The Distribution of Deontic Modals in Jane Austen's Mature Novels

Lauren E. Levine
Georgetown University
lel76@georgetown.edu

Abstract

Deontic modals are auxiliary verbs which express some kind of necessity, obligation, or moral recommendation. This paper investigates the collocation and distribution within Jane Austen's six mature novels of the following deontic modals: *must*, *should*, *ought*, and *need*. We also examine the co-occurrences of these modals with name mentions of the heroines in the six novels, categorizing each occurrence with a category of obligation if applicable. The paper offers a brief explanation of the categories of obligation chosen for this investigation. In order to examine the types of obligations associated with each heroine, we then investigate the distribution of these categories in relation to mentions of each heroine. The patterns observed show a general concurrence with the thematic characterizations of Austen's heroines which are found in literary analysis.

1 Introduction

Jane Austen is a celebrated British author whose classic works have endured over the centuries. From 1811 to 1818, her six mature novels: *Northanger Abbey* (NA), *Sense and Sensibility* (S&S), *Pride and Prejudice* (P&P), *Mansfield Park* (MP), *Emma* (E), and *Persuasion* (P) were published, *Northanger Abbey* and *Persuasion* being published posthumously.

In recent years, efforts in the area of digital humanities have grown, and the interest for having accessible tools and methodologies for approaching a quantitative analysis of Jane Austen's writings has increased (Runge, 2019). Analytical techniques such as keyword analysis, phraseological research, and distribution analysis have been performed on the works of Jane Austen in order to gain literary, structural, and linguistic insights into the texts (Fischer-Starcke, 2010). This includes investigation into modal verbs: Wijitsopon (2013) has noted modal auxiliary verbs to be a linguistic feature particularly characteristic of Jane Austen's novels, and

argued that an analysis of clusters containing the modal auxiliary *must* provided evidence of literary critics' claims that Austen's novels are framed more by the internal thoughts and perspectives of the characters rather than by the physical events of the stories. Previously, Burrows (1986) also explored the differing frequency patterns of modal auxiliary verbs, finding differences of statistical significance between different modes of narrative as well as between different characters within Jane Austen's works. There have also been less quantitative discussions of Jane Austen's use of modals, such as Boyd (1984), which offers a close analysis of the character intricacies and social commentary expressed by Austen's masterful use of modals.

We aim to continue such investigation into modal auxiliary verbs in Austen's work, specifically looking at a selection of deontic modals. Deontic modality is a linguistic mode that expresses how the world ought to be relative to some normative standard, such as moral norms, and frequently carries a sense of necessity or obligation. Deontic modals can narrowly be defined as the set of modal auxiliary verbs, such as *can* and *should*, which have the potential to express deontic modality (Carr, 2017). In this paper, we examine the distribution in Jane Austen's mature novels of the following deontic modals: *must*, *should*, *ought*, and *need*.

All six of Jane Austen's mature novels embody themes of obligation, propriety and duty, but the presentation of these themes is not uniform across the novels. While *Pride and Prejudice* has a heroine who balances propriety and wit in Elizabeth Bennet, *Mansfield Park*'s heroine Fanny Price is so obliging to her position in society that the book is often considered to be "moral at the cost of comedy and vigour" (Todd, 2006). Deontic modals can be used by an author to stylistically incorporate themes such as societal obligation and duty into the prose. By investigating the deontic modals that Austen employs in relation to her heroines, we aim
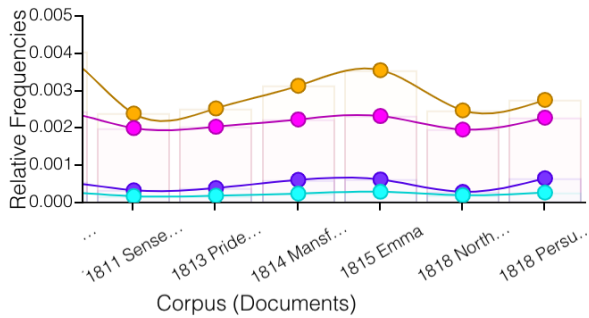
Figure 1: Relative frequencies of deontic modals in Jane Austen's novels. (From top to bottom) Orange: *must*, Pink: *should*, Purple: *ought*, Blue: *need*

Obligation Modal	Top Collocates (Verbs)
must	think, know, said, make, feel
should	think, like, said, know, come
ought	know, think, feel, say, make
need	fear/be afraid, say, tell, ask, think

Table 1: Top verbal collocates of deontic modals.

to characterize how different aspects of the theme of obligation are associated with each of Austen's heroines.

For the purposes of this investigation, we leverage *Voyant Tools*, an open access web-based suite of text analysis tools that allow for easy visualization of corpus data, including relative frequency of terms, collocations, and occurrence contexts (Sinclair and Rockwell, 2021). *Voyant Tools* contains a Jane Austen corpus compiled from digital texts freely available from Project Gutenberg.

Voyant Tools was selected to be used in this investigation for its ability to provide easy access to a full corpus of Jane Austen's works, as well as for the particular analysis functions included in its interface, such as the collocates tool and the relative frequency tool, which provide adequate functionality to engage in meaningful analysis through relatively simple corpus linguistic methods. In addition to providing specific literary analysis, in this paper, we demonstrate how accessible digital tools, such as *Voyant Tools*, allow for researchers without significant technical expertise to leverage relatively simple corpus-based methods of analysis in order to gain literary insights.

2 Relative Frequency of Deontic Modals

The relative frequency of a term in a corpus refers to the ratio between the number of times the term occurs in the corpus and the total number of tokens in the corpus. A high relative frequency is typically taken to be an indicator that the linguistic feature (or term) under investigation is of significance in the corpus (Fischer-Starcke, 2010).

Using *Voyant Tools*, the relative frequencies of the modal verbs *must*, *should*, *ought*, and, *need* were calculated for each of Austen's six novels.

The visualization of this data is presented in Figure 1. We note that *must* and *should* have higher relative frequencies across the board (the highest being *most* in *Emma* with a relative frequency of 0.0035218), reflecting that they are more common words than *ought* and *need*.

The figure proceeds in order of publication of the novels, but it should be noted that the manuscript for *Northanger Abbey* was actually completed around 1803, making it the earliest of Austen's novels. With this in mind, we see that there is a general increase in the relative frequency of all four deontic modals investigated as time moves forward. This could indicate that the use of deontic modals to express obligation became more of a characteristic feature in Austen's later works: *Mansfield Park*, *Emma*, and *Persuasion*.

3 Collocates of Deontic Modals

Collocates of a given term are other terms that appear with greater than random probability in proximity to that term in a corpus. This greater than random probability indicates that the co-occurrence of two collocates depends on a relationship between the lexical terms, and examining such collocational patterns can reveal meaningfully ways in which terms are related in texts such as Austen's novels (Wijitsopon, 2013).

Leveraging the *Voyant Tools* collocates function using a window size of 5 tokens, we examined the collocates for the four deontic modals previously mentioned. As we are investigating auxiliary verbs, some of the top verbal collocates for each modal are highlighted in Table 1. As we can see, the top collocates for all of the modals are relatively similar verbs. They revolve around thoughts, communication, feelings, and perception, which supports the notion that Austen's stories are more strongly framed around feelings and perceptions than actions.

When further examining the collocates of the modals within each of Austen's novels, we found that the top collocates between novels were also

fairly similar, with the exception that the name of each heroine was also a top collocate within her own novel. This presents an opportunity to investigate how the distribution and usage of deontic modals differ between the heroines of Austen's novels.

4 Relative Modal Obligation per Heroine

As the heroines of all of Austen's novels were shown to have a high level of co-occurrence with the deontic modals being investigated, we set out to examine the occurrences of these modal verbs and if/how they were used in context to convey a sense of obligation in the narrative.

First, we looked to establish what proportion of deontic modals co-occurring with the mentions of the heroines' names were actually relevant to the given heroine in the discourse. This required a manual evaluation of all of the instances of a deontic modal that occur within 5 tokens to the left or 5 tokens to the right of a mention of one of the heroines' names in all six novels to determine whether or not they had some relevance to the heroine in question. This manual evaluation was completed by the author of this paper. This filtering brought the total number of occurrences of deontic modals to be further investigated from 244 occurrences down to 154 occurrences. (It should be noted that only co-occurrences with heroines' names/nicknames were investigated. There are many more co-occurrences with pronoun mentions of the heroines which could be examined in future work.)

After establishing for each heroine which occurrences of deontic modals were directly relevant, we then reviewed those occurrences to determine whether or not their usage introduced some intent of obligation into the discourse. This was a subjective annotation based on review of the surrounding context in the novel for each occurrence. The cases that remained (116 in total) were then compared to the total number of relevant modal occurrences for each heroine in order to determine the ratio of how many of the relevant co-occurrences of deontic modals actually introduced some element of obligation in the discourse for each heroine. The proportions for each heroine are listed in Table 2.

Looking at Table 2, we see that most of the heroines are around the same level, with Elinor and Marianne (S&S) on the lower end, and Fanny (MP) very much on the high end with a proportion of 0.97 of co-occurring deontic modals conveying some in-

Heroine	Relative Modal Obligation
Catherine (NA)	0.85
Elinor (S&S)	0.59
Marianne (S&S)	0.63
Elizabeth (P&P)	0.8
Fanny (MA)	**0.97**
Emma (E)	0.75
Anne (P)	0.75

Table 2: A ratio of the occurrences of deontic modals which are (1) co-occurring with the heroine name (2) relevant to the heroine in the discourse and (3) express some intention of obligation in the discourse relative to the occurrences of deontic models that just fulfill (1) and (2)

tent of obligation into the discourse that is relevant to Fanny. This tracks with literary critics' understanding of Fanny as a character that is largely propelled by other (generally male) characters (Todd, 2006). However, based on the relative frequencies of deontic modals across the novels that were presented earlier in the paper, one might expect the proportions in Table 2 to be uniformly higher of Austen's later heroines. In order to understand why this is not the case, we must also investigate the different types of obligation being introduced into the discourse by the these modals.

5 Categories of Obligation

For the purpose of this investigation, a small set of 5 categories for differentiating senses of obligation was developed in order to sort the 116 occurrences of deontic modals that were judged to invoke some sense of obligation in the discourse. The categorizations were developed from examining those 116 cases and considering what groupings might be relevant to character analysis. A brief description of each category and an accompanying example are listed below:

Obligation of Action (OA): Some obligation is placed on the action or the potential action of the heroine (either by another party or by the heroine herself).

> e.g.: *You have had a long run of amusement, and now you **must** try to be useful."* *Catherine took up her work directly. . .* (NA)

Obligation of Feeling (OF): Some obligation is placed on the feelings or the potential feelings

Category	Catherine (NA)	Elinor (S&S)	Marianne (S&S)	Elizabeth (P&P)	Fanny (MP)	Emma (E)	Anne (P)
OA	**0.73**	0.4	0.7	0.5	0.38	0.5	0.33
OF	0	**0.3**	0.1	0.06	0.24	0.21	0.17
OO	0.18	0.2	0.1	0.13	**0.21**	0.07	0.08
OE	0.09	0.1	0.2	0.31	0.15	0.11	**0.42**
O	0	0	0	0	0.02	**0.11**	0

Table 3: This table shows the proportional distribution of obligation categories associated with each heroine. The highest proportion among the heroines for each obligation category is highlighted in bold.

of the heroine (either by another party or by the heroine herself).

> e.g.: *He was evidently oppressed, and Fanny **must** grieve for him...* (MP)

Object of Obligation (OO): The heroine is the object of some sense of obligation that is placed (not by the heroine) on another party.

> e.g.: *... that Elinor's merit **should** not be acknowledged by every one who knew her, was to her comprehension impossible.* (S&S)

Expression of Obligation (EO): The heroine expresses some obligation or moral judgment on another party or a state of affairs. (If the expression is directed towards the heroine herself, first choose OA or OF if applicable.)

> e.g.: *...Mr and Mrs Musgrove," exclaimed Anne, "**should** be happy in their children's marriages.* (P)

Other (O): There is a sense of obligation expressed that does not fit into one of the above categories.

6 Distribution of Obligation Categories amongst Heroines

The categories described above were used to manually annotate the deontic modals occurring with each heroine. These manual annotations were completed by the author of this paper. Once these annotations had been made, the number of modal occurrences annotated with each obligation category were totaled for each heroine. Each of these category totals was then divided by the total number of modal occurrences associated with the heroine in order to get the proportional distribution of obligation categories associated with each heroine. The results of these calculations are presented in Table

3. Each column of the table shows the proportional breakdown for how the obligation categories are distributed for a single heroine.

Looking at Table 3, we see that Catherine (NA) has the highest Obligation of Action (OA) proportion. This is fitting, as Catherine is one of Austen's younger heroines, whose character at the outset of the novel is considered to be naive. As such, she is frequently instructed in her behavior. Throughout the course of the story, she learns what behavior is proper and what is not. Elinor (S&S) has the highest Obligation of Feeling (OF) proportion, which is also fitting because Elinor's central story line in *Sense and Sensibility* revolves around how she is obligated to suppress her feelings because the man she loves is already engaged to another woman. The highest Object of Obligation (OO) proportion comes from Fanny (MP), whose character is commonly criticized for being overly moralistic and lacking in agency (Todd, 2006). In *Mansfield Park*, Fanny's feelings are rarely consulted on any matter and she is often treated as an object. Finally, Anne (P) has the highest proportion for Expression of Obligation (EO). Anne is Austen's final heroine, and arguably her most mature. Anne is the oldest heroine, and she is also the most self-reflective and self-controlled. As such, it is fitting that Anne is the heroine who has the highest proportion in the category for expressing her own understanding of how things should be. Overall, we see that the breakdown of obligation categories amongst the modal occurrences for each heroine have a resemblance to the characterizations of Jane Austen's heroines that are understood through literary analysis.

7 Conclusion

In this paper we examined how deontic modals are distributed in Jane Austen's novels, and how they contribute to elements of Austen's style and themes. We found that the relative frequencies of deontic

modals are higher in Austen's later novels, and that the verbal collocates of these modals generally revolve around perceptions and communication, which support the literary notion that the framing of Austen's novels favors perception over physical action.

Upon examination of the co-occurrences of these modals with name mentions of the heroines of Austen's novels, we found that the proportion of deontic modals that worked to convey a sense of obligation in the discourse varied from heroine to heroine, and we found that the type of obligation most commonly conveyed by these modals also varied from heroine to heroine. As discussed above, these distributions reflect some general characterizations of Austen's heroines.

As these findings corroborate existing literary perspectives, we see that there is merit in engaging in quantitative corpus analysis to gather supporting evidence for literary claims, as well as to potentially seek out new patterns of interest for literary analysis. In addition these specific insights, this paper provides a methodology for gaining literary insights through the use of simple corpus linguistic methods. This paper also serves to demonstrate how accessible digital tools, such as *Voyant Tools*, allow for researchers without advanced programming skills to leverage quantitative methods of analysis to engage with literary texts.

References

Zelda Boyd. 1984. Jane austen's" must": The will and the world. *Nineteenth-Century Fiction*, 39(2):127–143.

J. F. Burrows. 1986. Modal verbs and moral principles: An aspect of jane austen's style. *Literary and Linguistic Computing*, 1(1):9–23.

Jennifer Carr. 2017. Deontic modals. In *The Routledge handbook of metaethics*, pages 194–210. Routledge.

Bettina Fischer-Starcke. 2010. *Corpus Linguistics in Literary Analysis Jane Austen and Her Contemporaries*. Continuum.

Laura L. Runge. 2019. Austen and computation 2.0. *Texas Studies in Literature and Language*, 61(4):397–415.

S. Sinclair and G. Rockwell. Voyant tools [online]. 2021.

Janet Todd. 2006. *The Cambridge Introduction to Jane Austen*. Cambridge University Press.

Raksangob Wijitsopon. 2013. A corpus-based study of the style in jane austen's novels. *MANUSYA*, 16(1):41–64.