

# Counting Complexity: Syntactic change in the style of Henry James

**Lesley Moss**  
**UCL**

This paper describes a pilot project for my investigation into the contrast between Henry James's early and late styles. In it I will describe the issues in the description of Henry James's style which motivate my research, my methodology and the results gleaned at this stage from my data.

## The 'difficult' late style

The style in which Henry James wrote his novels is frequently commented on and described as complicated or difficult. Marion Wynne-Davis in *The Bloomsbury Guide to English Literature* says that 'His second period shows a much more concentrated and difficult style of treatment' and she calls his last period 'the most intensive and subtle in style'. (1995 p.632) Detailed reference is often made to the particular style used by James in his later work.

Over the next three years, J. published the trilogy of novels that has come to define the style of his late phase, *The Wings of the Dove* (2 vols., 1902), *The Ambassadors* (1903), and *The Golden Bowl* (1904). Their prose is characterized by the increased length and complexity of sentences; multiple clauses serve to incessantly pursue and refine meaning, such that facts remain tentative, intentions fluid, and conclusions evanescent. (Taylor 2003 pp. 519-520)

Sometimes the description is linked to an explanation of James's purpose in using such a style.

Henry James's method of interior narration ... reaches maturity in the late fiction. Punctuation and italics, colloquial diction and phrasing, a predominantly loose sentence structure - these techniques are developed to heighten the consciousness of the central character and diminish the presence of the author. (Menikoff 1971 p.441)

Some critics note that the late style has been controversial from James's own time.

the style that characterizes the late fiction attempts to render the "appalling complexity" of the world and yet to render it as clearly as is possible. James's "late style," to which his brother William was one of the first, but hardly the last, to object, is highly abstract, relying on metaphors more common to poetry than to prose; ... Highly, figurative, James's late style demands of its reader an active participation; we must ourselves become those "on whom nothing is lost" if we are to grasp the late fiction. (Fowler 1993 p.181)

Gifford vividly expresses the difficulty of the reader who feels that James's style deteriorated through his career.

[The three late novels] 'are notorious for their difficulty ... The notation is almost excessively fine, the issues often appear tenuous, the atmosphere has been pumped "gaspingly dry". Readers who delighted in the pictorial brilliancy of his earlier work and its neatness of style, must now grope in a world where for all the animation of James's figurative speech both meaning and action often hang in suspense; (Gifford 1984 p.126)

Descriptions of James's style have rarely dealt in any detail with the syntax he uses. Chatman's *The Later Style of Henry James* is an important exception. Chatman recognised the accepted view of the late style - 'I start with the usual characterization of James' later style as "difficult".' (1972 p.2) – and locates the source of at least some of this difficulty in its 'abstractness'. He then examines in detail 'the devices or choices – syntactic and lexical – that evoke the "specific effect" of abstractness in the reading of James.' (Chatman 1972 p.3)

It is interesting to see Chatman's use, without computers, of quantitative methods. He begins his book with an examination of samples of some James novels, comparing novels from different stages of James's career and other contemporary works. He is apparently wary of numerical data, starting this section with 'I begin, tediously but unavoidably, with figures' (Chatman 1972 p.6) and giving the figures integrated into the text. They can, however, be summarised in a table as in Figure 1 (with James's novels in order of publication). The other novels Chatman used are:

- Joseph Conrad *Typhoon* (1902)
- Samuel Butler *The Way of All Flesh* (1903)
- George Gissing *Veranilda* (1904)
- E. M. Forster *Where Angels Fear to Tread* (1905)

Figure 1 Chatman's figures

	Sentences	Human	Tangible	Intangible	Thought verbs
<i>The American</i>	200	71%	6%	23%	221
<i>Portrait of a Lady</i>	200	56%	2%	42%	
<i>Wings of a Dove</i>	200	55%	1%	44%	250
<i>The Ambassadors</i>	196	45%	6%	49%	165
<i>The Golden Bowl</i>	200	53%	1%	46%	228
<b>Other novelists</b>	200	67%	15%	18%	

Chatman looks at the use of nouns as grammatical subjects and whether they referred to human beings, tangible objects or intangible elements. The samples show that there is a progression from the early to the late novels, with more subjects referring to concrete objects in the early novels, and more intangible reference in the later ones, with *The Portrait of a Lady* intermediate in both date and intangibility. The novels by other authors are most like *The American* in this respect. Chatman also includes some quantitative comparison of other elements. For example, he tries to see if there is more use of verbs of mental action in the late style, but the comparison is less satisfactory. *The Ambassadors* inexplicably breaks the pattern of the other 2 late novels and *The American* is not very different from *The Golden Bowl*.

Chatman introduces the idea of 'plenitude' – the great wealth of detail in the protagonist's consciousness which is conveyed by abstraction, ellipsis and complexity. He cites Ohmann's description of an embedded syntactic structure as the source of much of the difficulty of the later style, but feels that it is the combination of embedding and ellipsis which makes for the real difficulty, and describes how the reader has to 'hold open a mental slot' (Chatman 1972 p.126) while they wait for the ellipted item to be given, only to find that it never is. Chatman places the 'interpolation' which James introduces into his syntactic structure as usually occurring between a verb and its complement, although it can also occur between an adjective and its complement or between an auxiliary and main verb.

For Chatman, the significance of this complex syntactic structure is that it is a way of conveying a 'character's consciousness as he struggles with his problems' (Chatman 1972 p.127). James does not use first person narration so this is not a stream-of-consciousness style: Chatman suggested that James wishes to avoid the impression of the character directly addressing the reader. Instead he uses his abstract, complex, elliptical late style to show the workings of the protagonists' minds through a narrator, which importantly also allows an ironic tone.

More recently, Leech and Short examined James's 1891 short story 'The Pupil', looking at various grammatical features. According to Hoover's analysis, described later in this paper, the date of the story places it just after the end of James's intermediate period of novel writing. Leech and Short look at various lexical features, and then comment on James's sentence structure, noting the complexity and also the use of embedded clauses and parenthetical phrases.

James's predilection for parenthetical constructions[,] may find its justification in the impression that it gives us of complex wholeness (Leech and Short 1981 p.228).

James uses dependent clauses in a ratio of 3:1 with independent clauses – a level of use much greater than Leech and Short found in Conrad. They note that *that* clauses, with or without *that* actually appearing, are the dominant form, and relate this to the psychological verbs, such as *know that*.

These apparent perversities of James's syntax become meaningful in the light of an appraisal of his particular concern with psychological realism: his unremitting endeavour to pin down the psychological moment (Leech and Short 1981 p.102).

Though centred on this particular story, Leech and Short's explanation of James's stylistic decisions is generally applicable.

James grapples with this insuperable problem: that whereas for the human sensibility one moment holds a myriad of simultaneous conditions and possibilities, for the writer and reader one thing must come after another. Although he cannot escape from the linearity of language, James does the next best thing, which is to fasten our attention initially on the most immediate feature of Pemberton's predicament: his uncomfortable sense of indecision, and then to expatiate on it so that by the time we have threaded our way through two paragraphs, we have built up a sensitive grasp of the coexisting intricacies and ironies of that predicament (Leech and Short 1981 p.103).

Clearly, the distinctiveness and difficulty of James's late style is widely acknowledged, and the contrast with the early style often implied and sometimes stated clearly. However, the exact differences between the two styles are not fully explained. Where the late style is described, the description is often metaphorical, as in Gifford's description. It is not clear how the writers can be sure that there is a substantive change between early and late novels; it is the impression they have gained from their reading and experience, but the difference is not quantified. For example, where Taylor claims that there is an 'increased length and complexity of sentences' it is unclear whether length and complexity are synonymous, or always go together, and, if both have increased in the late style, to what extent this is true. Although Chatman has employed some quantitative methods, and has described in detail some linguistic elements of James's writing, corpus stylistics with its automated processing can examine large quantities of data and therefore make greater claims to representativeness. Chatman's samples are only 200 sentences long, and the sentences are consecutive, so that only one portion of the work is examined. (It is not clear how Chatman selected his samples – they are from different parts of the different novels.) Leech and Short's study is very helpful but it is based on one story, and quantitative data is limited. The purpose of my research is to explore the use of corpus stylistics to examine quantitatively the differences between James's early and late styles (and by extension its use in describing literary style in general) so that any claims made can be supported by data.

## A corpus stylistic approach

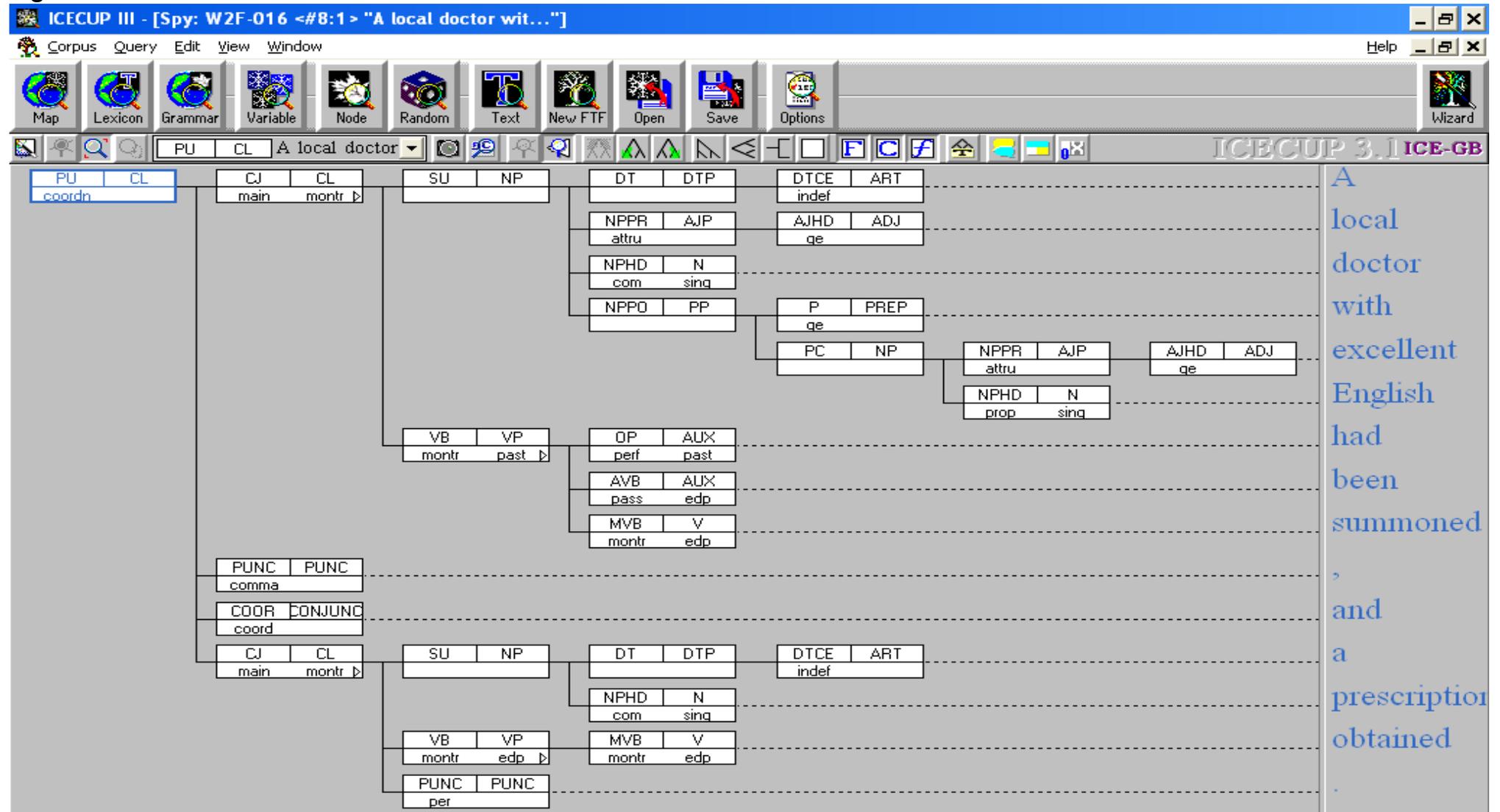
I define corpus stylistics as applying the methods of corpus linguistics to stylistics. While corpus stylistics (under this or another name) is growing, other studies have overwhelmingly been vocabulary-based, such as Stubbs's paper on Conrad (Stubbs 2005). This looks at collocation and word sequences to find themes underlying the obvious topics of *Heart of Darkness*. Mahlberg's work on Dickens is also centred on collocation and word sequences, which she calls clusters (Mahlberg 2007).

David Hoover analysed James's vocabulary in a paper given at the Style in Fiction symposium at the University of Lancaster in 2006 (Hoover 2007). He examined frequency of word use, first checking his methodology by comparing Henry James's work to that of other contemporary writers. He found that he could clearly distinguish James's novels from the other works, using lists of the 3-4,000 most frequent words in a text, which represented about 95% of all the words of each novel. He then selected the words from the one-thousandth to the two-thousandth in the most frequent word list, so that they were relatively rare. Hoover analysed his data using principal components analysis and cluster analysis. Using these methods, he was able to chart James's novels, finding that they fell neatly into chronological order and into three clear phases of work – 1871-81, 1886-90 and 1897-1917. The gaps between these three phases coincide with breaks in James's novel writing. The correlation Hoover found was also seen when James revised early novels for his New York edition; they moved in the direction of the late style, although they could still be identified as early works by the statistical analysis. This chronological development was not seen in the work of other novelists such as Dickens and Gaskell. I have used Hoover's three periods in my choice of novels.

My methodology makes use of the ICECUP 3.1 program which was developed at the Survey of English Usage and is used for the ICE-GB corpus of written and spoken British English which is held there. ICECUP presents each sentence in a 'tree' which displays the syntactic structure of the sentence, including clause, phrase and word levels with both function and word labels. There are also features of most words or clauses shown in small type within the box under the main labels. The example in Figure 2 is from ICE-GB and shows a sentence from one of the literary samples in the corpus. The tree 'grows' from the 'PU' (parsing unit) at the top left of the page, and here is a coordinated clause. This divides into the two clauses which make up the sentence with their intervening coordinating conjunction and punctuation, and then breaks down the subject and predicate of each clause into its component parts. Or, starting from the word level, the word *prescription* is shown as a noun, which is common and singular and has the

function of being head of a noun phrase. It is preceded by the indefinite determiner *a* and forms part of the subject noun phrase of the second of the two conjoined clauses. As can be seen from this brief explanation, even a simple ICECUP tree contains an enormous amount of information.

Figure 2 An ICE-CUP tree



A local doctor with excellent English had been summoned, and a prescription obtained.

One of the challenges of corpus stylistics is obtaining (or otherwise having to make) good electronic versions of the texts which are to be studied. My texts were chosen at random as far as possible, using David Hoover's 3 periods of James's work (which seems to be in line with other divisions of the novels) but the choice was constrained to texts which had already been prepared in electronic form to a high degree of accuracy and with information about the source text used. It was vital that the early novel database be based on a text which used the original version rather than a later revision. Most of the electronic texts of James's work, where they give a source at all, are taken from the extensively revised New York edition, which would reflect, at least in part, the late style. The early novel I am using is *Washington Square* which was published in 1881, and the late novel is *The Golden Bowl*, published in 1904, although the version I am using is the revised 1909 edition. For the late novel, I have assumed that the later the revision, the more it will reflect any characteristics of the late style.

For this pilot project, I have analysed the first chapter of each novel. The first chapter may not be representative of the novel as a whole, but I have made the assumption that one first chapter can be valuably compared to another first chapter. Having chosen the novels, they had to be prepared for passing through the ICECUP tagging program version 1.0, processed by that program and then hand-checked and corrected. The tagging program attempts to assign parts of speech and function labels at the word level. I have added by hand some clausal structure, as well as some special annotation to allow me to search for features of interest which are not part of the ICECUP system. For example, where there is direct speech in a sentence, I have inserted <ds> into the code for that sentence so that these sentences can be separately retrieved. This annotation does not appear in or affect the tree diagram produced by ICE-GB. In putting together the database, I have had to decide what a 'sentence' should consist of. A sentence here is a piece of text which begins with a capital letter and ends with a full stop. As will be seen below, this sometimes includes more than one clause which would be acceptable as a free-standing sentence, but I have chosen to take James's decision as to what should be included in one sentence as part of his style.

Figure 3 shows one of the sentences from my database. Comparing it to the ICE-GB tree shown in Figure 2, it can be seen that, due to the pressure of time, I have not inserted the phrasal level of analysis which would be required for a fully parsed corpus.

# The Golden Bowl Sentence 61

Figure 3 The Golden Bowl Sentence 44

ICECUP III - [Spy: GB01 <#44:1 > "He had been intereste..." - OK]

Corpus Query Edit View Window Help

Map Lexicon Grammar Variable Node Random Text New FTF Open Save Options Wizard

ICECUP 3.1 JAMES

PU	CL	NPHD	PRON	He had been intereste in his discrimi , which amused him .
main	cop	past		
		OP	AUX	
		perf	past	
		MVB	V	
		cop	edp	
		AJHD	ADJ	
		edp		
		P	PREP	
		ge		
		<b>Features: general</b>	PRON	
		poss	sing	
		NPHD	N	
		com	sing	
		PUNC	PUNC	
		comma		
		NPPO	CL	
		depend	rel	
		montr	past	
			NPHD	PRON
			rel	
			MVB	V
			montr	past
			NPHD	PRON
			pers	sing
			PUNC	PUNC
			per	

He had been interested in his discrimination, which amused him.

start | Micro... | ICEC... | ICEC... | WS01... | 2007 ... | Address | 11:54

This example sentence from *The Golden Bowl* illustrates the annotation of the texts. The relationship between the main and dependent clauses is shown in the shape of the diagram.

## Chapter comparison

It may be helpful first to give some basic statistics about the two chapters. These are shown in Figure 4. The first chapter of *Washington Square* is referred to throughout as WS01, and the first chapter of *The Golden Bowl* as GB01.

**Figure 4 Basic statistics**

	<b>Words</b>	<b>Sentences</b>	<b>Words per sentence</b>
<b>WS01</b>	1386	39	35.54
<b>GB01</b>	6268	290	21.61

The first chapter of *The Golden Bowl* is substantially longer than the first chapter of *Washington Square*, which might fit with the idea of the later novel being more complex and difficult. However, the sentences in WS01 are very much longer, which seems surprising. We might expect 'difficulty' to correlate with long sentences. The explanation may be that the difficulty of the late style is, in part, a question of syntactic complexity. There is a distinction between a compound sentence, which may be very long, and a complex sentence, which despite its formal complexity, may be short.

This pattern is complicated in James's work by two other features of his style. As mentioned above, James sometimes includes more than one main clause in a sentence without any co-ordination. These clauses may be joined by a colon, a semi-colon or a dash. For example:

1. *So is his whole life over here – it's the most romantic thing I know.* (GB01:124)

(Sentences are identified by their text and their sentence number in the database.)

I call this phenomenon, which is by no means unique to James, an 'additional main clause'. It is necessary to identify and mark these sentences specifically because the usual practice in ICE-GB would be to divide the linked clauses into separate sentences.

James also sometimes breaks into a sentence to insert an element which is not part of the syntax of the sentence. This may be a short noun phrase or a whole clause, as in:

2. *His relation to the things he cares for – and I think it beautiful – is absolutely romantic.*  
(GB01:123)

Such items are marked in ICECUP as a Detached Function and they are additionally analysed according to their form – whether the interruption is made by a clause or noun phrase etc.

An analysis of the number of co-ordinated clauses, dependent clauses, additional main clauses and detached clauses in the two chapters yields the result in Figure 5.

**Figure 5 Sentence features**

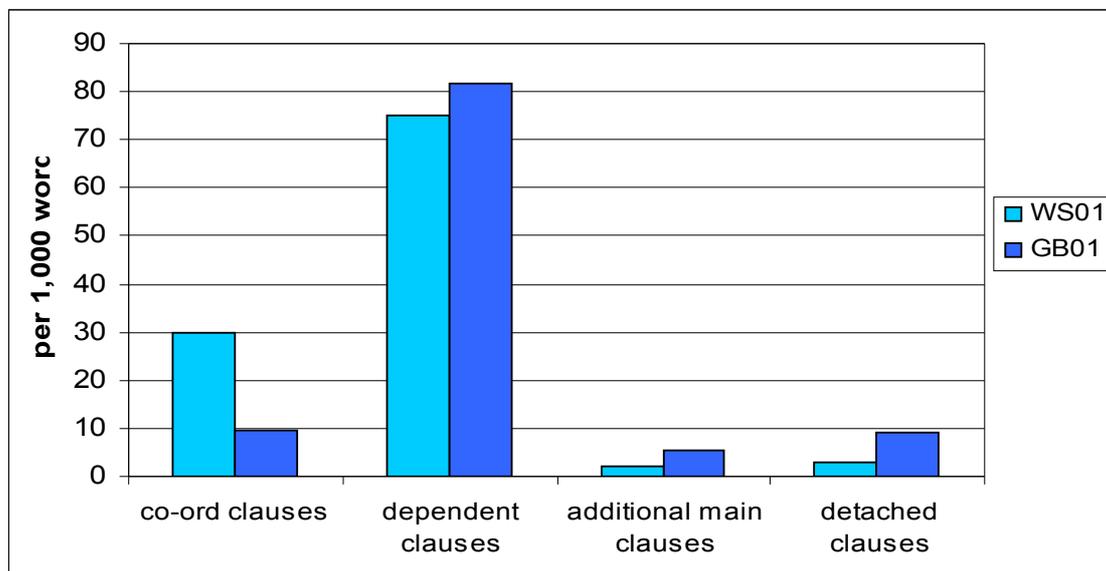
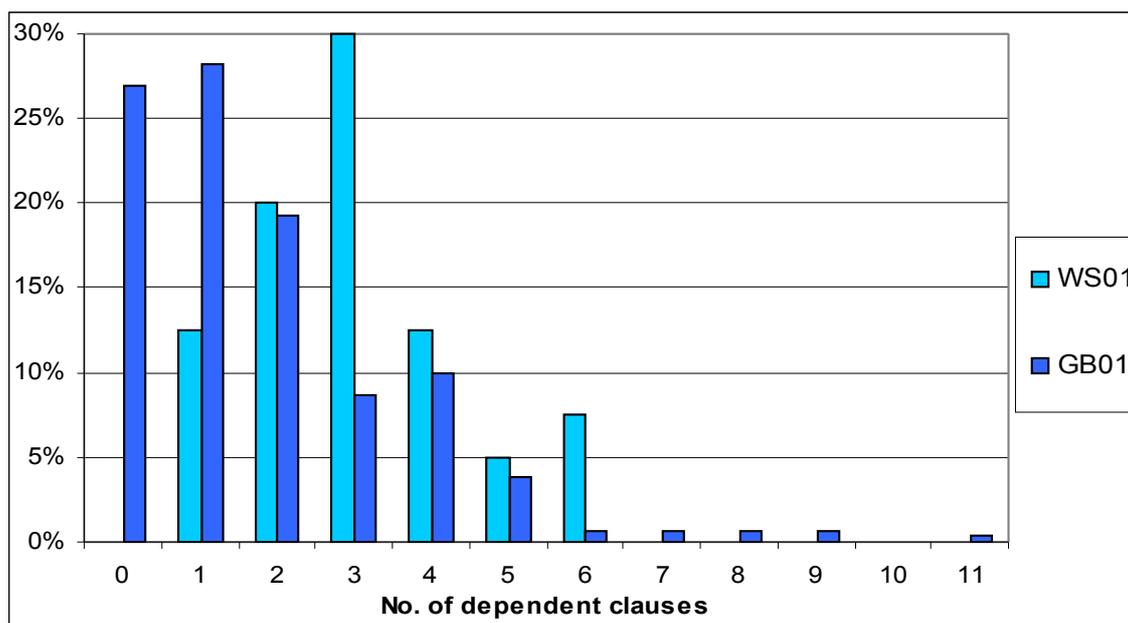


Figure 5 illustrates the occurrence of these features per 1,000 words of text so that the two chapters, which are unequal both in number of words and sentence length, can be compared. The longer sentences of WS01 are explained by their having a greater number of co-ordinated clauses, and GB01 has more dependent clauses, additional main clauses and detached clauses. This suggests that the 'difficulty' commented on by critics is not a matter of sentence length, but does correlate with syntactic complexity. However the contrast between the two chapters in these measures of syntactic complexity is not very great and hardly seems to justify a perception of a distinct development between the early and late styles. This is a point I will return to later in this paper.

A different way of looking at the data on dependent clauses, which are the main measure of syntactic complexity, is to see how many dependent clauses occur in each sentence, as in Figure 6:

Figure 6 Dependent clauses per sentence



This graph illustrates that, if having more dependent clauses is the measure of syntactic complexity, GB01 has much more complex sentences than any in WS01. No sentence in WS01 has more than six dependent clauses, whereas there is a sentence in GB01 which has 11 (GBO1:264). On the other hand, there are far more sentences with three, four, five and six dependent clauses in WS01. This may, of course, be an anomaly due to the very small number of sentences in WS01. The 7.69% of sentences in WS01 which have 6 dependent clauses represents only 3 sentences. With this caveat, the distribution of dependent clauses remains interesting; it appears that in this sense GB01 is only very partially more complex than WS01, despite the very different impressions the two chapters make on the reader. It may be that the reader's perception of complexity relies on a few extremely complex sentences, whereas the bulk of the texts are not so different.

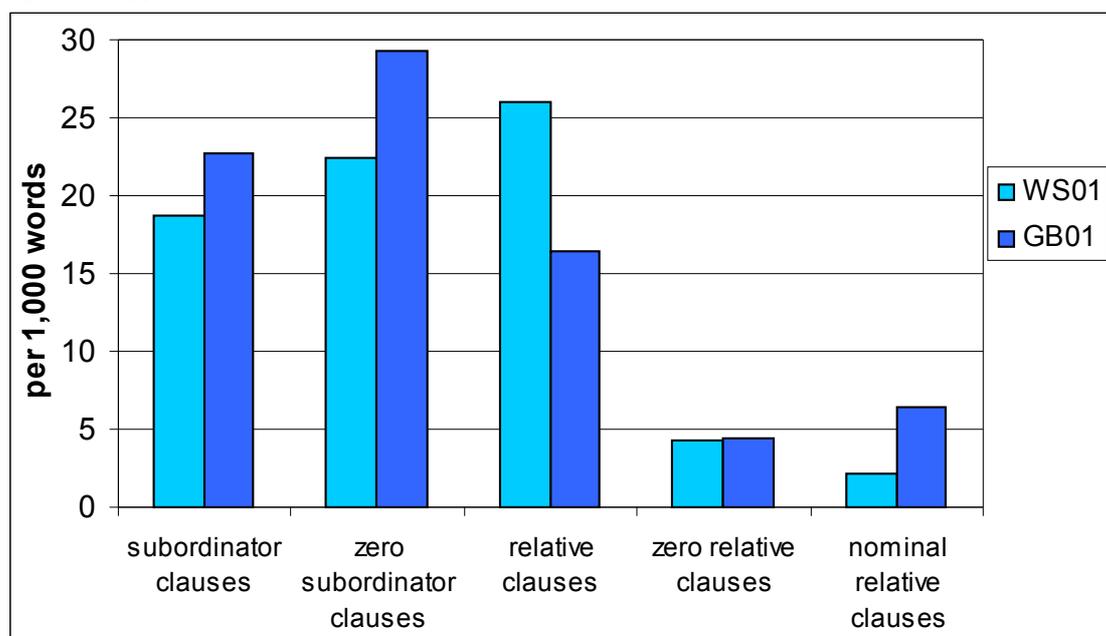
A more detailed analysis of the types of dependent clauses used in the two chapters gives the breakdown shown in Figure 7. 'Subordinator clauses' are dependent clauses, excluding relative clauses, with an overt subordinator, such as:

3. *I hasten to add, to anticipate possible misconception, that he was not the least of a charlatan.* (WS01:11)

'Zero subordinator clauses' similarly exclude relative clauses but have no overt subordinator, as seen in the same sentence:

4. *I hasten to add, to anticipate possible misconception, that he was not the least of a charlatan.*

**Figure 7 Types of dependent clause**



Most categories show that GB01 has more examples of each kind of dependent clause than WS01, making up the overall figures shown earlier. However, the difference in the use of zero relative clauses is so close that it is unlikely to be important. An increase in this type of clause, for example:

5. *“Good, bad or indifferent, I hope there’s one thing you believe about me.” (GB01:171)*

does not seem, on the basis of this small sample, to be a distinction between the early and late styles. Nominal relative clauses do show a more marked difference in usage, but they are also very rare. An example is:

6. *It would come to asking what they expected him to do. (GB01:282)*

The total number of these clauses in both chapters is only 39. Until I have more data, it seems that these constructions are too rare to draw conclusions on style development.

The main increase in dependent clauses is in subordinator clauses, with and without the subordinator present ('zero subordinator'), which was noted by Leech and Short to be the dominant form. It is interesting to see that this increase is paralleled by a decrease in relative clauses. The reason for this change is difficult to identify, and it is unclear whether the decrease

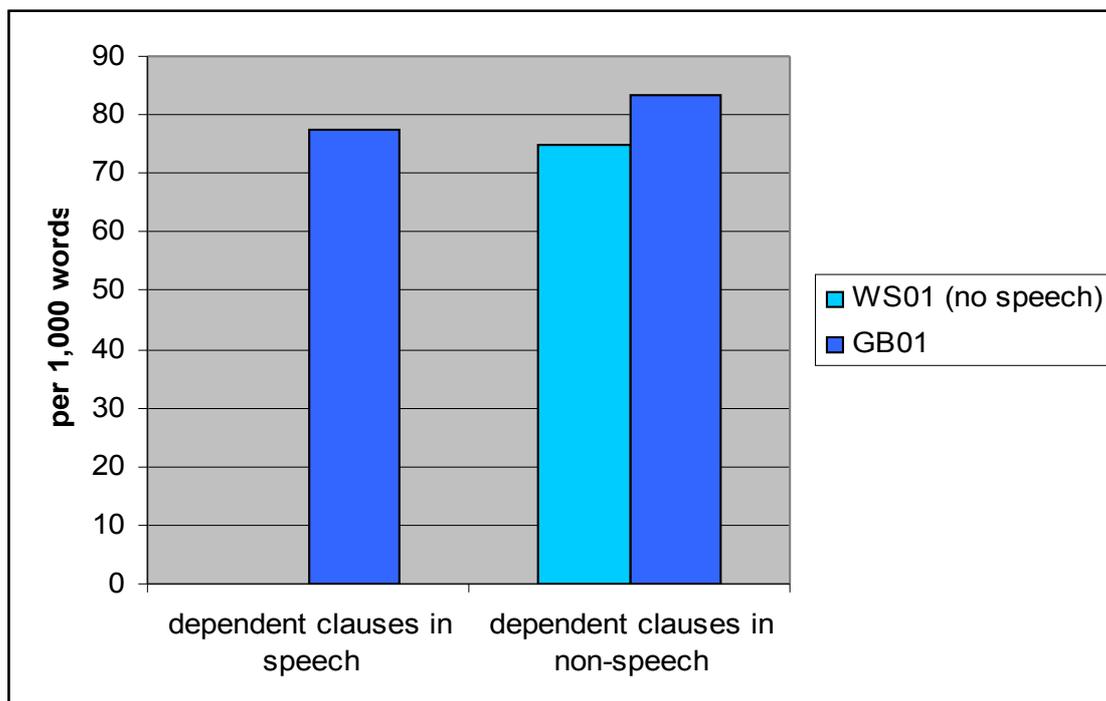
in the use of relative clauses and the increase in the use of other subordinate clauses in the late style are related phenomena. Wierzbicka (Wierzbicka 1988) suggests that *that* clauses carry an inherent meaning of knowledge, which would support the idea that in the later novels James is increasingly interested in conveying the workings of his protagonists' minds – literally, what they know. However, this conflicts with Chatman's difficulty in finding a distinct increase in the use of verbs of mental action in the late novels.

### Speech v. non-speech

One of the differences between the two chapters is that GB01 includes 127 sentences which have an element of direct speech whereas WS01 has no direct speech at all. There has been a great deal of study of the differences between speech and writing, and it may be that, when James represents speech, his style is different from non-speech narrative. It seems likely that his characters' speech might be simpler than the involved sentences of his narrative style. If so, it is worthwhile separating out these two kinds of writing in the analysis.

The result is shown in Figure 8:

**Figure 8 Comparison of direct speech and narrative**



As expected, once sentences containing direct speech are removed, the contrast in the number of dependent clauses per 1,000 words in each chapter is greater, and suggests that this may be one of the real differences between the styles of the two chapters. It is notable, however, that even the sentences containing direct speech in GB01 have more dependent clauses than the

narrative sentences of WS01. This may be partly because, for practical reasons, I have counted as ‘speech’ sentences which contain direct speech (within speech marks) but these sentences often contain a non-speech element. This may be just a reporting clause:

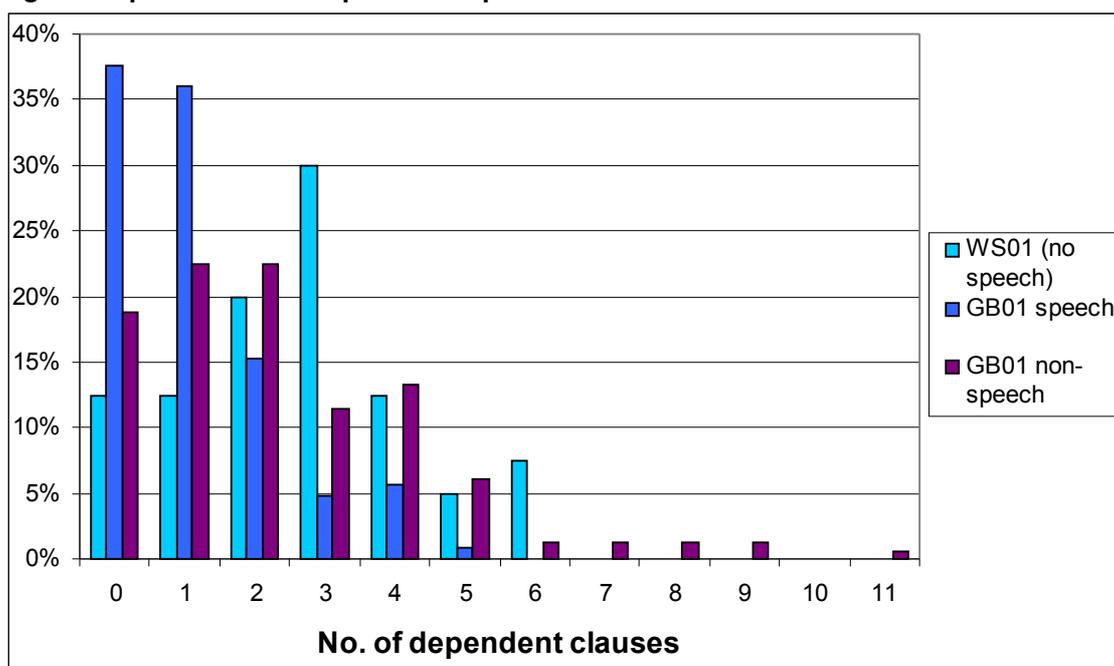
7. *“Do you think it would be good for YOU?” Maggie Verver had smilingly asked.* (GB01:33)

but may include a good deal of narrative as well.

8. *“When I speak worse, you see, I speak French,” he had said; intimating thus that there were discriminations, doubtless of the invidious kind, for which that language was the most apt.*(GB01:20)

This is, however, an unusually elaborated example. Most of the speech sentences contain only a small amount of narrative. In Figure 9 the number of dependent clauses in speech and non-speech sentences in GB01 is compared with the sentences of WS01, which has no direct speech at all. From this it emerges that there are really three distinct patterns to compare in these two chapters.

**Figure 9 Speech and non-speech compared**



As shown before, WS01 has sentences with from zero to six dependent clauses, with the highest percentage of sentences having three. The portrayal of direct speech in GB01 is in sentences with few dependent clauses; most have none at all, none have more than five. Non-speech sentences in GB01 have more dependent clauses, on the whole, than those of WS01, but not for some reason with sentences with three or six dependent clauses. However the GB01 non-speech sentences include examples with up to eleven dependent clauses.

## Delay

Another notable component of Henry James's style is what I have called 'delay', which I define as the interpolation of a word or words between an element and its complement. Delay has similarities with what are elsewhere called parentheticals or disjuncts, although the focus here is particularly on perceived complexity of style. It can be illustrated with the sentence WS01:15, which is arranged below in such a way as to emphasise its interrupted syntax:

9. *He had married,*  
    *at the age of twenty-seven, for love,*  
*a very charming girl,*  
    *Miss Catherine Harrington, of New York,*  
*who,*  
    *in addition to her charms,*  
*had brought him a solid dowry.*

'Delay' includes but is not limited to, the detached clauses shown in Figure 5. In this example 'a very charming girl' is the direct object NP of the verb 'married', and the reader has to keep in mind the possible existence of a direct object (as 'married' could also be intransitive) while reading the intervening phrases. Similarly there is a delay between the word 'girl' and the relative pronoun which refers to it and is the subject of the relative clause which describes it. Finally, there is a delay between 'who' and the rest of the clause of which it is the subject.

This phenomenon was described by Ohmann as self-embedding, an alternative to left- or right-branching constructions. He analyses the opening sentence of one of James's short stories and finds that the interpolated parts of the sentence are longer than the main sentence itself.

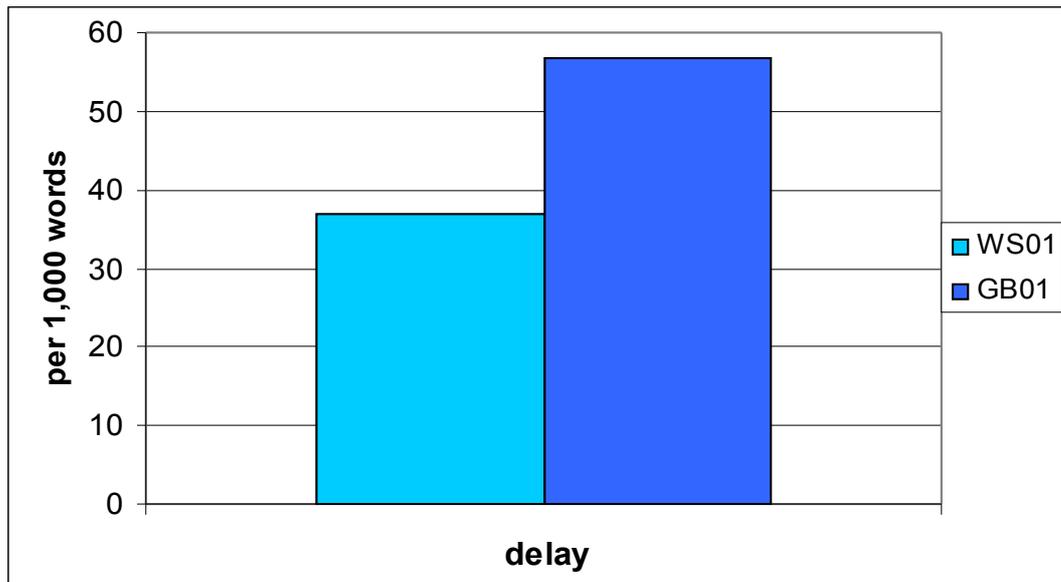
the strain on attention and memory required to follow the progress of the main sentence over and around so many obstacles is considerable. (Ohmann 1964: 437)

Ohmann considers this to be an important defining part of James's style:

It seems likely that much of James's later style can be laid to this syntactic device - a matter of *positioning* various constructions, rather than of favoring a few particular constructions. (Ohmann 1964: 437)

I have added special annotation to the computer code of the sentences in my database where delay occurs. The number of separate occurrences of delay per 1,000 words of text is shown in Figure 10, and gives quite a clear contrast between the two chapters. As Ohmann and Chatman have suggested, delay or embedding does seem an important characteristic of James's late style.

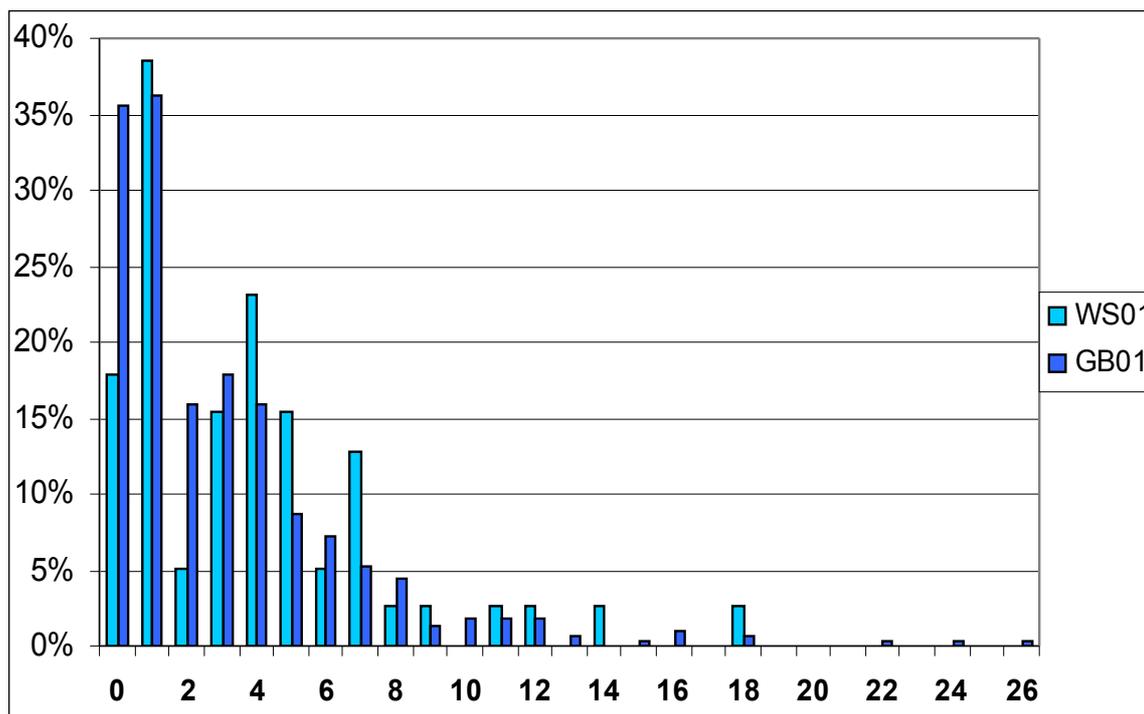
**Figure 10 Instances of delay**



I have also quantified the delay by counting the number of words which intervene between an element and its complement. Thus, in the sentence above, there is a delay of 7 between 'married' and 'a very charming girl'. It is important here to define what a 'word' is, as this directly affects the counting of delay. ICECUP groups together ('ditto-tags) the proper names 'Miss Catherine Harrington' and 'New York', counting them as one word. It also ditto-tags complex prepositions, like 'in addition to'. I have accordingly given this sentence a delay score of 13 (7+3+3).

Delay score can also be analysed on a sentence by sentence basis, and this is shown in Figure 11. It resembles the analysis of dependent clauses. There are sentences in GB01 which have very large delay scores and are undoubtedly very difficult for the reader to process.

Figure 11 Delay score



The sentence with a delay score of 26 is the fourth sentence of *The Golden Bowl*:

10. *It was not indeed to either of those places that these grounds of his predilection, after all sufficiently vague, had, at the moment we are concerned with him, guided his steps; he had strayed simply enough into Bond Street, where his imagination, working at comparatively short range, caused him now and then to stop before a window in which objects massive and lumpish, in silver and gold, in the forms to which precious stones contribute, or in leather, steel, brass, applied to a hundred uses and abuses were as tumbled together as if, in the insolence of the Empire, they had been the loot of far-off victories.* (GB01:04)

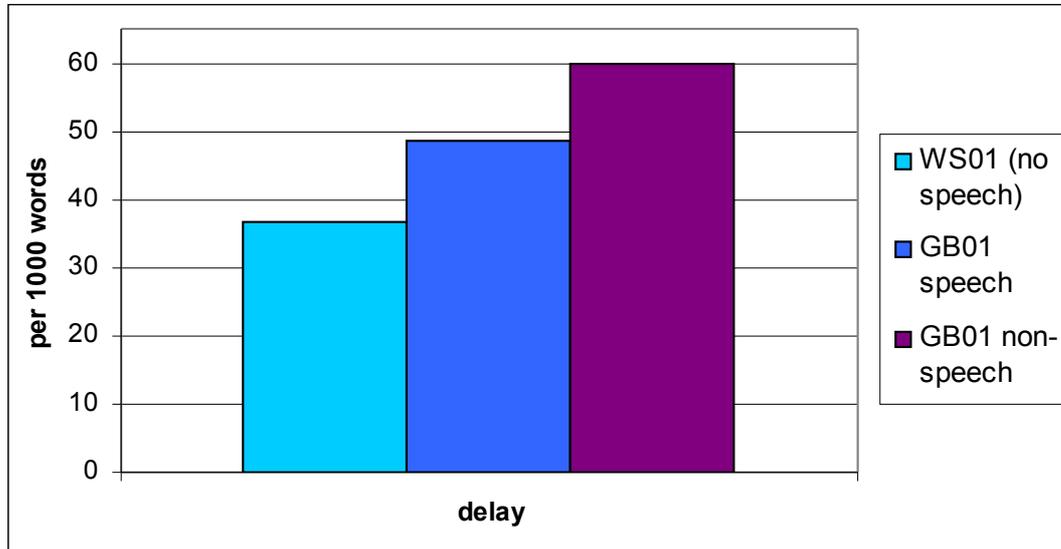
(The sentence also contains an additional main clause.)

However, the contrast between the 2 chapters is not entirely clear. There are a higher percentage of sentences from WS01 which have delay scores of 1,4,5,7,11,12,14 and 18, although there are no sentences at all from the early chapter which have delay scores beyond 18. Again it seems that the impression of difficulty may come from a few, very memorable sentences rather than a homogeneously higher level of difficulty throughout.

An analysis of the difference between the representation of direct speech and of non-speech narrative is also useful in comparing delay. As shown in Figure 12, there is a clear distinction between the three sections of writing. Although the speech sentences in GB01 have more

instances of delay than the WS01 narrative, the contrast between the two non-speech narrative texts is very marked.

**Figure 12 Counting delay**



## Conclusion

Any conclusions from this study are extremely tentative. A comparison of one chapter from each of two novels may not represent the style of each overall novel, and in turn the novels may not be representative of the style of all of James's novels from that period. However, from the data examined thus far, there does seem to be a difference between the two chapters in the amount of syntactic complexity as measured by the use of dependent clauses, particularly if sentences representing direct speech are excluded. GB01 is clearly more complex by this measure, and also has a greater use of 'delay' or embedded phrases within syntactic constructions. However, a straightforward increase in complexity from WS01 to GB01 is complicated by the higher percentage of dependent clauses and delays in the low to medium range in WS01.

Writing sentences with many dependent clauses may be an attempt to be extremely precise about what is being described by giving a great deal of detail about a scene or thought, as in sentence GB01:04 shown above. However, it is not necessary for such detailed description to be used in an unusual word order. Rather it seems that James is trying to convey something of

the quality of a musing mind, with its twists and turns and interruptions, which is also analogous to the Prince's physical wandering through the streets of London.

This pilot project will be succeeded by a larger study, which will include 15 chapters from each novel, selecting the first five, the middle five and the last five chapters. When a greater quantity of data is available, it will be possible to see if clearer differences emerge between the early and late styles, or if a few sentences continue to make the greatest contrast. Other areas to explore include analysis of the functions of the dependent clauses – for instance, whether the clause is a direct object, adverbial, subject complement, etc. This might give further insight into the uses James makes of dependent clauses at different stages in his career. It may also be fruitful to examine where the most complex phenomena arise in the texts to see what James is describing when his style is most 'difficult'.

Finally, this methodology allows an examination of much larger volumes of text than was possible in earlier work like that of Chatman, and an objective identification of syntactic differences between James's novels, which could be applied to other literary works.

## Notes

. Further information on the ICECUP program can be found at <http://www.ucl.ac.uk/english-usage/resources/icecup/index.htm>

2. The *Washington Square* text is taken from <http://www2.newpaltz.edu/~hathawar/washsq.html>. The *Golden Bowl* text is taken from <http://www2.newpaltz.edu/~hathawar/goldenbowl1.html>.