents the usage-features ‘1st personal pronoun us, our’, ‘2nd personal pronoun you’, and ‘story tense in the past’ as positive predictors of the outcome. This means they are associated with and predict novels from the turn of the century. In the same list, the negative predictors, which are associated with the 19th century, include ‘generalisation’, ‘irony’, ‘reference to literary conven-
tions’, and the interaction of the use of ‘1st personal pronoun us, our’ and metalepsis. The association of generalisation, irony and reference to literary conventions with the 19th century is not surprising, as a close reading of the novels would confirm this, especially the association of generalisation with Silas Marner and of the two other features with Northanger Abbey. However, the association of ‘1st personal pronoun us, our’ and ‘2nd personal pronoun you’ with the later novels is not an expected result. What it seems to indicate is that the narrators more often refer to themselves and the reader as a unit, creating a sense of community between narrator and reader, in the novels from the turn of the century than in the 19th-century novels, and that the reader is directly addressed more often in the later novels. The past story tense is also more closely associated with the novels from the turn of the century, which perhaps indicates that the 19th-century authors were more creative in varying the story tense. The importance of interaction between the ‘1st personal pronoun us, our’ and metalepsis in the 19th-century novels is probably influenced by Silas Marner, where a close reading reveals this connection, but it is possible that this is a more general feature of 19th-century narratorial commentary. In order for these findings to be explained with greater certainty, the association of the features with the different periods needs to be investigated further.

Secondly, having established which features are significant in predicting whether a narratorial comment is typical of one period or the other, we can now determine their relative importance. The numbers in the second column are indicators of relative predictive power. The greater the number indicated, either negatively or positively, the greater the importance of that feature in predicting the outcome as either 19th century or early 20th century. We see immediately that the use of the ‘2nd person pronoun you’ is the most important feature in the metanarrative commentary of the turn of the century period. On the other hand, the interaction of the use of ‘metalepsis’ with a lack of ‘1st person pronouns’ is a stylistic feature that is strong indicator of the 19th century.

Lastly, the model tells how accurately, overall, we have captured the difference between the two periods in question. At the bottom of the table of coefficients, there is a short list of statistics. Especially important is the C-value and the R^2. In linear regression modelling, the R^2 is typically considered the best indicator of predictive strength, however, in logistic regression, it is a pseudo R^2 and is often considered less reliable. Normally, the R^2 should be above 0.3 to be considered representative of a strong model. The R^2, at 0.45, is a clear indicator of predictive strength. Perhaps, more
reliable for this kind of model, but less conventionally used, is the C-value. This statistic can be thought of a kind of percentage of accuracy. The model here can accurately predict over 84% of the outcomes. In other words, a combination of features analysed can determine, with more than two thirds accuracy, the era of narratorial comment an example represents. Having established this, we can say with confidence we have captured the characteristics of the narratorial commentary for the two periods.

It should also be noted that the usage-features marked with a spot rather than an asterisks are on the margin of statistical significance. They are included because these features are interesting from an interpretative point of view. Various models were considered, including models with exclusively features of a statistical significance of \( p<0.05 \), and all models proved predicatively powerful with C-values above 0.8.

4. Summary
The study has shown how corpus-driven Cognitive Linguistics can inform the study of stylistics. The case study focused on two forms of narratorial commentary and found patterns between authors and across periods. Certain hypotheses about narratorial commentary have been confirmed and other facets, previously unconsidered, have been revealed. The study confirms the connection between metanarrative comments, references to literary conventions and references to the narrator as author of the work. With regards to metalepsis, the correspondence analysis showed that it occurs most frequently in Peter Pan and Silas Marner. This, on its own, is not a novel insight. However, we found that the use of the imperative mood was distinctly part of the stylistic choices associated with metalepsis in Peter Pan where in Silas Marner, metalepsis is associated with the narrator’s future tense.

The study also examined changes in the stylistics for narratorial commentary between the 19th century and the turn of the 20th century. A logistic regression analysis identified a range of grammatical features that distinguished the style of narratorial commentary across the two periods. These findings were confirmed and found to be statistically significant and predictive. The most important stylistic differences were found to be in the use of personal pronouns, the lack of the first person pronoun ‘I’ distinctive of the 19th century and the use of the second person ‘you’ distinctive of the beginning of the 20th century.
The next step, of course, is to interpret these findings. For this, a combination of traditional stylistic analysis and the analytical constructs of Cognitive Linguistics need to be employed. For the above examples, it seems that differences in the grammatical profiling of metalepsis warrants investigation. However, in general, the questions — why do these correlations occur and what does such grammatical variation say about the stylistic choices of the different authors — remains to be answered. Further research will add a larger number of examples as well as a wider range of authors and novels but also attempt to answer these questions. This study has, nonetheless, demonstrated the possibility of using the new corpus-driven techniques, applied within the theoretical framework of Cognitive Linguistics, to approach literary questions.
References


Schmid & S. Handl (Eds.), *Cognitive Foundations of Linguistic Usage-Patterns*, Berlin: Mouton. 89-118.


