

**SHAKESPEARE'S LATE SYNTAX: A COMPARATIVE
ANALYSIS OF *WHICH* RELATIVISATION IN THE
DRAMATIC WORKS**

by

Jacqueline E. Mullender

A thesis submitted to
The University of Birmingham for the degree of
DOCTOR OF PHILOSOPHY

Department of English
College of Arts and Law
The University of Birmingham
June 2010

UNIVERSITY OF
BIRMINGHAM

University of Birmingham Research Archive

e-theses repository

This unpublished thesis/dissertation is copyright of the author and/or third parties. The intellectual property rights of the author or third parties in respect of this work are as defined by The Copyright Designs and Patents Act 1988 or as modified by any successor legislation.

Any use made of information contained in this thesis/dissertation must be in accordance with that legislation and must be properly acknowledged. Further distribution or reproduction in any format is prohibited without the permission of the copyright holder.

ABSTRACT

This thesis combines corpus stylistic, literary and historical linguistic approaches to test critical observations about the language of Shakespeare's late plays. It finds substantial evidence of increased syntactic complexity, and identifies significant linguistic differences between members of the wider group of later plays. Chapter One outlines the critical history of the late works, including consideration of stylometric approaches to Shakespeare's language. Chapter Two describes the stylistic methodology and corpus techniques employed. Chapter Three reports the finding of salient *which* frequency in the late plays, and details the analytical categories to be used in the examination of *which* usage, the results of which are discussed in Chapter Four. Chapter Five describes two further analyses, where a broader group of ten late plays is considered on the basis of their high *which* frequency. The relativisation syntax of the five post-1608 plays (*Cymbeline*, *The Winter's Tale*, *The Tempest*, *Henry VIII* and *The Two Noble Kinsmen*) is found to distinguish them unequivocally as a group, while *Pericles* stands out as an anomaly. Finally, in Chapter Six it is argued that Shakespeare's syntax reflects a stylistic phenomenon unrelated to individual dramatic characterisation, motivated by his re-association with the Elizabethan romance writers of the sixteenth century.

ACKNOWLEDGEMENTS

My thanks first and foremost go to Professor Michael Toolan, for his support, enthusiasm and insight during my postgraduate studies at Birmingham. I am also very grateful to the following: The Arts and Humanities Research Council, for their financial support of this project; Dr. Catherine Alexander and Professor Kate McLuskie at the Shakespeare Institute, for their interest and encouragement; Dr. Martin Wiggins, for his rigorous responses to various presentations of this research; Dr. Jonathan Culpeper, Dr. Paul Rayson, Dr. Mike Scott, Professor Sylvia Adamson and Dr. Dawn Archer, for their kind help at workshops and conferences, and early access to forthcoming articles; Caroline Tagg and Garry Plappert, for their rating work and insightful feedback, and the other research postgraduates in Language at the University of Birmingham, who taught me a great deal. Finally, I would like to thank my family and friends for their kind tolerance of my neglect of them during this research.

TABLE OF CONTENTS

Chapter 1: The Late Plays: Literary Critical Background

1.0	General remarks.....	1
1.1	The late plays: group identity.....	2
1.2	The language of the late plays.....	11
1.3	Authorship attribution studies.....	21
1.4	Summary.....	32

Chapter 2: Methodology and Method

2.0	General remarks.....	34
2.1	Stylistics.....	35
2.2	Corpus stylistics.....	38
2.3	Method.....	42

Chapter 3: Analysing the Corpora

3.0	General comments.....	49
3.1	The prominence of <i>which</i> in the late plays.....	51
3.2	Paradigmatic analysis.....	56
3.3	Syntagmatic analysis.....	71
3.3.1	Main functions of <i>which</i>	74
3.3.2	Antecedents.....	76
3.3.2.1	Types of relative clause according to antecedent.....	76

3.3.2.2	The animacy parameter.....	80
3.3.2.3	Distance from the relative.....	81
3.3.2.4	Word groups intervening to the left.....	86
3.3.3	The information parameter.....	89
3.3.3.1	Restrictive and non-restrictive clauses.....	89
3.3.3.2	Continuative clauses.....	98
3.3.4	The role of <i>which</i> in the relative clause.....	118
3.3.4.1	Nominal roles.....	118
3.3.4.2	<i>Which</i> as prepositional complement.....	120
3.3.4.3	Relativiser-only <i>which</i>	120
3.3.4.4	Cataphoric <i>which</i>	125
3.3.4.5	Pushdown constructions.....	126
3.3.4.6	Recursions.....	131
3.3.5	Interventions to the right of <i>which</i>	135
3.3.5.1	The number of words intervening to the right.....	135
3.3.5.2	Word groups intervening to the right.....	140
3.3.6	Clausal length.....	144

Chapter 4: Results and Discussion

4.0	General remarks.....	153
4.1	Main functions of <i>which</i>	154
4.2	Antecedents.....	157
4.2.1	Types of relative clause according to antecedent.....	157
4.2.2	The animacy parameter.....	164
4.2.3	Distance from the relative.....	166

4.2.4	Word groups intervening to the left.....	168
4.3	The information parameter.....	172
4.3.1	Overall statistics.....	172
4.3.2	Continuative clauses.....	175
4.4	The role of <i>which</i> in the relative clause.....	178
4.5	Interventions to the right of <i>which</i>	190
4.5.1	The number of words intervening to the right.....	190
4.5.2	Word groups intervening to the right.....	191
4.6	Clausal length.....	201
4.7	Summary of findings for the late plays corpus.....	203

Chapter 5: Additional Perspectives

5.1	A new late group?.....	207
5.2	Findings for the new late group.....	212
5.3	The individual plays.....	218
5.4	Chapter summary.....	229

Chapter 6: Conclusions

6.1	Review of method.....	232
6.2	Overview of findings.....	238
6.3	Related studies.....	249
6.4	Further interpretation.....	254
6.4.1	General comments.....	254
6.4.2	The dramatic function of complexity.....	254
6.4.3	The Elizabethan revival.....	259

Appendices

1	Non-Shakespearean material removed from the corpora.....	313
2	Wordlists (earlier and late plays).....	cd
3	Keyword list for the late plays.....	315
4	Keyword list for the late plays (Excel table with full data).....	cd
5	Paradigmatic statistics.....	316
6	Paradigmatic relatives: all concordances.....	cd
7	Concordances for <i>which</i> (earlier and late plays).....	cd
8	List of <i>which</i> occurrences given to raters.....	cd
9	Tables of results for the analysis of the ‘new late group’.....	270
10	Tables of results for the analysis of the individual plays.....	281
10	The reconstructed <i>Pericles</i> text: principal differences from Craig (1914).....	318

	Bibliography.....	321
--	-------------------	-----

LIST OF TABLES

Chapter 3

3.1	The late plays: top 49 keywords.....	52
3.2	“Overall survey of relativisers in Modern English” (Dekeyser 1984, 64).....	61
3.3	Dekeyser’s survey of relativisers (1984, 64) without zeros.....	64
3.4	Raw frequency and proportion of relatives in Shakespeare.....	65
3.5	Proportion of running tokens that are relatives in Shakespeare.....	66
3.6	Shakespeare’s relatives: raw and relative frequencies of types	67
3.7	Shakespeare’s relatives: proportion of each type that function as relatives.....	68
3.8	“Relativisation correlated with register” (Dekeyser 1988, 31).....	69
3.9	<i>Wh-</i> forms and <i>that</i> in Shakespeare’s earlier and late plays.....	70
3.10	Rating results: divergent ascriptions.....	111

Chapter 4

4.1	Main functions of <i>which</i>	156
4.2	Types of relative clause according to antecedent.....	164
4.3	“The de-humanization of (THE) WHICH” (Dekeyser 1984, 71).....	165
4.4	The animacy parameter.....	165
4.5	Distance between antecedent and relative.....	167
4.6	Word groups intervening to the left.....	171
4.7	The information parameter.....	174
4.8	Analysis of continuative clauses.....	177

4.9	The role of <i>which</i> in the relative clause: principal functions.....	188
4.10	The role of <i>which</i> in the relative clause: additional constructions.....	189
4.11	Number of words intervening to the right.....	191
4.12	Word groups intervening to the right.....	198

LIST OF FIGURES

Chapter 3

3.1	Analysis of <i>which</i> clause antecedent types.....	81
3.2	Single word antecedents.....	94
3.3	Analysis of the information parameter.....	117
3.4	Analysis of <i>which</i> function in the relative clause.....	135

Chapter 4

4.1	Clausal length.....	202
-----	---------------------	-----

CHAPTER 1

THE LATE PLAYS: LITERARY CRITICAL BACKGROUND

1.0 General remarks

The group of Shakespeare's plays written at the end of his career has generated a large body of critical commentary, often focusing on what critics perceive to be features of the plays that set them apart in some way from those that precede them. Very rarely, however, have commentators described such features with any real concreteness or objectivity, leaving the field open to varying opinions, without a sound basis for further discussion. This research aims to address that problem by using some of the most recent developments in the computational technology and methodology of corpus stylistics, combined with a systematic and thorough manual analysis, to try to ascertain whether there is any empirical evidence for some of the most frequently recurring of these observations.

In 1960, Roman Jakobson published his clarion call for stylistics, or the linguistic analysis of literature, declaring that "a linguist deaf to the poetic function of language and a literary scholar indifferent to linguistic problems and unacquainted with linguistic methods are equally flagrant anachronisms" (1960, 377). Despite the frequency of citation of that announcement since, and the flowering of the discipline of stylistics, it is still unusual for rigorous linguistic analysis to be incorporated in literary interpretation. It can only be a positive development, therefore, that some stylisticians (such as Mahlberg (2007), Toolan

(2009), Archer (2007) and Mandala (2007)) are now advocating and practising the incorporation of literary critical responses in their linguistic analyses of literary texts. Similarly, some literary critics are also beginning to embrace not only linguistic but corpus computational methods of enhancing literary interpretation, and these studies are finding their way into literary publications (for example Witmore and Hope (2007)¹). This research focuses on the linguistic topic of syntax, but nonetheless aims to contribute to this disciplinary integration, however modestly, by addressing as fully as possible the relevant concerns in both the literary and linguistic fields.

1.1 The late plays: group identity

Despite both the First Folio's disregard for the chronology of Shakespeare's canon, and its ascription of varying generic categories to the late plays (*The Tempest* and *The Winter's Tale* as comedies, *Cymbeline*, *King of Britain* as a tragedy, and *The Life of King Henry VIII* as a history), and despite its omission of *Pericles* and *The Two Noble Kinsmen* altogether, these six works written between 1607 and 1613 are now most often seen as having some kind of group identity. Philip Edwards, for example, claimed that the late plays "seem more closely related than any other group of Shakespeare's plays" (1958), but debate as to which plays belong in the group has been prevalent for over a century. Edward Dowden identified them as comprising Shakespeare's "Fourth Period" and applied the generic term "romance" to *Cymbeline*, *The Winter's Tale*, *The Tempest* and *Pericles* (Dowden, 1901²) based on their shared thematic preoccupations. For Lytton Strachey in 1904³ the group

¹ Discussed further in sections 1.3 and 3.0 below.

² Originally published in 1875, 12th edition in 1901.

consists of three plays - *Cymbeline*, *The Winter's Tale*, and *The Tempest* - and three fragments - the Shakespearean parts of *Pericles*, *Henry VIII.*, and *The Two Noble Kinsmen*. All these plays ...form a distinct group; they resemble each other in a multitude of ways, and they differ in a multitude of ways from nearly all Shakespeare's previous work.

(Strachey 1904, 417).

Since then, however, there has been protracted critical disagreement over which plays to include in the group. E.M.W. Tillyard in 1938 admitted only *Cymbeline*, *The Winter's Tale* and *The Tempest*. Derek Traversi (1954) elected to treat the four 'romances' identified by Dowden (Tillyard's three with *Pericles*) separately from the final two. This choice was popular through the middle decades of the twentieth century, commonly defended in terms of perceived differences in genre, and of the collaborative authorship of *Henry VIII* and *The Two Noble Kinsmen*, the last two to be written. Curiously, though, the objection of collaborative authorship is commonly overridden in the case of *Pericles*, also co-authored, which is usually included in the core group.

One mid-twentieth century scholar who ventured beyond the boundaries of the four romances in his choice, although for what seem rather unusual reasons, was G. Wilson Knight, as Raphael Lyne, in his recently published *Shakespeare's Late Work* points out:

G. Wilson Knight's *The Crown of Life* [Essays in Interpretation of Shakespeare's Final Plays] stretches its canon to include *Henry VIII*, but it does so in order to define another climactic point in Shakespeare's career [other than the finality of *The Tempest*]... The *Crown of Life* was written mainly in summer 1944, and the presence of war is felt in the book's profound patriotism.

(Lyne 2007, 14).

³ As McDonald points out, "[t]here is confusion over the date of the essay, chiefly because the reprint in Strachey's *Books and Characters: French and English* (London: Chatto and Windus 1922) gives an erroneous date (1906) for the original publication. It first appeared in August of 1904" (2006, 13).

Knight saw in Archbishop Cranmer's prophecy in the final scene of *Henry VIII* an account of "an undying purity and blessedness settling here, on England" (256), suggesting that his choice was heavily dependent on the historical context of his writing.

Frank Kermode in 1971 asserted that "the Last Plays...form a distinct group, being, in the words of Philip Edwards, 'more closely related than any other group of Shakespeare's plays'"(219) although he regarded *Henry VIII* and *The Two Noble Kinsmen* as only "loosely associated" with the others, and for him 'The Final Plays' also meant only the four romances. Robert Adams later followed this choice of grouping, as is clear from the title of his (1989) *Shakespeare: The Four Romances*. Anne Barton, however, in her essay "Enter Mariners wet": realism in Shakespeare's last plays' took the broader view and included all six late plays. For Barton the group is linked in terms of the stagecraft and special effects they demanded:

That the King's Men, working at the Globe and at their new indoor theatre in Blackfriars, were now capable of much that would have seemed almost miraculous to the audiences of the 1580s and even 1590s is obvious from the texts of *Pericles*, *Cymbeline*, *The Winter's Tale*, *The Tempest*, *Henry VIII* and *The Two Noble Kinsmen*

(Barton 1994⁴, 196)

In his 1990 monograph *Staging Shakespeare's Late Plays*, Roger Warren (later editor of both the Oxford *Cymbeline* (1998) and *Pericles* (2003)), opts to exclude the last two plays. Although he asserts that there has been considerable interest in staging the late plays as a group, and believes that "they seem to belong together more closely than any other group of Shakespeare's plays except the histories" (1990, 5), he too is really only talking about the romances. *Henry VIII* is discussed and found wanting, with a citation from the theatre critic Benedict Nightingale for the case against the play: "[L]ate Shakespeare's preoccupations can

⁴ Originally published in 1986.

only be deduced from *Henry VIII* with some effort: they don't stand out as they admittedly do at times in *Pericles...*" (1990,7). *The Two Noble Kinsmen* is not even mentioned. Balancing this view of *Henry VIII* is the *Oxford Companion to Shakespeare* entry for that play, which calls it "the most romance-oriented of all the histories" (Dobson and Wells, 8).

The inclusion or otherwise of plays in editions of the complete Shakespeare works reveals another aspect of the picture, particularly concerning the last two collaborative plays. *Pericles* is included by editors throughout the twentieth century, even given its non-appearance in the First Folio, and the fact that it is a dual-authored play. The inclusion of *Henry VIII* in the *First Folio* has ensured the unequivocal acceptance of that play in the general canon, despite its dual authorship, although complete editions reveal different attitudes to its inclusion in the late group. Those that list the plays generically often include it as a history, as is the case with the 1942 Riverside edition, for example (Neilson and Hill 1942) and the 1969 Pelican complete works (Harbage 1969). Even the recent Pelican editors, Orgel and Braunmuller (2002), prefer to keep *Henry VIII* in the histories list, and out of the late group, as does Bevington (1997), although Hardin Craig (1961) goes to unusual lengths to delineate genres, including *Henry VIII* as "A History Play of the Fourth Period", under a more general heading entitled "The Period of Romances". *The Two Noble Kinsmen*, on the other hand, was not welcomed in from the apocrypha to the canon until fairly recent decades, so it is not surprising that critics have been slower to include it in the late group. The Signet complete works of 1972 was an example of adventurous early inclusion of the play, and all modern editors of the last two decades now tend to follow this lead.

Parallel to this change in attitudes to *The Two Noble Kinsmen* in the last years of the century, and largely due to challenges from new historicist and post-structuralist approaches to

the canon, a shift in critical opinion occurred away from the generic concept of romance, and towards inclusion of the final two collaborative plays in notions of the late group, as can be seen from the fact that writers on the late plays wishing to exclude them make at times rather tenuous, even whimsical cases. Kiernan Ryan, in his 1999 editorial introduction to *Shakespeare: The Last Plays* believes that the four romances share “a family resemblance sufficiently striking to warrant all four texts being treated, by common consent, as engrossed in the same basic endeavour” (1) and argues interestingly that “we have not yet mastered their formal grammar and poetic idiom, and so have not yet learned how to read them” (18). However, he excludes *Henry VIII* and *Two Noble Kinsmen* on the grounds of their having been written collaboratively (although *Pericles* is included), conceding then that “[a] strong case could doubtless be made for expanding this group to include [those plays]”, going so far as to admit that their omission “leaves this reader on Shakespeare’s last plays open to the charge of being incomplete” (1-2).

Writing at almost the same time as Kiernan Ryan, Simon Palfrey in *Late Shakespeare: A New World of Words* (1997) also debates the choices and opts to exclude the last two plays, but not on the basis of any properties of the plays themselves:

The critical conventions surrounding the four earlier works require, I think, sustained challenge; sticking to the traditional group has helped define my work’s scope. Furthermore, the nature of my task has made necessary extensive explications of each of the plays. Simple restrictions of time and space have contributed to my policy of exclusion.

(Palfrey 1997, 31)

A useful account of some of the issues at stake in assigning members to the group is offered by Jennifer Richards and James Knowles:

The late plays are often considered to form a distinct group because they share common themes and motifs – death and rebirth, family unity, repetition, time, riddles – a theatrical emphasis on spectacle and music, and a reconciliatory or optimistic mood. There are, however, many problems in this grouping, not only in its ‘thematic’ coherence (itself a matter of interpretation) but particularly in the key matters of dating and naming which themselves shape these interpretations. For instance, if we accept an early date for *Pericles* (1607), it precedes both *Coriolanus* (1608), a play which seems antithetical to normal critical perceptions of the late plays, and the revised (Folio) version of *King Lear* (c.1610).

(Richards and Knowles 1999, 2).

The conflict between strict chronology and thematic content as criteria for selection can perhaps be seen here. Interestingly, although they do not ultimately include the play in their grouping, Richards and Knowles proceed to argue the similarities of the Folio *Lear* with the romances (not easy to reconcile with their ascription to the late plays of “a theatrical emphasis on spectacle and music” or “a reconciliatory or optimistic mood”). The problem of simple adherence to chronology is clearly as great in that case as with *Coriolanus*. But they continue, with a compromise between the poles of the polemic, including the later plays in the group to be discussed in their volume, but identifying differences:

Similar problems follow their classification. We describe them here as ‘late’ Shakespeare, but we would like to distinguish between the early plays in this group, the so-called ‘romances’, and the three last plays identified above. Elsewhere, however, they are variously depicted as Shakespeare’s later, last or final plays, the romances, the late comedies, and the tragicomedies.

(Richards and Knowles 1999, 2).

There is a third play considered here after the four ‘romances’ because in their anthology they include an essay on the social and political context of *Cardenio*, believed to have been another Fletcher and Shakespeare collaboration from 1612-13, now lost except for “tantalising glimpses” in a probable adaptation of 1728 (Dobson and Wells 2001, 66) but not considered as

a contender for group inclusion in this research because the focus here is textual, and there is no confirmed Shakespeare text in existence in that case⁵.

Raphael Lyne, writing recently on the late plays, asserts that “there are characteristics special to Shakespeare’s late work, and these characteristics might well fit with ‘lateness’ in a literary career” (2007, 10) but argues that “one should not overstate this: it is impossible to show that Shakespeare managed his career to this extent, and there is no pressing need to consider these works as anything other than coincidentally ‘late’” (11). The well-worn biographical approach to Shakespeare’s last plays, referred to by Lyne here, is tackled at extended length in the recent monograph *Shakespeare and the Idea of Late Writing: Authorship in the Proximity of Death* by Gordon McMullan (2007), but (more to the point in the present research) McMullan identifies the group in the following terms:

Why do critics (with a few honourable exceptions, but only a few) focus solely on the four plays dating from 1608 to 1611 – that is, *Pericles*, *The Winter’s Tale*, *Cymbeline* and *The Tempest* (and often, in fact, on only three of these, casting *Pericles* adrift) – and not the two (or three) plays that postdate them – *Henry VIII*, *Cardenio* (lost, but perhaps not wholly irrecoverable) and *The Two Noble Kinsmen*?

(McMullan 2007, 4).

He argues that critics’ reluctance to include works after *The Tempest* results from concerns about collaboration, the difficulties involved in deciding on a “clear, overarching definition of ‘late Shakespeare’” (4) and from “an unspoken sentimental belief that *The Tempest* represents Shakespeare’s valedictory gesture... Any play that post-dated *The Tempest*, therefore, was considered *de trop*, in frankly poor taste, as unappealing as an ageing pop star’s comeback tour” (4). It is interesting that, a hundred years earlier, Lytton Strachey should have argued

⁵ The very recent publication of *Double Falsehood: or The Distressed Lovers* (2010, ed. Brean S. Hammond, London: Arden Shakespeare) may make that text a contender for future analysis in the same way as the present research.

against the biographical approach in not entirely dissimilar terms, also asserting an identical late grouping. It seems that “the tacit assumption, that the character of any given drama is, in fact, a true index to the state of mind of the dramatist composing it” (Strachey, 41) has been largely responsible for the frequent rejection of the co-authored plays from the late group for the century following Strachey, despite it still being the case that, as Strachey says, “the validity of this assumption has never been proved” (42).

Raphael Lyne declares an intention to “explore aspects of Shakespeare’s late work that separate it from the rest of his career” (2007, 24) and offers a detailed discussion of the grouping problem. He finds that “there are characteristics special to Shakespeare’s late work” but concurs with others in arguing that “the late work is not a closed or stable category” (11). While the end of the canon seems largely unproblematic, he says, definition of the beginning of the late period is problematised by the possibility of a diverse range of potential links between the six late plays and other works from around that period (23). Lyne argues (with Richards and Knowles) that temporal proximity is one such link, so that the probable late revision of *King Lear* in 1610 “makes the Folio *Lear*, in effect, late Shakespeare” (25), and the publication of the book *Shakespeare’s Sonnets* in 1609 raises the possibility of late writing revisions there too (29). He also sees the potential for thematic links between the last six and *Antony and Cleopatra*, “most likely written in 1606, [and] featur[ing] an interplay of spectacle, wonder, and sceptical reason that has much in common with extravagant displays in *The Winter’s Tale*, *The Tempest*, and elsewhere” (24). However his overview of the situation with regard to the later works in the group is expressed thus, affirming in the final sentence here the unity of the last six plays:

The Tempest has some climactic and final qualities, but it is by no means the last play Shakespeare wrote. He wrote at least two, and probably three, later plays in collaboration with John Fletcher, his successor as the leading dramatist of the King's Men. *Henry VIII*, *The Two Noble Kinsmen*, and what little can be made from the difficult case of the lost play, *Cardenio*, all have qualities that connect with those outlined as the core of the late Shakespearean mode.

(Lyne 2007, 11).

Lyne's approach to the grouping, and the discomfort of Richards and Knowles with the inclusion of the 1613 plays, are encompassed in the assessment of the situation offered in Russ McDonald's *Shakespeare's Late Style* (2006), an extended study of some aspects of the language of the late plays, considered separately for its linguistic analysis in the next section. McDonald provides in his introduction an extensive account of the critical history of the plays over the last century. He charts the biographical and allegorical espousals of the generic category of 'romance,' the adoption of 'tragicomedy' as an alternative nomenclature, the rejection of 'romance' and its associated notions by the post-structuralists and cultural materialists, but the tenacity of the term nonetheless, and concludes that the six as a group are united stylistically, with certain caveats:

Henry VIII has enough in common with the four major romances to warrant consideration along with them, as commentators since Coleridge have recognized: he referred to it as "a sort of historical masque or show play." But it would be idle to deny that it is primarily a history play and that its effect on audiences differs significantly from that of *Pericles*, for example. *The Two Noble Kinsmen* resists all the familiar Shakespearean categories, although its heroic story, taken from Chaucer and Boccaccio, and its mixed ending give it certain affinities with the romances and with other examples of Jacobean tragicomedy. Stylistically, at the level of sentence and line, these two plays exhibit most of the same properties as do the four romances that precede them.

(McDonald 2006, 25).

So it seems that although Shakespeare's canon does not helpfully divide itself into segments for our convenience, there is nonetheless now a certain amount of consensus that

these six last plays have a distinct identity. The three most recent accounts of the grouping discussed above informed the initial choice in the present research of the last six plays as the group for analysis, although it must be emphasised that any ascription of grouping remains somewhat arbitrary pending the establishment of more empirical foundations for a decision. In Chapter 5 a different perspective on the grouping is employed and tested, relating to the possible inclusion of four slightly earlier plays, *All's Well That Ends Well*, *Antony and Cleopatra*, *Macbeth* and *Coriolanus*.

What can also be seen from the foregoing, however, is that most of the above series of scholars share a difficulty in articulating with any specificity or concreteness the precise features found in the late plays. Thematic preoccupations are clearly a strong candidate, but even these are not seen as uniform or predictable across the group. With McDonald's account here of the group identity in terms of "the level of sentence and line," however, another aspect emerges that critics have for some time believed the plays to have in common, namely their language.

1.2 The language of the late plays

The Early Modern linguist Norman Blake recently asserted that "[t]he language of Shakespeare is a topic to which literary scholars pay lip-service, though they rarely pay as much attention to it as it deserves" (2002, x). This thesis enquires into the language of the late plays, because there is a widespread perception amongst generations of literary critics that what makes them distinctive has something to do with language, that something changes in the linguistic fabric of the Shakespeare canon around the time of their writing. Yet there seem to

be no attempts to describe exactly, empirically, what the linguistic changes are in the late group.

Revisiting Strachey's early twentieth century account of the "final period" we can see that he distinguishes this group in terms of its language. Although his famous invective is usually understood as uncompromisingly disparaging, it is actually employed for its rhetorical effect, contrasting as it does with his eulogistic praise of the language of these plays, as we can see in his use of adjectives such as "faultless":

It is difficult to resist the conclusion that he was getting bored with himself. Bored with people, bored with real life, bored with drama, bored in fact with everything except poetry and poetical dreams. He is no longer interested, one often feels, in what happens, or who says what, so long as he can find a place for a faultless lyric, or a new, unimagined rhythmical effect, or a grand and mystic speech.

(Strachey 1904, 52).

And he argues similarly that the denouement of *The Tempest* "is hardly more than a peg for fine writing" (47).

James Sutherland's 1959 article 'The Language of the Last Plays' identifies "a kind of writing that seems to be almost peculiar to the later plays" (144). Sutherland finds a different linguistic quality from Strachey, but like him, often conceives his observations in terms of conjecture as to the dramatist's mental processes. For Sutherland there is: "a sort of impressionism," (147), "an almost violent forcing of expression" (148), "rapidity of composition", a "recklessness of expression" (149), and "violence or abruptness or obscurity" (150). He summarises by saying that

The Shakespeare that I am offering for contemplation is a writer who, in those last plays, sometimes appears to be swept along by some inner compulsion, and whose mind seems at times to be generating an immense energy which he is applying, as a man might apply a pneumatic drill, to the immense problems of composition.

(Sutherland 1959, 151).

Both Strachey and Sutherland give examples of the effects they find in the plays, but articulate what they see in affective, evaluative and metaphorical terms. Readers who might wish to identify and study further the distinctive features of these interesting late plays are unfortunately left empty-handed.

Frank Kermode finds in the “final plays...a metrical freedom which goes beyond anything in the earlier plays” (1971, 219), and in his much-cited introduction to *The Tempest* declares that “the last plays...exhibit a control of language and imagery formerly unequalled” (1954, lvviii). Stephen Orgel sees *The Winter’s Tale* as “syntactically and lexically often baffling” (1996, 7) and he finds a “sort of linguistic opacity” to be “a feature of the play” (9). Roger Warren discovers comparable features in *Cymbeline*, such as “verbal complexity” (1998, 21), also saying that “its language...ranges from convoluted complexity to limpid simplicity” (21), and he concurs with Sutherland in finding “impressionism” (21) in the language of the play. Virginia Mason Vaughan and Alden T. Vaughan describe the language of *The Tempest* as “elusive”, saying that it “creates a dreamlike effect” and that “the verse is often elliptical” (1999, 20-21). In the introduction to her influential 1968 edition, Anne Barton talks at some length about the language of *The Tempest*, finding the play as a whole “deliberately enigmatic” (12), and saying that

This state of affairs derives in part from a compression, a stripped-down quality, more extreme than anything else in Shakespeare's previous work. Linguistically, the play is remarkably intense. Over and over again, the verse achieves an uncanny eloquence by way of what it omits or pares away.

(Barton 1968, 13).

She comes closer than many other critics to objective description of linguistic effects, identifying for example in an extract from act one the "suppression of the expected personal pronoun 'I' after two participial phrases which have seemed to lead deliberately towards it"

(13). She writes that:

The Tempest deals in compound words: 'sea-change', 'spell-stopped', 'cloud-capped', 'pole-clipt', 'still-closing', 'hag-seed', 'ever-angry', 'man-monster', 'sight-outrunning', 'sea-sorrow.' In compounds like these – and the play is filled with them – nouns strike against nouns or other parts of speech with an immediacy which jars both components out of their accustomed meaning. Sometimes this stark relationship is signaled in the Folio text by a hyphen, sometimes not. The important factor is the brusque juxtaposition of two words neither of which appears to modify or be syntactically dependent on the other in any normal sense

(Barton 1968, 14).

Barton concludes her essay on the late plays 'Leontes and the spider' by remarking on "the subtlety with which Shakespeare has adjusted his language and dramatic art to the demands of a new mode" (1980, 180), and elsewhere characterises Shakespeare's philosophical and dramatic approach to language, in *The Tempest* in particular and the late plays in general, by saying that it "foreshadows the techniques of modern theatre" in its "insistence on the difficulty of communication" (1971, 68-9).

In the editorial introduction cited in the previous section, Kiernan Ryan, in many respects a cultural materialist critic but one who (unusually) argues for a synthesis of formal and political analysis of Shakespeare's final plays (1999, 19), declares:

It is to the deliberate detail of their language and form that we must look, if the last plays are to be released from both the retrospection of old and new historicism and the abstractions of the allegorists.

(Ryan 1999, 18).

Indeed the essays that follow in his volume do focus far more on the texts of the plays than has been the tendency amongst the critical traditions he mentions. But none of them goes as far as to describe the plays in fully linguistic terms. Instead, very interesting literary readings of the language are offered: dramaturgical, political, psychological accounts that illuminate contextual and intertextual phenomena, but what they do not accomplish is an objective linguistic account upon which subsequent commentators could base their own hypotheses.

J.H.P. Pafford finds similar qualities to those identified by Barton, in his introduction to *The Winter's Tale*: “the tangled speech, the packed sentences, speeches which begin and end in the middle of a line, and the high percentage of light and weak endings” (1963, xxiii) and he goes on to make stronger generalisations about the group of plays as a whole: “[I]n *The Winter's Tale* the language has all the marks of that of Shakespeare’s late period: the play could be given an approximate date from its language alone” (lxxxiii). The prosodic features identified are specific, but he does not reveal any further detail about the means he would use to conduct such an ascription.

A tendency towards impressionistic accounts of the language is not only a feature of older criticism. In one reference to the language of the plays, Raphael Lyne describes *The Two Noble Kinsmen* in the following terms:

Alongside the stillness and gravity there is the rhetorical intensity that can be seen elsewhere in the romances, particularly in scenes where a spiritual scale is approached. The language – ‘mute contemplative’, ‘maculate’, ‘scurrile’, ‘bride-habited’ – is more ornate than usual. It works in the same way as Cranmer’s prophecy, creating a heightened atmosphere very much in keeping with climactic moments in romance.

(Lyne 2007, 19)

Such accounts are certainly evocative, and can strike sympathetic chords in the reader, but they do not offer any foundation of linguistic evidence upon which a further discussion could be based. For example, it is not quite clear in what sense the listed examples are “ornate”, what the “usual” style consists of, or indeed where it can be found. Many respondents agree that there is surely something distinctive about the language of the late plays, but greater specificity is needed than this. That the numerous commentators mentioned above identify features of the plays’ language is surely interesting, and their comments are often insightful, but little objective substance is offered on the whole. There are links between some of the observations made, for example between “elliptical” and “paring away”; and “enigmatic”, “baffling” and “obscurity”, which concur with some of the findings of McDonald below, and will be pursued later here. But essentially, although there is perhaps enough material in common in the critical responses to provoke further investigation, nothing offered so far could provide the basis for the kind of detailed linguistic analysis that would be needed in order to establish features of the plays’ language on a more empirical footing.

The considerable interest in the language of the late plays is evidenced by the presence of some book-length studies focusing on that topic, such as Maurice Hunt’s 1990 *Shakespeare’s Romance of the Word*. Hunt’s title suggests that language is the pre-eminent arena for discussion, but this is in fact not principally a linguistic study, despite the fact that Hunt is influenced in an interesting way by certain linguistic philosophers. He suggests for

example that “[i]n the field of Shakespeare studies, [the] argument that verbal models determine ways of perceiving and interpreting not only literary texts but reality as well unintentionally validates the Sapir-Whorf hypothesis” (16) and he makes some reference to the speech act theories of Austin and Searle (e.g.17, 24). But this is in most respects a literary critical account of the plays, and so it does not attempt to provide empirical evidence for their linguistic identity.

Another full length account of the late plays that identifies and discusses them in linguistic terms is Simon Palfrey’s *Late Shakespeare: a New World of Words*, (mentioned briefly in the previous section). He says “the modern academic tradition has lost access to something of the plays’ keenness and vim” (2) and sketches his aims in terms of wanting to “tease out the politics of a particular phrase or character”, to discover “how the language operates ‘in and around’ its unique spoken moment” (1990: 10). Palfrey’s principal approach is deconstructive, and he focuses on linguistic tropes such as metaphor frequently throughout the study, but his interest in their language is literary rather than linguistic.

There are, of course, accounts of Shakespeare’s language from a historical linguistic perspective, most often as part of studies of Early Modern English in general. However they rarely treat on diachronic or stylistic change in the canon. For example, Cusack (1970) gives an interesting account of Shakespeare’s use of linguistic variants in the English of his day, convincingly demonstrating it as a significant tool in his dramaturgy and characterisation, but she does not consider the changing patterns of such usage through the canon. Mulholland’s interesting article on second person pronoun usage in *Much Ado About Nothing* and *King Lear*, for example, (1967) does not take account of the fact that the *thou* form was a declining variant in Shakespeare’s lifetime, (as discussed in Hope 1994). An analysis of the kind she

performs, but of an early and a late play instead, might reveal interesting results. Elsewhere there are occasional brief, tantalising references to the possibility of differences through the canon, for example in Kathleen Wales's article on Shakespeare's use of conversion:

Particularly numerous are his shifts from nouns to verbs, which suggest that his own 'world view' tended to be 'dynamic' rather than 'stative'... Some of Shakespeare's most striking examples of 'dynamic' conversion are found in the later plays, particularly the tragedies

(Wales 1978, 396)

But here, as elsewhere, the idea is not developed. Estelle Taylor's contribution, 'Shakespeare's use of 'eth' and 'es' endings of verbs in the First Folio' is a notable exception to this rule, since she argues that "one may trace chronologically the author's dramatic change or shift in attitude toward the use of the old formal *eth* ending of verbs and the newer colloquial *es* ending" (1976, 437), but the late plays are not identified as a distinct group.

In recent years there seems to have been a resurgence of interest in Shakespeare's language, with the publication of several accounts of the subject, from different perspectives. Frank Kermode's *Shakespeare's Language* (2000) is another book about literary aspects of Shakespeare's language rather than a linguistic analysis. Unlike other writers, his focus is commendably on Shakespeare's development, although for some reason "what happened in the fifteen years or so between *Titus* and *Coriolanus* is the main subject of this book" (13). In his introduction Kermode describes the language of *Coriolanus* as possessing qualities (reminiscent of Orgel, Warren and Pafford earlier) such as "obscurity of syntax and vocabulary" (14) and again "these refined and occasionally excessive obscurities" (21), and promises much with "they are also found in Shakespeare's late plays", but then these issues are not addressed in his sections on those plays.

Further evidence of recent interest in Shakespeare's language is the fact that, in response to the gap of 130 years since E.A. Abbott's *A Shakespearean Grammar* (third edition 1870), two new accounts of Shakespeare's grammar have appeared: Norman Blake's *A Grammar of Shakespeare's Language* (2002), and Jonathan Hope's *Shakespeare's Grammar* (2003), although diachronic stylistic changes are beyond the scope of such works. Jonathan Hope, however, also provides the introductory essay to *Shakespeare and Language* (2004), a collection of contributions to the journal *Shakespeare Survey*⁶. In a rare and important articulation of linguistic and literary interests in Shakespeare criticism, he addresses the differences between literary and linguistic approaches to Shakespeare's language, raising the centrally important issue of empirical method:

Throughout these essays, there is a tension between literary and linguistic approaches to texts, which could crudely be represented as one between counting and not counting. Linguists see language as an object of study, a thing in itself, while literary scholars are more likely to treat language as a gateway to some other object of study. Even when linguists and literary scholars both fix on language as an object of study, they are likely to read it differently: linguists for features which can be shown to be characteristic by way of their relative frequency in contrast to relatively infrequent items; literary scholars for relatively infrequent items which can be argued to be salient in some way

(Hope 2004, 15).

He goes on to say: "I read these essays with great profit, but my most frequent marginal comment was probably 'Figures?' or 'Prove it!'" (15).

Another recent contribution to the interest in Shakespeare's language is Russ McDonald's *Shakespeare's Late Style* (2006), a detailed and insightful account of the language of the last plays. McDonald aptly identifies the gap in literary accounts of the late plays' language. Referring amongst others to Kermode (2000), cited above, he says that

⁶ Cusack, 1970, discussed above, is included in this anthology.

Those twentieth-century critics who did address themselves to the topic of Shakespeare and language rarely had much to say about late developments, a neglect that seems attributable partly to fatigue: however comprehensive their stated intentions, most of them begin to flag as they near the finish line

(McDonald 2006, 19).

McDonald takes as his late group *Pericles*, *Cymbeline*, *The Winter's Tale*, *The Tempest*, and after some caveats (25), *Henry VIII* and *The Two Noble Kinsmen*, although he does not seem centrally preoccupied with definition of the group. He includes extensive discussion of *Macbeth*, *Coriolanus* and *Antony and Cleopatra* as “late tragedies” (52), demonstrating their generic differences from the late plays, yet also their linguistic similarities with the late group at times (42-76). McDonald’s analysis deals with a rich mix of literary and linguistic features including rhetoric and prosody, and he identifies as typical in the late language “elision”, syntactic “divagation”, syntactic “suspension” and repetition. Although aspiring to “attention to stylistic detail” (18), McDonald declares his aim of “employ[ing] a non-technical vocabulary” and “avoiding the diction of linguistic scholarship”, suggesting that “in the study of Early Modern dramatic verse, such specialized terminology usually impedes rather than fosters clarity and understanding” (33). Technical rhetorical and prosodic terms are used throughout the work, however, so this caveat seems odd. Nonetheless, the book offers considerable insights into the late style. McDonald offers a series of interesting hypotheses, finding in the late syntax an “idiosyncratic mixture of contradictory effects” (134), with conflicting “aggregation and ellipsis” (135), simplicity and complexity, hypotaxis and “a high degree of subordination” (135) but at the same time “relatively simple coordination” (135) and asyndeton (e.g. 134). The range of the work is ambitious, and includes an attempt to link linguistic features with narrative and dramatic structures and dynamics, as well as biographical

insight concerning the author of the plays. McDonald invokes the narratologist Gerard Genette in his assumption that “the smallest grammatical and poetic detail not only correspond to larger narrative or dramatic preferences but also serve, especially in aggregate, as reliable indicators of an artist’s way of apprehending the world” (27). However, McDonald’s preferred method of supporting his (sometimes rather impressionistic) insights is textual exemplification, meaning that Jonathan Hope’s call to “Prove it!” (2000, 15) would be relevant here. In fact, in the original context of Hope’s plea, McDonald (a contributor to the anthology) was one of its addressees, who “have a tendency to make claims which linguists expect to see backed up with figures” (15). Importantly nonetheless, McDonald develops earlier critical observations much more fully and identifies in detail some of the features of “the increased complexity of the style, particularly the syntax” of the late plays (2006, 4). He also offers interesting observations about the “masculine”, “Senecan” and “laconic” language of the character Coriolanus (57-64), and the relationship or otherwise of Shakespeare’s syntactic style to characterisation in general (e.g. 96, 113).

In summary, then, the late plays have been perceived as a group with a distinct identity for over a hundred years, often with reference to their language, which is frequently seen as complex, opaque or convoluted. Recent responses from literary critical scholars have begun to articulate a more precise frame of reference for these perceptions, but still leave considerable work to be done.

1.3 Authorship attribution studies

Although there are as yet no accounts that attempt to describe empirically the linguistic qualities in the late plays, there is a vein of Shakespeare study, rather different from those

discussed above, which looks in detail at certain aspects of the language of the canon in order to distinguish his style from that of his contemporaries, in cases where authorship is undecided. Although there is a considerable variety of methods and approaches within this large field, these studies (sometimes referred to within the field as ‘stylometrics’) often use computational methods, with quantified, tabulated results, and tend to be produced by those who are “prepared to do the detailed and often tedious work involved in authorship studies, computing minute qualities of style” (Vickers 2002, 354). The present selective account focuses particularly on those works related in some way to this study, although fuller overview and discussion from different perspectives can be found, for example, in Vickers (2002), Jowett (2007, 17-26) and Archer (2007, 245-6).

There are increasingly vocal calls for a rigorously empirical approach to be taken in stylometric approaches to authorship attribution. According to Brian Vickers, an authoritative voice in this field, “[i]n recent years authorship studies have developed far more detailed methods for determining linguistic preference, and with a comparable increase in methodological awareness” (2002, 132) and he applauds attributionists whose “primary concern was the empirical study of the plays’ language” (2002, 397).

Marcus Dahl addresses the problem of attitudes to empirical methods in literary critical circles, in his recent thesis on the authorship of Shakespeare’s *King Henry VI Part One*:

Although most readers of ‘Shakespeare’ have a conception of Shakespeare’s language, they will seldom go to the trouble of attempting to locate that conception in an objective context. Moreover, when in fact a putatively ‘objective’ method is established (such as the testing of sub-literary features), there is often a distinct reluctance by more traditionally-minded literary critics to accept the significance of the results. Thus there can be a perceived divide between the dry word-counting and statistical charts of the stylometricists [sic], and the ordinary aesthetic/political readers of the accepted canon (although literary/editorial critics as diverse as Edward Malone

and Gary Taylor have employed enumerative and statistical techniques in order to endorse their literary-critical judgements).

(Dahl 2004, 2002-3)

Dahl is undoubtedly right in highlighting the reluctance of many literary critics to embrace computational findings. The Shakespeare editor and critic Gary Taylor's use of such methods, mentioned here, was described by the computational analyst of authorship attribution, M.W.A Smith, as "such academic bravery" (1991, 73) for the same reason.

Another important respect in which authorship attribution studies are relevant to this research concerns their interest in stylistic changes through the Shakespearean canon. The establishment of the chronology of the plays has always been a closely related preoccupation to that of authorship, and as a result there are fairly frequent, if fleeting, references to differences between Shakespeare's stylistic phases. Similarly, because three of the principal candidates for dual authorship (*Pericles*, *Henry VIII* and *The Two Noble Kinsmen*) lie in the late period of the canon, there has been some attention (albeit rather peripheral and unsystematic) to Shakespeare's 'late style'. A number of analyses of Shakespeare's metrical prosody and verse patterns have addressed the question of chronology in the canon (such as Jackson's (2002) work on pause patterns, and Tarlinskaja's (1987) study of the iambic pentameter), but the work has not so far been extended to other linguistic features.

Some studies have focused on lexical choices in Shakespeare's work, such as that of the nineteenth century German scholar Gregor Sarrazin (1897)⁷ who pioneered the use of comparative lexical analysis to determine the chronology of the plays, looking at frequency patterns of rare words (those occurring once, twice and three times in each play), and finding

⁷ Cited in Vickers (2002), in Vickers's translation from the original German. Originally published as 'Wortechos bei Shakespeare, I', *Shakespeare Jahrbuch*, 1897 (33), pp. 121-65.

that certain words were specific to the late phase. A comparable approach was taken some decades later by Alfred Hart in his ‘Shakespeare and the vocabulary of *The Two Noble Kinsmen*’ (1934), as he analysed the comparative frequencies of compound words and new usages in order to distinguish Shakespeare and Fletcher’s hands in the play. Another authorship test which comes interestingly close to focusing on stylistic development through the canon is Brian Vickers’s refutation of Donald Foster’s⁸ attribution of Shakespeare’s authorship to *A Funerall Elegye*, based on Vickers’s analysis of Latinate lexis in plays written close to the time of its composition in 1612. Vickers (2002, 60-79) analyses the three late plays known to be entirely by Shakespeare (*Cymbeline*, *The Winter’s Tale* and *The Tempest*) and finds that their use of Latinate words is significantly higher than in the *Elegye*, and the poem is attributed to Ford as a result.

In *Defining Shakespeare: Pericles as Test Case* (2003), Jackson argues for a new method of authorship attribution, based on the identification of short parallel passages that suggest identity of authorship. Jackson argues laudably that “[m]istakes in attribution arising from the haphazard and biased accumulation of verbal parallels can be avoided through systematic and comprehensive electronic searches” (193). His method involves selecting certain passages from the text under scrutiny and searching for them in the ‘Literature Online’⁹ database, which “contains all Shakespeare plays and poems that are commonly included in editions of the complete works” (196). He describes one such search for selected phrases from two passages from *Titus Andronicus*, (where debate exists as to whether George Peele or Shakespeare wrote certain scenes):

⁸ Foster’s assignation of the work to Shakespeare appeared in 1989 as *Elegy by W.S. A Study in Attribution*. (Newark and London).

⁹ <http://lion.chadwyk.co.uk>

Words, phrases, and collocations from the two short passages [in *Titus Andronicus*] were methodically keyed in, one at a time, to be searched in Peele's and Shakespeare's works... For example, for the play's opening line, 'Noble patricians, patrons of my right', searches were made of 'patrician', 'patron', 'my right', and of 'FBY right' (where FBY stands for 'followed by'), the contents of any hits being visited and checked. Where words are fairly rare, the best policy may be to search for the single word and look for collocations within the contexts

(Jackson 2003, 196)

There are, however, some problems with this approach. A comprehensive survey of the text is not in any sense achieved, since only individual passages are tested. Jackson does not tell us the criteria he employs for the selection of his words and phrases, so it would seem that the findings are rather subjective. Definitions of central terms and concepts used in the study are not explained, such as "linkage" (between words and phrases) and "collocation", nor what constitutes success, or a "hit" (e.g. 196). Linguistic analysis is not attempted, nor is any other model offered as a way of structuring the findings. As a result, the outcome is unfortunately as "haphazard and biased" (193) as many others. Very recent authorship attribution work by Craig and Kinney, employing the same text database as Jackson¹⁰ but with computer software processing the text to look for salient lexical frequency in Shakespeare, and using considerably more rigorous methods, includes a discussion of the parallel passages technique:

The parallel-passages method is attractive for literary scholars, since it pays attention to actual passages, rather than to wider patterns, and since...it allows readers to make their own assessment of the evidence. Intuitively it is plausible: of course writers will have characteristic turns of phrase, word pairings or extended phrases that are highly peculiar to themselves. In many cases, one must concede, it will give a true result... We believe, however, that it is a method that may also produce some false classifications. Used with no comparative data, it may be seriously deceiving... The method relies on what Vickers calls 'the accumulated evidence of highly individual usages in thought

¹⁰ The Literature Online database at <http://lion.chadwyk.co.uk>

and words occurring too often for coincidence'¹¹; this can be restated, in Elliot and Valenza's words, as 'Christmas trees full of unique quirks equals proof'¹²

(Craig and Kinney 2009, 59).

In his analysis of "function-word modifiers and complements in 500 word samples" of play texts by Shakespeare and Fletcher, W. M. Baillie aimed "to find stylistic features which consistently discriminate Shakespeare's late dramatic writing from Fletcher's of the same period" (1974, 75-76). However, Baillie does not give details of his sample selection method, nor of what he understands by "function word modifiers" or "complements". Neither does he go into detail about the technical functioning or reliability of the "Eyeball" tagging software he employs, which tends to render the findings suspect, since there are acknowledged problems for the analysis of Early Modern English even with the most recently developed taggers and parsers today (see for example Archer 2009, 13). There would therefore seem to be more questions begged by this work than answered. Nonetheless, the perception that the analysis of function, or grammatical, words can be a useful indicator of style is fundamentally important, and another approach shared by many authorship attributionists and stylisticians. Brian Vickers articulates the authorship perspective thus, making at the same time the key point that concordances and lexicons of Shakespeare, because compiled on the basis of deductive selection by their author, can be empirically unreliable:

But some authorship tests evade the normal reading experience, demanding special notation and statistical analysis. One such is the analysis of 'function words', those parts of speech which perform essential grammatical functions so unobtrusively that compilers of concordances regularly omit them. These include the definite and indefinite articles, conjunctions, prepositions, pronouns, auxiliary verbs, all elements

¹¹ Vickers 2002, page 205.

¹² Craig and Kinney cite the source of this quotation as "Elliott and Valenza, 'Oxford by the Numbers', p.339".

that recur frequently in any piece of verse or prose and which are used by many writers with distinct preferences

(Vickers 2002, 90).

Vickers himself uses an analysis of the function word ‘of’ in his case against Shakespeare’s authorship of *A Funerall Elegye*, mentioned above (2002, 92). In an attempt to provide a basis for establishment of the Shakespearean canon, Taylor (1987) conducted a function word analysis of Shakespeare’s core, undisputedly authored plays:

In the hope that they might prove useful in solving certain problems in the Shakespeare canon, we asked Dr. Marvin Spevack to provide us with relative frequencies, averages, and standard statistical deviations for 25 such common words: all, as, at, but, by, for, from, how, if, in, it, most, much, never, no, nor, not, now, or, out, so, some, such, than, that, the, then, there, these, this, to, upon, what, when, where, which, who, why, with. This information was derived from computer transcripts of the Riverside edition, from which Spevack had generated his *Concordance*. On the basis of the data Dr. Spevack supplied, we selected ten words for which Shakespeare’s usage was most consistent

(Wells and Taylor 1987, 80).

Taylor then reduced the group of twenty-five to “ten function words which Shakespeare uses relatively consistently” in order to “determine what proportion of the total vocabulary of a work those ten words normally constitute, and how frequently they occur relative to one another” (1987, 81). Asserting the success of the analysis, he claims that

it does represent an undeniable pattern in the uncontested works of Shakespeare’s authorship. From that perspective, Shakespeare’s acknowledged writing constitutes a homogeneous statistical population. Any candidate for inclusion in that population can be judged statistically, in order to measure the degree to which it fits that pattern

(Wells and Taylor 1987, 81)

This “homogeneous statistical population” is assumed here to be an unchanging sub-literary feature of Shakespeare’s style, with no acknowledgement of the possibility of diachronic development, a potentially complicating perspective to which only a few writers seem alert (e.g. Smith 1990, 198). The methodology and findings of this study have been criticised by

more than one writer, both for the weakness of the test as a positive determiner of authorship by Brian Vickers (2002, 94), and on the basis of its lack of empirical rigour, by the statistician M.W.A. Smith, who regards Taylor's methodological approach (principally concerning the selection of words) as indefensible (1991, 78).

Relative pronouns have been seen as indices of authorial style by a number of writers on the authorship question. H. D. Sykes (1919) identified in the writing of George Wilkins, Shakespeare's collaborator in *Pericles*, "the frequent omission of the relative pronoun in the nominative case" (1919, 150-152), something that he argues was not a feature of Shakespeare's style (although again there is no acknowledgement of a diachronic dimension in this case). In *Defining Shakespeare: Pericles as Test Case*, discussed above, Jackson analyses the frequency of a set of function words in samples from a series of plays by various Renaissance dramatists, although clear cut results were not forthcoming (2003, 116). More useful as a test was a count of frequency of *to* infinitives (2003, 119-123), which identified Wilkins as collaborator in *Pericles*. In another section of the work he identifies "three clearly definable types of unusual *which* in *Pericles*" (2003, 123) which, he argues, Wilkins uses more frequently than Shakespeare.

A widely-cited account of function words in this field is the socio-linguistic authorship study conducted by Jonathan Hope (1994). Hope analyses three sets of function words in Shakespeare and his contemporaries, aiming to establish authorship of, among others, *Pericles*, *Henry VIII* and *The Two Noble Kinsmen*. His hypothesis is that certain choices of relative pronouns, second person pronouns (*thou* forms as opposed to *you*), and certain types of usage of auxiliary *do*, are sociolinguistic and diachronic variants in Early Modern English, and are therefore potentially successful tests for authorship because they would differ

according to the age and generation of the writer (useful in distinguishing Fletcher and Shakespeare, for example, because the latter was fifteen years older than his co-author). Discussing the case of auxiliary *do*, Hope encounters the question of a late style, not because that is the principal object of his research, but as a result of the methodological rigour he employs. He recognises the possibility of the added complication in authorship attribution represented by diachronic change in an author's own style: "a problem which has been encountered by other workers seeking to use linguistic features as evidence for authorship... [H]ow can we know whether or not an author's usage of a feature is constant over his career, or whether he changes?" (1994, 21). In response to this problem he analyses auxiliary *do* usage in two groups of plays, an early and a late group, and discovers that the percentage of 'regulated usage' (1994, 11-15) remains sufficiently constant for diachronic change through the canon not to undermine the authorship statistics. In other words he finds that this usage is not a diachronically marked style feature. Elsewhere in Hope's study he examines relativisation strategies in first Shakespeare, then his alleged collaborators, and reveals in the process numerous insights into syntactic habits of these writers.

In a recent contribution to an anthology dedicated to Early Modern tragicomedy and co-written with Michael Witmore, Hope does address the issue of diachronic stylistic change directly, attempting to distinguish the late plays generically from the rest of the canon "not by qualitative analysis of tone, atmosphere or plot, but by quantitative analysis of linguistic features" (2007, 136). They use Docuscope, a "text analysis and comparison program...designed for use in writing/rhetoric classes (as understood in the North American model)" (137), which identifies a wide range of predetermined lexical and grammatical features in an electronic version of a text. Their main premise is that

such analysis calls attention to a heretofore invisible set of dramaturgical strategies at work in the late plays, strategies that mobilise language so consistently and on such a pervasive verbal level that their effects have gone unnoticed by more traditional literary genre criticism

(Witmore and Hope 2007, 136).

Witmore and Hope's findings in the late plays include "an increase in the direct representation of the past," identified apparently through the frequency of phrases such as "but also" and "at the same time" (146). They also notice "a significant increase" in the use of "strings of words including 'which' as a non-restrictive relative pronoun" over the course of Shakespeare's career, "particularly in the last plays" (146). The fact that the work is published in a literary critical collection may account for the lack of numerical tables (perhaps reflecting some of the problems articulated by Dahl above). As it stands, however, this approach to the late plays in terms of their background linguistic features is a very interesting and important step, developing Hope's 1994 study, discussed both above and in 6.2.2. As will be seen in Chapter 3 below, some of the findings of Witmore and Hope are also confirmed by the present research. The studies by Jackson (2003) and Hope (1994) are discussed at more length in the concluding chapter here, since they touch explicitly on the findings of this research.

Finally, Brian Vickers's influential work *Shakespeare, Co-Author: A Historical Study of Five Collaborative Plays* is a meticulous and comprehensive survey of two hundred years of attribution studies, including the majority of those discussed above, making the cases for collaborative authorship in five Shakespeare plays. Vickers consistently advocates increased empirical rigour in the field, "lament[ing]" on a number of occasions "the absence of system or method" in the writers he surveys (297), and he includes chapters on three of the late plays: *Pericles*, *Henry VIII* and *The Two Noble Kinsmen*. Substantial sections of these chapters deal

with metrical and phonological features of the plays, which (although extremely interesting) are not within the scope of this research. Similarly, much attention is devoted to the description by generations of attributionists of aspects of the style of Shakespeare's collaborators, rather than that of Shakespeare himself, in order to distinguish their work from his. But Vickers also reports on linguistic features that do have a bearing on this analysis, particularly syntax and function words in Shakespeare. Among many others, he reports findings that the Shakespeare sections of *Pericles* contain "awkward ellipses" (292) and have more complex syntax than that of Wilkins (308), both of which relate directly to observations made by other critics, cited in the previous sections here. This theme is reiterated with reference to *Henry VIII*, where Vickers cites Partridge describing Shakespeare's syntax as follows:

On the track of the telling and indelible image, he may leave behind anacoluthons and hanging relative clauses in the most inconsequent fashion; he compresses his meaning and tortures his syntax, so that while the effect of the passage may be poetically grand, the meaning is wrung from it with extreme difficulty

(Vickers 2002, 377¹³).

Very similarly, in his discussion of the authorship of *Henry VIII*, Vickers cites Jackson as saying that Shakespeare is "more apt than Fletcher to use longer, more complex phrasing, as sentences straddle line endings in energetic imitation of the tortuous movements of the mind" (Vickers 2002, 393¹⁴). These accounts have interesting similarities with the descriptions of the language of the late plays seen in 1.2 above, suggesting that some of the same features have been observed. The possibility that the Shakespearean style under scrutiny could be

¹³ From A.C. Partridge, 1964, *Orthography in Shakespeare and Elizabethan Drama: A Study of Colloquial Contractions, Elision, Prosody and Punctuation* (London).

¹⁴ Vickers is citing Jackson 1997, 78.

diachronically specific is not discussed, however, although elsewhere in the work authorship tests are reported that do incorporate chronological sensitivity (e.g. 307). One centrally significant case made by Vickers is for the importance of function words in the analysis of literary style, a perspective that stylometrics and stylistics have very much in common¹⁵:

Whether consciously or not, writers tend to reveal distinct preferences in the type of words that they use, and the frequency with which they draw on them. Comparable preferences, patterns of choice and avoidance, can be traced at less significant verbal levels, such as a seventeenth-century writer's liking for *ye* rather than *you*

(Vickers 2002, 80).

Vickers presents a number of accounts of authorship studies which examine function word usage in the co-authored plays, including those by Hope (1994) and Jackson (2003) considered separately in this thesis.

It can be seen from this survey that the approaches taken by analysts in the field of authorship attribution studies have some relevance to this research. However, if they offer several pointers for an analysis of Shakespeare's late style, they tend to do so obliquely, as possible starting points for searches, and at times as salutary reminders of appropriate methodological practice.

1.4 Summary

We see from the foregoing discussion that a number of literary and linguistic sub-disciplines bear on the language of the late plays. There is longstanding interest among literary critics in the notion of a distinct late group at the end of the Shakespeare canon, in its constituents and characteristics, and although empirical analysis is yet to be used to address

¹⁵ See, for example, Culpeper (2001, 199), and Scott (2008, Help Menu) cited in 3.1 below, p.49.

this question, some recent studies (such as McDonald 2006) offer important starting points. Historical linguists (some of whose work will form an essential part of the analysis later in this thesis) offer detailed accounts of his language from a number of perspectives, yet have not so far focused on identifying diachronic stylistic features of the canon that might illumine questions about the late group. From the point of view of method, increasingly rigorous studies carried out by authorship attributionists present models of practice that can be taken on board in an attempt to address the literary issues in this case. This research therefore can be seen to be positioned at the junction of these different fields. The next chapter presents in more detail the precise approach taken in this thesis.

CHAPTER 2

METHODOLOGY AND METHOD

2.0 General remarks

This thesis aims to test critical claims about the language of the late plays, by using corpus techniques to identify salient, relevant features in the late plays, should any appear. Appropriate linguistic models will then be used in order to investigate the findings manually. This chapter comprises first an account of the general methodological approach taken, then of the corpus stylistic techniques chosen as the initial part of this research, with accompanying methodological discussions there, also. Finally, in section 2.3, it offers a detailed account of the procedures employed in the preparation and execution stages of the corpus analysis.

One of the problems of an interdisciplinary approach is that the cautionary exigencies of not one but both disciplines weigh on the researcher and require accommodation. For Shakespeare studies, in the case of the present research this means principally editorial concerns relating to the instability of texts, and the resultant difficulty in approaching them with the aim of analytical comprehensiveness. From the point of view of stylistics, however, there are other problems to face. Having spent the first chapter of this thesis arguing for a more objective and empirical approach to the analysis of Shakespeare's language, it is now necessary to acknowledge the limitations of such an approach, indeed the pitfalls involved in

attempting a systematic statistical analysis of any aspect of a text. Both sets of issues will be addressed here.

2.1 Stylistics

This research takes as its methodological starting point the approach advocated by contemporary practitioners of stylistics, characterised in one sense by the following statement by Leech and Short, from their widely-cited introduction to prose stylistics, *Style in Fiction*, and echoing the sentiments expressed by Jonathan Hope in the previous chapter¹⁶:

Aesthetic terms used in the discussion of style (urbane, curt, exuberant, florid, lucid, plain, vigorous, etc) are not directly referable to any observable linguistic features of texts ... The more a critic wishes to substantiate what he [sic] says about style, the more he will need to point to the linguistic evidence of texts; and linguistic evidence, to be firm, must be couched in terms of numerical frequency

(Leech and Short 1981, 47).

Statistics cannot replace literary criticism, but they can substantiate its insights. As Michael Stubbs points out, in his analysis of Conrad discussed in the next section, “quantification can make more explicit the evidence on which interpretation is based” (2005, 6). And a sub-literary style feature can most often be revealed only by statistical analyses, as Burrows also argues in his study of Jane Austen’s novels (1987, 3). But stylistics goes beyond the purely statistical analysis of texts, such as that seen in the discussion of authorship attribution studies in section 1.3. Stylistics has a theoretical underpinning which is missing in purely computational work, evidenced in the following account by Carter and Simpson. Not only do linguistic accounts of literary texts need to be supported by statistical evidence: it is important for a stylistic analysis to

¹⁶ See section 1.2, page 19.

aim to be sufficiently detailed, explicit and *retrievable* for other analysts, working on the same texts, to check or retrieve the original analytical decisions and procedures. These other readers may not indeed share the writers' intuitions, and this may lead them to their own different interpretations of the texts; but it is regarded as essential that they should be able to follow the steps by which particular analyses are made. And because the analysis is systematic and according to clearly defined models or procedures, such readers are in a position to argue against the positions adopted...should they wish to do so

(Carter and Simpson 1989, 14).

Deirdre Burton goes further, highlighting in her "four crucial criteria to be borne in mind" (1980, 119) the need to account for as much of the available data as possible. This means that not only is exemplification (typically seen as sufficient evidence in literary criticism) unlikely to be an adequate means of supporting an analytical case, but that where possible, a description of all of the data, rather than a sample, is desirable.

But stylistic theory argues that this approach is not without its attendant problems. Leech and Short outline the difficulties in making assertions such as that a writer "tends to use short sentences" for example, since without a reliable point of comparison (how might one determine the average length of an English sentence?) that seemingly objective claim is almost meaningless (1981, 43). They also caution against attempts to "provide an objective measurement of style" (47), since the concept is too elusive and all-embracing to be susceptible of accurate definition in anything like its entirety, and (citing Dr. Johnson) point out the pitfall of "the stately son of demonstration, who proves with mathematical formality what no man has yet pretended to doubt." (47). They conclude that "[t]he essential point is that the use of numerical data should be adapted to the need." (48).

Jonathan Culler and Derek Attridge derive salutary arguments concerning the pitfalls of ostensible objectivity from critiques of Roman Jakobson's formalist approach (exemplified

in his famous Closing Statement (1960), cited in 1.0 above). Culler highlights the dangers (distinctly relevant to the present thesis) of unfettered and self-indulgent linguistic analysis:

[O]ne must seriously question the claim that linguistics provides a determinate procedure for exhaustive and unbiased description. A complete grammar of a language will, of course, assign structural descriptions to every sentence, and if the grammar is explicit two analysts using it will assign the same description to a given sentence; but once one goes beyond this stage and undertakes a distributional analysis of a text, one enters a realm of extraordinary freedom, where a grammar, however explicit, no longer provides a determinate method. One can produce distributional categories almost *ad libitum*

(Culler 1975, 57).

Furthermore, Attridge cautions against the illusion created by “the rhetoric of objectivity” which, “like all rhetorics, reveals itself as such in its language; to define is to attempt an act of exclusion, keeping out that which would threaten the logical form of the definition” (1987, 46¹⁷). Even an analyst with the technical expertise and scrupulousness of Jakobson must declare his vested interests, it would seem. Incorporated in Jakobson’s ostensible scientific rigour, according to Attridge, is a project “to persuade [his reader] by the amassing of detail that this verbal artifact [the poem] has remarkable qualities” (Attridge 1987, 46). Carter and Simpson make the same point, that “[t]he assignment of meaning or stylistic function to a formal category in the language remains an interpretative act and thus cannot transcend the individual human subject who originates the interpretation” (1989, 6). Ultimately all responses to a text are interpretations, determined by the manifold contexts and values of the analyst, and to pretend otherwise is to practice a deception. Where then does this leave the stylistician? If objectivity is impossible, are we then awash in relativism? Michael Toolan suggests not, arguing instead a dialectical approach, where “what is essential is a renewed self-critical

¹⁷ Page numbers refer to the 1996 publication in Weber (ed).

awareness of the provisionality of one's reading, of the roots of description in rhetorical persuasion" (1990, 11). The present research proceeds with this advice very much in mind, aiming to be "rigorous, systematic and replicable" (Weber 1996, 3), but also to appraise methods and findings realistically, and to acknowledge limitations where appropriate.

2.2 Corpus stylistics

Although much of this thesis comprises a detailed manual analysis of Shakespeare's language, the field of focus was determined by the findings of a corpus stylistic analysis, and the linguistic analysis itself would have been impossible without that technology, so an account of this relatively new discipline is presented first.

One of the longstanding limitations in the linguistic analysis of literary works has been the difficulty of coping with large bodies of texts. Close, detailed analysis was just not possible on a large scale. However, computer programs have appeared, particularly in the last two decades, which make possible the analysis of very large amounts of text. Some of these have been mentioned in passing in the previous chapter, in discussion of authorship attribution studies. J.F. Burrows (1987) was among the early group of writers to use this kind of technology to analyse literary texts from a stylistic perspective, with a study of function or grammatical words in Austen's novels. Louw (e.g. 1997) was one of the first to argue for the use of corpus methods to challenge literary intuitions. Both he and Michael Stubbs (2005) are among those stylisticians who have focused principally on the use of lexical as opposed to grammatical words, using one of the large general language corpora (in their cases the British National Corpus) to compare frequency and collocation in individual literary texts with contemporary background linguistic usage. The use of general corpora is one respect in which

corpus stylistics can overlap with its parent discipline, corpus linguistics¹⁸, although as Mahlberg explains:

Stylistics focuses on what makes a text, or a group of texts, distinctive, and it investigates deviations from linguistic norms that trigger artistic effects and reflect creative ways of using language. Corpus linguistics, on the other hand, mainly focuses on repeated and typical uses...

(Mahlberg 2007(a), 221).

There is however a software program, increasingly widely used among corpus stylisticians, that has not so far been taken up by authorship scholars, yet which provides a very valuable, highly transparent method of calculating word frequencies and related data in texts. This program is WordSmith Tools, created by Mike Scott in 1996, and now in version 5.0. Although the current version offers a number of more sophisticated analytical procedures, essentially it is structured around three main functions: Wordlist, Keywords and Concord. An electronic text is uploaded into the program and a wordlist is created from that corpus, showing the raw frequencies of all the words in the text. Although some interesting findings can arise from wordlists, as Jonathan Culpeper points out in his 2001 study of Shakespeare's characters, which uses corpus stylistic analysis, "high frequency words...have little stylistic significance, unless they are compared with an appropriate norm" (199). For example, the three most frequent words in the Shakespeare canon are likely to be the same as those in a Present-day English text of almost any kind, and usually appear as *the*, *and* and *I*. Not until the raw frequencies are compared with those of a reference corpus (usually a larger body of text than the first corpus) can more insight into salience and significance of individual frequencies be gained. Mike Scott himself describes the procedure: "The idea is quite simple: by

¹⁸ For a further discussion of this interrelationship see Mahlberg 2007(a), 219-222, and Archer 2007, 245-246.

comparing the frequency of each item [usually each individual word] in turn with a known reference, one may identify those items which occur with unusual frequency” (2009, 80). A list of keywords is thus produced, with the most frequent and infrequent items relative to the reference corpus listed in order of salience, or keyness. The choice of reference corpus is of course important, and will affect the results (Scott 2009, 80). Finally, Concord can then produce a set of concordance lines for any of the words in any of the lists generated. This displays the whole set of occurrences, with the word shown as the central (node) word in each occurrence, with co-text to either side. The individual instances can then be analysed manually.

There are several possible approaches to this kind of analysis, however, as Paul Rayson articulates very usefully. Corpus linguistics researchers, working in linguistic and not literary linguistic fields, with one of the large general corpora of contemporary language, as suggested above, typically proceed along the lines outlined by the first of Rayson’s process models:

[I]t begins with the identification of a research question, continues with building and annotating a corpus with which to investigate the topic, and finishes with the retrieval, extraction and interpretation of information from the corpus which may help the researcher to answer the research question or confirm the parameters of the model

(Rayson 2008, 521).

Depending on whether the researcher is focusing on the use of a particular linguistic feature, or looking at “whole texts or varieties of language” (521), this process can be either what Rayson calls Type I or Type II respectively. He identifies a third type, however, which has been used in the present research, and which he describes as follows:

[D]ecisions on which linguistic features are important or should be studied further are made on the basis of information extracted from the data itself; in other words, it is *data-driven*. I will call this Type III. It combines the approaches of Types I and II by first focussing on whole texts and then suggesting specific linguistic features to study in further detail. In other words, the ordering of the five main steps above will change to the following... 1. *Build*: Corpus design and compilation. 2 *Annotate*: Manual or automatic analysis of the corpus. 3. *Retrieve*: Quantitative and qualitative analyses of the corpus. 4. *Question*: Devise a research question or model (iteration back to step 3). 5. *Interpret*: Manual interpretation of the results or confirmation of the accuracy of the model

(Rayson 2008, 522).

The crucial difference in Type III is the fact that it is “data-driven”. In other words, the area for focus is not known until the initial results are produced from the corpus analysis. The researcher only then decides what to take further, based on whatever salient findings appear. One of the methodological advantages of this approach is that the area for analysis is based more on empirical evidence than would be the case if selected entirely according to the researcher’s choice. Culpeper highlights this aspect of keyword analyses in general:

[Keyword analysis] can reveal lexical and grammatical patterns without reliance on intuitions about either which parts of the text to focus on or what the relevant dimensions or features are. Many other quantitative methods – multi-dimensional analysis, for example – involve a priori decisions about what to count, but keyword analysis does not

(Culpeper 2009, 53).

Michaela Mahlberg, a corpus stylistician using WordSmith Tools to analyse novels by Dickens, also suggests a methodological approach to corpus analysis. With the aim of fostering integration of stylistics and literary criticism, she outlines the initial stages of a corpus stylistic analysis in the following helpful terms:

Literary insights and arguments from literary criticism can suggest items for a concordance analysis. Another option is to start with frequency information on words in the text and pick those whose frequencies appear noteworthy, as identified, for example, with the help of the KeyWords Tool in WordSmith... The preferred option is

a combination of corpus methodologies, previous corpus stylistic work and literary criticism

(Mahlberg 2007, 22).

Mahlberg's second option here "to start with frequency information" corresponds with Rayson's Type III, although Mahlberg's approach means the additional input of literary critical insights into the rationale for selection of a field of focus. The present research conforms largely to this "preferred option" profile, as will be shown in the following section.

Whatever their preferred choices of method, however, corpus stylisticians tend to have in common the insistence that the machine is secondary in importance to the researcher:

Purely automatic stylistic analysis is [not] possible. The linguist selects which features to study, the corpus linguist is restricted to features which the software can find, and these features still require a literary interpretation

(Stubbs 2005, 6)

As Mahlberg points out, "the application of a corpus methodology is not an automated process" (2007, 31). And Archer articulates this further, saying that stylisticians (in this case the contributors to her edited volume of 2009)

are not seeking (or wanting) to suggest that the procedures they utilize can replace human researchers. On the contrary, they offer them as a way in to texts – or, to use corpus linguistic terminology, a way of *mining* texts – which is time-saving and, when used sensitively, informative

(Archer 2009, 4).

2.3 Method

The procedure employed follows Rayson's Type III above, although taking Mahlberg's "literary insights" as a starting point. In this case, observations on the late plays by

a number of critics (cited in section 1.2 above) form the point of departure for this research. These critics perceive in the language of these plays “obscurity” (Sutherland 1959, 150), “verbal complexity” (Warren 1998, 21), “linguistic opacity” (Orgel 1996, 9) “tangled speech” (Pafford 1963, xxiii), “excessive obscurities” (Kermode 2000, 21), and “[t]he increased complexity of the style” (McDonald, 2006, 4); the language is “syntactically and lexically often baffling” (Orgel 1996, 7), “elliptical” (Mason Vaughan and Vaughan 1999, 20-21) and “enigmatic” (Barton 1968, 12), and the “syntax becomes convoluted” (McDonald 2006, 33). There is a sufficient degree of critical consensus that linguistic complexity is a feature of these plays to make that idea worth testing and exploring, and a WordSmith Tools keyword analysis presents an effective method of identifying salient linguistic features in the plays. Clearly there are numerous different ways in which language can be complex, and a keyword analysis can only identify saliency on the basis of frequency, so there is no pretension to inclusiveness in this approach. But it seems feasible that if syntactic patterns are repeated sufficiently often, they are likely to manifest evidence in a keyword list. The next step was therefore to build a corpus of the texts under scrutiny. It was decided that recent critical consensus tends towards the inclusion of the two last collaborative plays¹⁹, so the texts chosen to comprise the focus corpus were *Pericles*, *The Winter’s Tale*, *The Tempest*, *Cymbeline*, *Henry VIII* and *The Two Noble Kinsmen*. At this point, however, certain difficulties needed to be addressed.

Although most texts present some challenges for editors, there are problems facing the would-be analyst of Shakespeare’s text that simply do not exist for someone approaching Dickens, Conrad or Austen:

¹⁹ See 1.1 above.

Shakespeare wrote his plays in the first instance to be performed in the theatre. Though he might have anticipated that some of them would be published, there is little evidence that he was actively concerned with their appearance in print. During Shakespeare's lifetime there was no collected edition of his works. They were issued typically two or more years after they had been written and performed, and the process leading from author's pen to printed book is highly variable

(Jowett 2007, 4).

The outcome of this variability is what Honigmann aptly calls "the unknowable in Shakespeare's text" (1965, 151), numerous textual emendations and differences that cannot be reliably attributed to a single source, let alone to Shakespeare himself, and the fact that some of the plays do not exist in a single version. Instead, different texts compete for authenticity, and each time a new editor approaches them, decisions have to be made (often quite subjectively) about what to include or cut. *Hamlet* for example has three main texts, all with some claims to authenticity, yet all with shortcomings (see for example Thompson and Taylor 2006, 74-94). *King Lear* was first printed in the typical single edition quarto format, yet seemingly revised extensively later (probably in 1609) into the form in which it appears in the First Folio of 1623 (see for example Orgel 2000, and Foakes 1997, 110-133); modern editions with which audiences and readers are familiar are usually confections of the two. The First Folio is a relatively carefully edited, first complete works of Shakespeare, announcing itself to be "a defining, complete and authentic text" (Jowett 2007, 69). A number of the plays (including *The Tempest*, *The Winter's Tale* and *Cymbeline*) only exist in that volume (see e.g. Jowett 2007, 75), which can make editorial matters somewhat less challenging. But in a context where "the modern concept of literary authorship had relatively little purchase and collaboration was common" (Jowett 2007, 25), and scribes and compositors unhesitatingly altered or added to texts (e.g. Jowett 2007 57-61, and 1983), it is unsurprising that even the

First Folio has inconsistencies and errors. Furthermore, it does not include *Pericles* or *The Two Noble Kinsmen* at all, and both are known to be collaborations, *Pericles* with George Wilkins (see Vickers 2002, 291-332), and *The Two Noble Kinsmen*, with John Fletcher (Vickers 2002, 402-432). *Henry VIII* did, however, make it into the First Folio, despite also being a collaboration with John Fletcher. On top of these considerations, there is only one surviving textual source for *Pericles*, a notoriously corrupt Quarto of 1609 (see for example Warren 2003, 71-80, and Vickers 2002, as above), which has resulted in the fact that an unusually high number of editorial variations exist between currently extant texts. As McDonald explains, “editors lack an alternative text that might help them clarify or emend a difficult or manifestly corrupt passage” (2006, 4). The current editions in both Wells and Taylor (2005), and to a large extent the Oxford single edition (Warren 2003), represent a “reconstructed text” (e.g. Wells and Taylor 2005, 1059) which draws heavily on Wilkins’s novel *The Painful Adventures of Pericles, Prince of Tyre* (written in the 1570s) to replace sections of the Quarto text perceived to be faulty. On the other hand, the current Arden editor, Suzanne Gossett (2004), retains the Quarto version of the text on almost all the occasions where it is rejected in the Oxford.

Another challenge for the potential corpus analyst of Shakespeare’s texts is the problem of variable spelling. Since lexical frequency software identifies items on the basis of their form, it would construe Early Modern variations such as *fair* and *faire* as different words. The plays needed therefore to be obtained in a modernised spelling version. A request to Oxford University Press for permission to use their digitalised Oxford Complete Works (Wells and Taylor 2005), was unfortunately declined, and other recent editions were also unavailable

because still in copyright. As a result, the W.J. Craig Oxford edition of 1914²⁰ was finally chosen, and unless otherwise stated all textual citations in this research are from that text. This is a scholarly edition, something that cannot always be said in the case of online texts, and it has been the choice of other corpus analysts such as Culpeper (2009, 32). Craig's editorial preface (1894, v-vi) declares a conservative approach to textual emendation:

But in the absence of Shakespeare's manuscripts, the seventeen early quartos and the folio of 1623 jointly present, despite defect of copyist and printer, the sole authorized version of the Shakespeare text. From that version I have only ventured to deviate where it seemed to me that the carelessness of either copyist or printer deprived a word or sentence wholly of meaning. Editors of Shakespeare have sometimes denounced as corrupt and have partially altered passages which owe their difficulty of interpretation to the presence of some word or phrase rare in Shakespeare's day and long since obsolete. It has been my endeavour to avoid this danger. I have only adopted a change after convincing myself that the characteristics of Shakespeare's vocabulary or literary style failed to justify the original reading.

(Craig 1894, v).

From a modern editorial perspective, its shortcomings are its more frequent scene breaks and heavily prescriptive stage directions (every scene begins with a precise location, such as "A room in the palace", for example) and its rather overgenerous use of commas and semi-colons, the latter resulting from the fact that "[t]he punctuation has been thoroughly revised" (1894, vi). It includes a version of *Pericles* where the Quarto text is faithfully adopted, and is therefore very close indeed to the recent Arden edition (Gossett 2004). It was considered acceptable on the basis that a modern editor had also made these choices concerning *Pericles*²¹. Craig's stage directions and scene breaks did not present a problem, since (as explained below) all non-speech material was removed from the text before analysis. One small shortcoming in the electronic version is its rather problematic line numbering of

²⁰ Etext from www.bartleby.com. See Bibliography.

²¹ But see comments about the syntax of the two versions, in section 6.2 below.

prose sections (not the fault of the original editor, of course, but probably the result of the electronic format), which means that some compromises have to be accepted in the accuracy of line references.

The Craig edition does not, however, include *The Two Noble Kinsmen*, so the online version of C.F Tucker Brooke's 1908 edition of *The Shakespeare Apocrypha*²² was used instead. This is an original spelling version, so it was modernised manually according to the text of the Oxford *Complete Works* (Wells and Taylor 2005). Citations in this thesis from *The Two Noble Kinsmen* are from that Oxford edition.

Finally, it seemed important to ensure that only Shakespearean text was analysed, since this enquiry relates to the development of individual style. Three of the plays in question are collaborations, so it was necessary to decide on which sections of them to remove. Vickers's (2002) comprehensive account of Shakespeare's collaborations is generally recognised as authoritative, so all the whole scenes were removed that are deemed by Vickers to be written by other authors. (Appendix 1 shows a list of all scenes removed).

All non-speech material (stage directions, speech prefixes and act, scene and line numbering) was then removed from the texts. From the resulting corpus of late plays a wordlist was produced, showing the raw and relative (percentage) frequencies of all the word types in the corpus. Rayson's step 2²³ was bypassed, since no annotation was used in this analysis, and step 3 was then undertaken. This involved the creation of a reference corpus for comparison with the late plays, in this case the previous thirty-two. Craig's edition was again used, and non-Shakespearean material removed where appropriate, from *Timon of Athens* and

²² Available online from Project Gutenberg: see Bibliography.

²³ See section 2.2, page 40 above.

Titus Andronicus.²⁴ The retrieval of a keyword list was now possible. On the basis of the keyword results a set of procedures and categories were decided upon for the manual analysis, using both inductive and deductive decision-making, as will be described in more detail in the next chapter.

The textual variations that occur in all Shakespeare plays, no matter which edition is used, mean that there is no completely authoritative text, but studies of Shakespeare's language always face these or comparable problems, as McDonald acknowledges:

While it would of course be desirable to have better texts, versions reflecting greater fidelity to the words that Shakespeare wrote or that the King's Men performed – especially in the case of *Pericles* – the existing copies nevertheless offer thousands of comprehensible and more or less authoritative lines and sentences, plenty of territory for noticing poetic choices and linguistic properties

(McDonald 2006, 5).

Decisions with an element of subjectivity have nonetheless been taken, even at this early stage, concerning the intended focus of the enquiry on linguistic complexity, although these are minimised by employing Rayson's Type III procedure, which opens the field to any data revealed by the keyword analysis. The confident "rhetoric of objectivity" (Attridge 1987, 46) would therefore be misplaced in this study, but with Toolan's advice firmly in mind, it is hoped that it will be enough that assiduous efforts have been made to maximise the exactitude of the research, and at all times to clarify as far as possible the choices that have been made.

²⁴ Again, see Appendix 1 for details.

CHAPTER 3

ANALYSING THE CORPORA

3.0 General comments

Much of the ensuing thesis concerns the presentation of data from the Shakespeare corpora. Given the fact that this research spans different disciplines, and readers may not be familiar with certain terms used, an account is now given of the key terms employed in the presentation of that data. Firstly, the term ‘frequency’ is used in corpus stylistics to mean either the number of actual occurrences of a feature in a corpus, or (if the figure is a percentage) the proportion of the corpus represented by occurrences of that feature (see for example Scott 2008, Help menu, and Hunston 2002, 3). In order to maximise clarity in this respect, the terms ‘raw frequency’ and ‘relative frequency’ will be used in this thesis to refer to actual and proportional occurrences respectively. On some occasions, when the term ‘relative’ is also being used close by to refer to a syntactic function, ‘relative frequency’ will be replaced by ‘proportion’ or ‘density’.

Counting numbers of words in a corpus requires the distinction between ‘running words’ or ‘tokens’ on the one hand, and ‘types’ on the other. An example will make this clear. In the previous paragraph the penultimate sentence (beginning “In order to maximise”) contains 29 tokens, or running words. However, the words “in” and “to” occur three times, and “this”, “frequency” and “and” occur twice, meaning that there are in fact only 22 types in

the sentence. It will be the practice in this thesis to show types in italics and textual quotations in quotation marks. Furthermore, it can also be necessary to distinguish words that occur in the corpus on the basis of their function in the text. As will be seen below, *which* is counted in these corpora not only as a type, but also in terms of the number of occurrences that function as relatives, as opposed to interrogatives, for example.

Although the method has been used fruitfully by some stylometric analysts, it is not appropriate to discuss the present findings in terms of their rate of occurrence per play, since the play texts are of such widely varying sizes. *Pericles* (which of course is missing the first two acts, because they are not thought to be by Shakespeare) has only 9,943 running words or tokens, whereas *Antony and Cleopatra* for example has 23,909, so stylistic changes cannot usefully be mapped that way. A rate of use is given sometimes, however, expressed as ‘1 in x’ occurrences. This figure is not as precise as the percentage proportion to which it equates, since it is rounded to the nearest whole number, but it is given where it is felt that it offers a useful clarification of the statistics.

The chi square test for statistical significance was used in this research. Chi square scores are calculated for each table or pair of figures²⁵. The chi square scores themselves are expressions of the level of significance of the differences between the data sets (the higher the number the greater the significance), but since they are always positive, distribution tables are then used to turn this into the level of risk involved in making a claim based on that data. These tables give the threshold chi square score below which the data is not significant, and then a figure for the level of statistical risk associated with that score. The range of scores for

²⁵ The chi square calculations for tables were performed with reference to Oakes 1998, pp.24-29. For individual pairs of figures the UCREL log likelihood test was used (<http://ucrel.lancs.ac.uk/llwizard.html>).

data that has been tested and found to be significant is 0.02 to 0.001, with 0.001 the most significant. Statistical significance test results are shown at the foot of tables, where appropriate, and discussed in the accompanying sections. They are almost always expressed as levels of risk from the distribution tables, and where this is not the case explanation is given.

3.1 The prominence of *which* in the late plays

The corpus of six late plays, (with a total of 99,976 tokens, or running words) was analysed first for the raw frequencies of all its word types, and the resulting wordlist can be seen in Appendix 2. A similar wordlist (with a total of 701,588 tokens) was produced for the thirty-two earlier plays (also see Appendix 2), and this then became the reference corpus for the keywords analysis, as described in the last chapter. WordSmith Tools uses Ted Dunning's Log Likelihood test to establish whether a word frequency is salient or otherwise, a test which according to Mike Scott "gives a better estimate of keyness" than the chi-square test often used elsewhere (2008, Help Menu). For this analysis, a very high probability threshold was used, with a keyness p value of 0.000001, or a one in a million likelihood of the result being pure chance, meaning that the keyword list produced will only include those words that are very key indeed. The resulting list contains 107 words, with keyness values ranging from the most key word *Camillo* (1) at 170.71 to *its* (the 107th keyword) at 23.95. The top 49 keywords are shown in sequence in Table 3.1 below, and the full list can be found in Appendices 3 and 4.

Table 3.1. The late plays: top 49 keywords

1-7	8-14	15-21	22-28	29-35	36-42	43-49
Camillo	Pericles	o' th	Cromwell	oracle	Tarsus	has
Palamon	Marina	Ariel	Cloten	Sicilia	Caliban	island
Posthumus	Imogen	Leonatus	Milan	'em	queen's	Haven
Britain	monster	Hermione	Naples	Polixenes	business	Thebes
Pisanio	queen	Paulina	sea	in't	her	beyond
which	Arcite	Milford	Stephano	Prospero	highness	cardinal
Bohemia	gods	Trinculo	Helicanus	i'th	Cymbeline	Cleon

Firstly, not all the results are 'words' as such. Some are abbreviations, such as *'em*; others are combinations of abbreviated words, such as *i'th*, and others still are only morphemes, suffixes such as *t*. It is likely that considerable orthographical intervention by the scribe Ralph Crane occurred in the transcription of the manuscripts of *The Tempest* and *The Winter's Tale*, particularly with regard to apostrophes and elisions (Howard-Hill 1972, 43-47) and these forms are therefore discounted in this thesis. Concerning the rest of the results, Mike Scott suggests that there are three basic kinds of keyword:

First, there will be proper nouns. Proper nouns are often key in texts, though a text about racing could wrongly identify as key, names of horses which are quite incidental to the story... Second, there are key words that human beings would recognise... Third, there are high-frequency words like because or shall or already. These would not usually be identified by the reader as key. They may be key indicators more of style than of "aboutness". But the fact that KeyWords identifies such words should prompt

you to go back to the text ... to investigate why such words have cropped up with unusual frequencies.

(Scott 2008, Help Menu).

As can be seen from Table 3.1, the majority of the top keywords in this case (31 out of 49, or 63%) are proper nouns. Although proper nouns in keyword lists can sometimes be seen as significant (see e.g. Culpeper 2009, 38) where the concern is for features of linguistic style they have, as Scott declares above, minimal relevance. They appear here simply because they occur in the late plays and not in the earlier canon. Lexical words “that human beings [as opposed to computers] might recognise”, for example *monster* or *island* in this corpus, are related to the “aboutness” of the texts, of *The Tempest* in those cases. Again however, their high frequency only tells us that the late plays are more preoccupied with monsters, islands and so on than the earlier canon. What is potentially highly significant here, though, is the appearance of a grammatical word, *which*, with a very high keyness score of 102.03. Such “high frequency words” (Scott, above) do not usually occur in keyword lists at all, because they tend to be similarly common in most texts. As Scott indicates, when they do appear there is probably something of stylistic interest happening in the corpus under scrutiny, particularly in this case since as Hope points out, there is minimal evidence of secondary textual intervention in the case of Shakespeare’s relative pronouns (1994, 33). The salient use of a closed class grammatical word such as *which* may well be an indication of a sub-literary style feature, therefore.

Further reason for interest in this result comes from the fact that Witmore and Hope (2007) also find an increase in linguistic formations using *which* in the late plays. Their

computer software produces results in terms of predetermined functional and semantic categories, as they explain:

The Language Action Type known as *Asides* identifies strings that introduce or end digressive comments ... A significant marker of asides for the developers of Docuscope is ‘which’ as a non-restrictive relative pronoun – that is, a relative introducing information which is non-essential, hence an aside. Indeed, when Docuscope analyses Shakespeare, Asides strings are almost wholly made up of ‘which’ forms, and there is a significant increase in these strings over his career, and particularly in the late plays

(Witmore and Hope 2007, 146).

They admit some reservations about the software, saying that they “remain agnostic on the question of whether or not the architecture of categories used by Docuscope represents the ‘best’ or most ‘functional’ way to organize language” (140), and it would seem that a keyword analysis such as WordSmith Tools fosters greater transparency for the analyst, allowing more freedom of interpretative choice. However, their assertion that “if Docuscope were redesigned to count different things ... we would still expect it to find statistically significant patterns of string usage” (140) would seem to be vindicated by the keyword analysis performed here, since the salience of *which* appears once again.

Which keyness is also particularly interesting because its most typical function is to introduce a relative clause, or the element of a sentence that renders it complex (see for example Rissanen 1999, 280). Complex sentences involve a hierarchy of syntactic levels, where one clause is subordinate to, or dependent on, the other, and hypotaxis is universally seen as a more complex syntactic feature than parataxis (for example Dekeyser 1984, 33; Romaine 1982, 54). A further link to this keyword finding is Russ McDonald’s observation that in the late plays “a frequent source of syntactic difficulty is the poet’s additive impulse, usually his continuation of an apparently completed sentence by appending non-restrictive

relative clauses” (2006, 46). It is possible that some of the increased linguistic complexity identified by other literary critics in the late plays may be related to this high frequency of *which* occurrence.

The WordSmith Tools data reveals that in the reference corpus of the first thirty-two plays of the canon occurrences of *which* represent 0.25% of the total tokens (running words). In the late plays group this percentage rises to 0.44%, a very substantial 76% increase, and it is distributed throughout the corpus²⁶. In other words, in the earlier canon *which* occurs once in every 400 words, whereas in the late plays this rises to one in 227 words.

The interpretation of these figures requires several important issues to be considered. The first step towards ascertaining whether or not they represent an individual authorial stylistic shift is the analysis of the background linguistic phenomena. The Early Modern period was one of considerable linguistic flux, and there were changes occurring in the use of relatives²⁷ such as *which*, which need to be examined in detail before any claims can be made for Shakespeare’s individual syntactic style. We need to establish as far as possible what choices might have been available for a writer of the time in the selection of one grammatical word over another. Once such paradigmatic matters have been investigated, the next step will be to analyse in detail the syntagmatic phenomena in Shakespeare’s use of *which*, the words that occur alongside *which* in the plays, to ascertain whether there are any observable changes in his use of the word from the earlier canon to the late.

²⁶ Distribution can be determined from the concordance lines display for *which*, in the column to the right which details (as a percentage) the place in the text file (the individual play) where the item occurs.

²⁷ Several different terms are in use for words such as *which*, which introduce relative clauses. There is debate over the status of *that* as a pronoun (e.g. Ryden 1966, xlv), so the term ‘relative pronoun’ has been avoided here unless more specific reference is required. ‘Relative’ has been adopted since it has the advantage of brevity over ‘relativiser’ and ‘relative marker’, but it is intended as synonymous with these two.

3.2 Paradigmatic analysis

We are concerned here with the whole group of relatives²⁸ available to Shakespeare when opening a relative clause. They are not completely interchangeable, but as will be seen, *which* is the most flexible of the set, potentially usable instead of *who* (although not in genitive or accusative cases), and also instead of *that*. In order to gain a clearer picture of his *which* usage, it is necessary to develop some idea of the choices made in such instances, and this requires an understanding of the background linguistic picture, and the current practices governing the use of relatives at that time.

Charles Barber summarises the situation with respect to relatives in that period as follows: “The relative pronouns commonly used in [Early Modern English] are *who* (and its oblique forms), *which* and *that*” (1976, 213). Although these relatives are in use in Present Day English, the manner and distribution of their use in Early Modern English was different in a number of respects, and was in the process of changing substantially, as Xavier Dekeyser explains: “No new elements have been added to the set of relativisers since the 15th century. The crucial difference between ME [Middle English] and PE [Present Day English] is not the number of relatives, but the system that governs their distribution” (1984, 61). In order to understand the significance of the increase in *which* in the late plays corpus, we need to understand this system, and the ways in which it was changing during the late sixteenth and early seventeenth centuries.

Diachronically speaking, *which*, *whom* and *whose* appeared as relative pronouns from the 12th century onwards, with *who* first recorded in a strictly relative function much later, in

²⁸ All of them are exclusively relative pronouns, except *which*, which can also be a relative determiner. See 3.3.1 below.

the 15th century (e.g. Dekeyser 1984, 61). As Rissanen explains in the *Cambridge History of the English Language*, “[t]he distribution along the animacy parameter is established in the course of the seventeenth century” (1999, 294). Prior to the establishment of that parameter, and to varying degrees during the process of its consolidation, *which*, *whom* and *whose* could all be used freely for both animate and inanimate antecedents (Rissanen 1999, 294), so *who* was not needed.

Mustanoja’s (1960) widely accepted account of the development of relatives through Middle and into Early Modern English explains the establishment of *which*, describing *that* as “the ME relative par excellence” (1960, 190), with *which* having subsequently evolved as a relative pronoun via its functions as a generalising relative (with the meaning *whichever*), and as an interrogative pronoun (Mustanoja 191-194). It seems that *wh-* forms replaced *that* in cases where increased semantic clarity was needed, as Rissanen explains:

There is little doubt that the spread of the *wh-* forms was supported by the heavy functional load of *that*. When the connection between the antecedent and the relative link was loose, the likelihood of ambiguity and misunderstanding of the meaning of *that* increased. Consequently, the *wh-* forms seem to be first established in contexts of loose relative link

(Rissanen 1999, 295).

This tendency seems to be related to “[t]he [Early Modern] practice to limit the use of *that* to defining (restrictive) relative clauses” (Mustanoja 1960, 197), since in restrictive clauses there is a tighter relationship between antecedent and relative than in non-restrictives. If we are interested in features of syntactic complexity in the late plays corpus, then, it is potentially significant that *wh-* forms are associated with clauses where there is a “loose relative link”, since these may well appear more complex to a reader or auditor.

Further evidence of this understanding of *which* is provided by Keenan and Comrie (1977), whose work has given rise to the association of *wh-* forms with syntactic complexity and abstraction (see for example Romaine 1980, 228 and Dekeyser and Ingels 1988, 28 and 35-36). Keenan and Comrie demonstrated in present day languages a universal hierarchy of ease of psychological processing of different restrictive relative clauses, dependent on the function of the relative in the clause, ranging from subject position (the easiest) through object and genitive positions, to comparative (the most difficult to process). This “Case Hierarchy” (sometimes known as the “Accessibility Hierarchy”) has been used by historical linguists (such as Romaine, 1980, 1982, and Dekeyser 1984) to describe the development of *wh-* relatives in Middle and Early Modern English. Romaine (1980, e.g. 221-226) shows how *wh-* forms took hold in the more complex functional positions in the relative clause (such as genitive and oblique) and only gradually became frequent in subject position, the least complex position. Dekeyser also demonstrates the connection between choice of relative and linguistic mode, revealing that *wh-* forms were connected strongly with written (rather than written-to-be-spoken) genres, and with the more formal registers such as informative scientific or legal prose (e.g. Dekeyser and Ingels 1988, 35-6).

In addition to *that* and the *wh-* forms, there was in Early Modern English, as now, the possibility for the relative clause “to be appended to the main clause by zero, i.e. without an expressed relativiser” (Rissanen 1999, 298). Rissanen gives an example of this kind of relativisation: “he hathe bene otherwise enformed of *them he put in trust*” (1999, 299). This form was confined to restrictive relative clauses, and occurred most often in object position. The use of relative pronoun deletion was (and is) more closely circumscribed than that of *which*, or even *that*, since the connection between antecedent and relative clause needs to be

tight and unambiguous in such cases. Typically “the [zero] relative clause immediately follows the antecedent...It is short, and has a personal pronoun subject” (Rissanen 1999, 298).

To summarise, the process overall in the sixteenth century was one where *wh-* forms were gradually replacing *that* as relatives in a variety of different syntactic situations, particularly in formal and literary contexts, and particularly where the semantic connection between the foregoing antecedent clause and the relative was loose or potentially unclear. Although, as suggested above, there is an interesting connection appearing between *which* usage and syntactic complexity or opacity, this information could lead to an interpretation of the Shakespeare data as simply reflecting the background trends in the language at the time, but as will be shown, that is not the case.

Before proceeding any further, it will be useful to clarify the different steps involved here in analysing the paradigmatic relationships of *which*.

- A. Raw and relative frequencies of the word types (*which*, *who* etc) are obtained. This is the information that WordSmith Tools retrieves from the corpus. The computer does not distinguish on the basis of function, only of the form of the words, but it is a useful starting point. In this case, as we saw above, it was the discovery of a very marked increase in relative frequency of the type *which* in the late plays corpus (expressed as high-ranking keyword status) that prompted further investigation.
- B. We ascertain the proportion of the word types in A that are relatives. In other words, the data is now analysed in terms of the function of the words. This separates off the non-relatives (such as interrogatives) from the relatives, and is the necessary precursor to C below.

C. Having (in B) established which of the words have relative function, the body of relatives is then counted, and analysed in terms of its internal proportions. This reveals the writer's paradigmatic choices.

Dekeyser's (1984) survey of relatives in modern English, represented in tabular form and reproduced below, approaches the data from the point of view explained in C above. He has identified all the words with relative function in his corpora, and then offers a diachronic overview of the paradigmatic usage of relatives between 1520 and 1649. The original table also included data for Present Day English, excluded because not relevant here. There are certain limitations to Dekeyser's data, which should be taken into account. The first sample (1520-1560) comes from Mats Ryden's (1966) study of Sir Thomas Elyot and his contemporaries (largely non-fictional, discursive prose). The second (1600-1649) is Dekeyser's own collection of texts, comprising non-fictional prose, dramatic works, formal epistolary prose, and poetry. As already suggested, variation in choice of relative can occur according to the register and form of the text in question (see also for example Rissanen 1981, 417; Romaine 1982, 69-71), so like should ideally be compared with like. There is also a gap of 40 years between the samples, and a considerable difference in their sizes. However, it is asserted confidently by Dekeyser as a representative overview (1984, 62), and Mats Ryden says of his own (early sixteenth century) samples: "The corpus of recorded data ... is deemed sufficient to give a survey of the use of the relatives in the literary language of the time" (1966, xx), so these accounts were accepted as suitable sources of comparative data for the present study.

Figures showing the percentage increase or decrease in the table have been added, for clarity. We see here the full set of relatives, with *who* and its inflections grouped together,

along with the so-called ‘zero’ form. Dekeyser’s inclusion of *the which* parenthetically in the title of his *which* column simply reflects one of the common collocation patterns found at the time. Norman Blake indicates that “[*t*]he *which* may be used instead of simple *which* in all

Table 3.2 “Overall survey of relativisers in Modern English” (Dekeyser 1984, 64)

	<i>Who(m)/se</i>	<i>(The)Which</i>	<i>That</i>	<i>Zero</i>	<i>Total</i>
1520-1560	2,442 (8.09%)	13,253 (43.90%)	14,038 (46.50%)	456 (1.51%)	30,189 (100%)
1600-1649	588 (17.43%) (+9.34%)	1,191 (35.31%) (-8.59%)	1,152 (34.15%) (-12.35%)	442 (13.11%) (+11.60%)	3,373 (100%)

grammatical positions” (2002, 64), and experience of the Shakespeare corpora in this research accords with that view. The addition of *the* does not seem to affect the clause either syntactically or semantically. Dekeyser summarises his figures as follows: “This table reveals the expansion of WHO and a sizeable increase of pronoun deletion in [Early Modern English], while both (THE) WHICH and THAT are on the decrease” (1984, 65). He continues: “The total share of WH- is remarkably steady ... the decline of the longstanding THAT strategy is entirely counterbalanced by an almost identical rise of [zero]” (65). This study suggests that the spread of *wh-* forms in the first half of the seventeenth century is confined largely to the replacement of *which* by *who* as the animacy parameter is established more firmly. This accounts for the decrease in *which*, while *that* occurrence is reduced in correspondence with the increase in zeros.

These findings suggest the possibility of a difference between Shakespeare's *which* usage and the background linguistic trends, but as explained above, Dekeyser only looks at relatives, and the keyword results for Shakespeare discussed above do not distinguish between the different functions of *which*, so will include some that are not relatives. In order to make the Shakespeare corpora comparable with Dekeyser's data, then, it was necessary first to proceed from A (the WordSmith data) to B, namely to identify and count the occurrences of all these word types in the canon (not only *which*) that were relatives, as opposed to interrogatives, complementisers or demonstratives, for example. WordSmith Tools was used to produce concordance lists of all the Shakespeare occurrences of *that*, *who*, *whom* and *whose* from both the corpus of the first thirty-two plays and that of the six late plays²⁹. All of these were then analysed manually. *That's* and *who's* were also examined, since they appear as separate types in the WordSmith Tools data, and *that's* results were simply added to those for *that*. In the case of *who's* there were no occurrences of that form as relatives: all were interrogatives. The occurrences of *who's* were not therefore included in the totals recorded.

More detail is needed at this point concerning the distinctions made between different functions of 'that' in the canon, since this word has a broad and sometimes complex range of functions in Shakespeare. The following occurrences are examples of clauses where *that* is not clearly either a relative pronoun or a conjunction:

- (1) And in the instant *that I met with you*³⁰
He had of me a chain

(*The Comedy of Errors*, 4.1.11-12)

²⁹ See Appendix 6.

³⁰ Throughout this thesis, italicisation of words in textual quotations from Shakespeare is used to highlight the analytical point in question.

- (2) Now it is the time of night *that the graves all gaping wide...*

(A Midsummer Night's Dream, 5.1.365)

In temporal clauses such as these, *that* seems to be functioning semantically as *when*. It does not function nominally in the relative clause, either as subject or object. However, because in all such cases it has a clear nominal antecedent (“instant” and “time of night” in these occurrences), and it seems to herald a relative clause in the same way as a more typical relative pronoun, *that* occurrences like these have been counted as relative pronouns. Discretion was also employed in the categorisation of oddities such as these:

- (3) ...there's no man is so vain
That would refuse so fair an offer'd chain.

(The Comedy of Errors, 3.2.148-9)

- (4) And sail so expeditious, *that shall catch*
Your royal fleet far off.

(The Tempest, 5.1.351-352)

In cases such as this, “that” seems to be what Quirk (1985, 999) calls a correlative subordinator, along with “so”, but that construction would normally require a pronoun following it: in these cases *he* and *it* respectively. Since Shakespeare has omitted the personal pronoun (possibly for metrical reasons), the effect of “that” becomes more like that of a relative pronoun, with “that” as the subject of a relative clause, and “man” and “sail” as antecedents. It was therefore counted as such on these occasions.

Finally, before comparing the statistics for Shakespeare's relatives with those of Dekeyser's corpus, clarification is needed concerning pronoun deletion, or zeros. Although this is a relativisation strategy, it clearly differs from other relative forms in that it does not have a word type as a relativiser, and is therefore not a suitable candidate for word frequency

analysis such as is used in this study, which means that the Shakespeare data does not include zeros. This issue is discussed at more length in Chapter 6³¹. In order to make possible direct comparison of Dekeyser's statistics with those for Shakespeare, Dekeyser's data is reproduced here again, but this time with the results for zeros removed, and the remaining totals recalculated. The layout of the table is also changed now, to correspond with the one used elsewhere in this thesis.

Table 3.3 Dekeyser's survey of relatives (1984, 64) without zeros. (Raw frequency and proportion)

	1520-1560	1600-1649
<i>Who/m/se</i>	2,442 (8.21%)	588 (20.06%)
<i>(The) which</i>	13,253 (44.57%)	1,191 (40.63%)
<i>That</i>	14,038 (47.21%)	1,152 (39.30%)
Total	29,733 (100%)	2,931 (100%)

³¹ See page 278.

Table 3.4 Raw frequency and proportion of relatives in Shakespeare

	First 32 plays	Late plays
<i>Who/m/se</i>	1,129 (18.05%)	325 (24.79%)
<i>(The) which</i>	1,631 (26.07%)	431 (32.87%)
<i>That</i>	3,495 (55.87%)	555 (42.33%)
Total	6,255 (100%)	1,311 (100%)

Statistical significance: risk level 0.001.

As before (in Table 3.2) Dekeyser's data shows the substantial increase (8.21% to 20.06%) in relative frequency of use of *who/m/se*, corresponding with the increase in *who* usage (although the three forms are not distinguished statistically). In the Shakespeare corpora this increase is less pronounced, because the proportion of *who/m/se* in the earlier canon is much higher (18.05%), probably resulting from the fact that even the first 32 Shakespeare plays were written at the end of the 16th century, from the 1590s onwards, when the animacy parameter would have been more fully established than in the first half of the sixteenth century. But in the Dekeyser data it is only *who/m/se* that increases its proportion of all relatives. In the case of Shakespeare, *which* also increases its share, going from 26.07% to 32.87%. Dekeyser's picture of Early Modern relativisation does not provide any explanation for this increase in *which*: there, the trend is in the opposite direction. It seems likely then that preference for *which* is an individual stylistic feature of Shakespeare's late plays.

Focusing now on just the two Shakespeare corpora, we can gain a fuller picture of the diachronic changes between the two by looking at the data described as step B above³². Firstly

³² See p. 59.

we can find out what proportions of all running words (tokens) in the corpora are represented by relative *who/m/se*, *which* and *that* together. These figures will help us to see whether relativisation itself increases (proportionally) or not in the corpora.

Table 3.5 Proportion of running tokens that are relatives in Shakespeare

	First 32 plays	Late plays
Total tokens in corpus	701,588	99,976
Total relatives (% of tokens; rate per 1,000 words)	6,255 (0.89%; 8.9)	1,311 (1.31%; 13.1)

Statistical significance: risk level 0.001.

Table 3.5 reveals an increase of 0.42% in the relative frequency of relatives in the late plays³³. In other words, the density of relatives in the late plays is almost half as great again as in the earlier plays, with one relative used for every 76 words (instead of for every 112 words in the earlier corpus).

Further results that develop this picture can be seen in Tables 3.6 and 3.7. In these we have data for steps A and B respectively, for the two Shakespeare corpora, shown for each word. Table 3.6 shows the number (raw frequency) of each word type found in each corpus, and (in parentheses) the proportion of all the running words (tokens) in the corpus that this represents. The rate of occurrence is also given here, for clarity. The most important feature to notice in Table 3.6 is the fact that all of the word types except *that* increase in relative frequency and rate in the late plays, and even *that* does not decrease, but stays the same, at 1.36% of all tokens in the corpus. This is another, confirmatory perspective on the increase in overall relativisation in the late plays, which we saw in Table 3.5. What this table also shows

³³ Full details of the paradigmatic statistics can be found in Appendices 5 and 6.

Table 3.6 Shakespeare's relatives: raw and relative frequencies of types

	Raw (and relative) frequency in first 32 plays	Rate of occurrence per 1,000 words	Raw (and relative) frequency in late plays	Rate of occurrence per 1,000 words	Statistical significance of change (Chi square score³⁴)
<i>which</i>	1,763 (0.25%)	2.5	444 (0.44%)	4.4	102.25
<i>who</i>	961 (0.14%)	1.4	199 (0.20%)	2	21.19
<i>whom</i>	351 (0.05%)	0.5	89 (0.9%)	0.9	20.94
<i>whose</i>	494 (0.07%)	0.7	99 (0.10%)	1	8.88
<i>that</i>	9,530 (1.36%)	13.6	1,365 (1.36%)	13.6	0.03

is that all of the first three *wh-* forms (*which*, *who* and *whom*) have statistically significant increases in the relative frequency of their types. All have increases in the late plays with the minimum 0.001 statistical risk of their being chance, since the threshold for significance in these cases is 10.83. But the chi square scores themselves are given here, in order to demonstrate the very marked prominence of the significance level of the increase in proportion of *which* in the late plays. The *which* increase is five times as statistically significant as that for either *who* or *whom*.

³⁴ See p. 50.

Table 3.7 Shakespeare's relatives: proportion of each type that function as relatives

	Number of type in first 32 plays that are relatives (proportion of type)	Number of type in late plays that are relatives (proportion of type)	Total in canon
<i>Which</i>	1,631 (92.51%)	431 (97.07%)	2,062
<i>Who</i>	385 (40.06%)	139 (69.84%)	524
<i>Whom</i>	287 (81.77%)	87 (97.75%)	374
<i>Whose</i>	457 (92.51%)	99 (100%)	576
<i>That</i>	3,495 (36.67%)	555 (40.66%)	4,050
Total	6,255	1,311	7,566

Statistical significance: risk level 0.001

We can see from this table that in the Shakespeare canon three out of five of the word types are very largely used as relatives. Only *that* and *who* have a substantial proportion of occurrences with different functions. And of course we see again the increase in *who* functioning as a relative in the later corpus, corresponding to the background increase in its use with animate antecedents. But most importantly in this table we see that the proportion of types that are relatives increases in all cases in the late plays. Not only are the types increasing their share of the running tokens in the corpus (Table 3.6), but the proportion of those types that are relatives increases. We now see that even *that* actually increases its density of occurrence as a relative in the late plays, since although its share of tokens in the late corpus remains the same (at 1.36%³⁵), the proportion of those that are relatives increase.

³⁵ See Table 3.6.

So these results reveal some important phenomena occurring in Shakespeare’s syntactic style over the period. Firstly, Shakespeare uses relative clauses proportionally more frequently in his late plays. As we have seen above, frequency of relativisation is directly linked to syntactic complexity by both Romaine (1984, 446) and Dekeyser (1984, 33). But more particularly, there is a highly significant increase in his choice of *which* to herald these clauses.

Four years after the account cited above, Dekeyser published another study with Mia Ingels, in which they fill the diachronic gap in the earlier data with information derived from a sample created from texts written between 1550 and 1600. Their main aim is to extend work by Romaine (1982), which examines Early Modern relatives simultaneously from a functional and historical perspective, focusing on paradigmatic choices linked to textual register (and to Keenan and Comrie’s (1977) Case Hierarchy). Their sample comprises approximately 48,000 tokens of “informative prose”, 36,000 of comic drama, and 36,000 of tragic drama (1988, 26). They present comparative data showing the distribution of *wh-* forms and *that* in their sample, reproduced below.

Table 3.8 “Relativization correlated with register” (Dekeyser and Ingels 1988, 31)

	<i>Wh-</i>	<i>That</i>	<i>Totals</i>
Informative Prose	562 (69%)	252 (31%)	814 (100%)
Comedy and Tragedy	286 (47.7%)	313 (52.3%)	599 (100%)
Totals	848 (60%)	565 (40%)	1,413

The comparable statistics from the two Shakespeare corpora are presented here, as in the tables above but with the statistics for the *wh*- forms conflated, for ease of comparison with Dekeyser and Ingels’s results.

Table 3.9 *Wh*- forms and *that* in Shakespeare’s earlier and late plays

	<i>Wh</i>-	<i>That</i>	<i>Totals</i>
First 32 (c.1590- c.1606)	2,760 (44.12%)	3,495 (55.87%)	6,255 (100%)
Late plays (c.1607- 1613)	756 (57.66%)	555 (42.33%)	1,311 (100%)
Totals	3,516	4,050	7,566

Statistical significance: risk level 0.001

It can be seen from a comparison of the above two tables that the statistics for Shakespeare’s earlier canon correspond fairly closely to the Comedy and Tragedy figures in the 1550-1600 sample. But instead of the *wh*- proportion staying “remarkably steady” into the seventeenth century (Dekeyser 1984³⁶), in the late Shakespeare plays the proportion of *wh*- forms increases substantially, from 44.8% to 58.6%, half way to the 69% figure found in Dekeyser and Ingels’ “informative prose” category. And, as in that category, in Shakespeare’s late plays *wh*- forms actually overtake *that* and become the more frequently occurring form.

Dekeyser and Ingels describe their informal prose category as follows:

We can safely assign what we have labelled [informative prose] to the register of written scientific English, and as it addresses itself to a scholarly audience, it most probably has the characteristics of a formal register.

(1988, 27)

³⁶ See p. 61.

I argue therefore that we are seeing in Shakespeare's late plays a movement towards a syntactic complexity and semantic abstraction (see Romaine 1982, 228) associated with the more formal, literary style of Early Modern prose than drama.

3.3 Syntagmatic analysis

Having considered the marked increase in relative frequency of *which* in the late plays, and the paradigmatic factors associated with its prominence, we now move on to a different analysis. Here we will not be concerned with *which* frequency or density in the plays: that will be the backdrop to the investigation, rather than its focus. Now what is at stake is a detailed analysis of the co-text surrounding *which*, in an attempt to describe Shakespeare's changing use of the relative, particularly with reference to syntactic complexity.

Shakespeare's syntax is rich and highly variable in its forms, however, and in order to facilitate a successful account even of the use of one word, an extensive analytical map is required, to chart the categories to be employed and the criteria for their use, before counting can begin. This is also a necessary provision if the research is to be "detailed, explicit and retrievable"³⁷ (Carter and Simpson (1989, 14). The rest of this chapter therefore presents such a map of Shakespeare's *which* relativisation, offers a detailed articulation of the categories employed, and discusses the principal variants and definition problems found in each case. (Appendix 7 shows complete lists of the occurrences of *which* in the earlier thirty-two and the late plays respectively, in their concordance format).

³⁷ See 2.1, page 36.

As mentioned in 3.2 above, WordSmith Tools makes possible this kind of detailed analysis through its Concord function, whereby all the occurrences of a selected type from the wordlist (or keyword list in this case) are presented as a set of concordance lines, facilitating the analysis of each occurrence in its textual context. The software also allows this to be expanded to whatever extent the researcher requires, from the fifteen or so words in the concordance window, to the whole text. This wider reference is of course essential in a study such as the present one, not only in order to see all the linguistic material actually under scrutiny on either side of *which*, but at times to provide a more extensive context, in cases where analytical decisions hinge on fine semantic nuances perhaps conveyed some way back in a speech, or by an earlier speaker in the drama.

All 1,763 occurrences of *which* in the earlier plays, and the 444 in the late group, were analysed manually according to a set of categories established for the purpose, and then counted manually, using a hand-held digital counter in order to minimize error. The process by which these categories were arrived at was a combination of deductive and inductive approaches. Existing commentaries on Early Modern syntax were consulted, (referred to where relevant in each case here), from which a list of potentially important features was compiled. In addition, some analytical criteria evolved inductively, with some key features of the clauses becoming apparent only during the analysis itself, making necessary reanalysis of earlier concordance lines and some revision of categories. This process meant that the research was considerably prolonged, but should have the advantage of producing more thorough and rigorous results.

It is important to note that Shakespeare's syntax is extremely variable, and in some respects does not correspond with syntactic categories in use today. The fine distinctions made

here are not intended to describe a system that he was (consciously or unconsciously) employing. They are rather a means of describing and mapping the linguistic phenomena in question, sometimes according to present day grammatical frameworks, so that they can be counted. In order to make possible the analysis of all the *which* occurrences in the canon, descriptive categories have been used which did not exist in the period in which Shakespeare was writing. The distinction between restrictive and non-restrictive clauses, for example, was not identified analytically until “much later” (Adamson 1999, 586). In all the following cases, however, the categories utilised (either according to existing linguistic practice, or created for the purpose of this analysis) are described in detail, exemplified, and explained in terms which aim to facilitate the replication of the analysis. Where subtle or finely-drawn distinctions are at stake, examples are offered which demonstrate the borderlines. Diagrams are included at the end of each main section, to clarify the relationships between categories and subcategories. Where these relationships are particularly complex (for example in the case of the role of *which* in the relative clause) an explanation is provided in addition to the diagram. In all sections of the analysis, where issues arise that pertain particularly to problems with individual cases, or anomalies in the counting and categorisation procedures (rather than to the definition of a type of formation), they are dealt with in Chapter 4.

Reviewing a doctoral dissertation on relative clauses, Mats Ryden observes that “[i]t is one thing to characterise and label type examples, real or invented. Another and much more difficult task is to identify various specimens as they occur in actual texts” (1983, 212). This is certainly the case, especially with a writer with the sophistication and linguistic richness of Shakespeare. However, one of the main purposes of the present thesis is to offer a systematic

and comprehensive, rather than merely selective and exemplifying, analysis of Shakespeare's use of *which* as a relative, so these difficulties had to be faced.

3.3.1 Main functions of *which*

The first distinction that needs to be made in a body of *which* occurrences is between interrogatives and relatives (either relative determiners or relative pronouns). Interrogatives, such as the following:

- (5) *Which* way is he, in the name of sanctity?

(*Twelfth Night* 3.4.50)

- (6) Prithee, Apemantus, read me the superscription of these letters: I know not *which* is *which*.

(*Timon of Athens*, 2,2,86)

have a very different, and arguably simpler, syntactic function from relatives. They were counted in this study, but not analysed further. The relative determiner³⁸, on the other hand, introduces a relative clause in a similar way to the relative pronoun, and so, given the interest in this thesis in syntactic complexity possibly created by the relative clause, it requires the full analysis afforded to the pronoun category. Rissanen describes it as follows: "The determiner *which* is popular in Late Middle English and Early Modern English. It always introduces non-restrictive – often continuative – clauses, mainly with non-personal antecedents" (1999, 296). Examples from the Shakespeare corpus include:

³⁸ "Relative determiner" is the term used by Rissanen (1999, e.g. 296) and Nevalainen (2006, e.g.86) and it has been adopted here, although the same function is referred to by Reuter (1938, e.g. 3) as "adjectival *which*", and Romaine (1980, 227) refers to the phenomenon in terms of the NP following *which*, as "the shared nominal or co-referential NP [which] is not deleted".

- (7) Sluic'd out his innocent soul through streams of blood;
Which blood, like sacrificing Abel's cries...

(*King Richard II*, 1.1.106-7)

- (8) And when thou fail'st – as God forbid the hour!—
Must Edward fall, *which* peril heaven forfend.

(*Henry VI Part 3*, 2.1.194-195)

Relative determiner *which* is characterised by being immediately followed by a noun, which either reiterates or paraphrases the antecedent of the relative clause, as can be seen from these examples, respectively. In the latter case the noun *peril* probably contributes something to the semantic clarity of the relative clause. In the former, however, as Reuter comments on a similar construction, “the [noun is] not added for the sake of clearness, but to serve a rhetorical end” (1938, 3). There are occasional cases of ambiguity involving determiners in the Shakespeare corpora. In the following example from *All's Well That Ends Well*, the noun “gratitude” cannot be said to be paraphrasing the antecedent of the relative clause, yet “which” here clearly is a relative determiner:

- (9) Time was, I did him a desired office,
Dear almost as his life; *which gratitude*
Through flinty Tartar's bosom would peep forth
And answer thanks.

(4.4. 7-10)

As seen in the first examples cited above, and as will be seen on later occasions in this analysis, Shakespeare is here prioritising metrical concerns over syntactic or even semantic clarity, and gives us “which gratitude” instead of perhaps “gratitude for which”. As Jonathan Bate recognises, he “cared more for the actor's tongue than grammatical nicety” (1995, 176 n.). Cases such as this exemplify “the additional complications arising from the conditions of

poetry and drama” (McDonald 2006, 108) which the analyst of Shakespeare’s language must deal with.

3.3.2 Antecedents

3.3.2.1 Types of relative clause according to antecedent

The first step in the analysis of *which* clauses must be to distinguish them on the basis of the syntagmatic relationships to the left of the relative, and first in this must be an examination of their antecedents. In his detailed study of Early Modern relatives, Mats Ryden identifies “the nature of the antecedent” as one of the principal “conditioning factors” which have “a particularly decisive distributional effect” on the occurrence of relatives (1966, xx).

Rissanen’s outline is useful here:

Relative clauses can be divided into adnominal, nominal and sentential, with reference to the type of their antecedents. The most common are the adnominal clauses, which have a (pro)noun as the antecedent ... In nominal clauses the relative pronoun ‘contains’ the antecedent ... and sentential clauses have an entire clause as the antecedent.

(Rissanen 1999, 292).

Which relatives do not normally introduce nominal clauses, the typical relative in those cases being *what*. The only form of nominal relative clause found in the Shakespeare *which* corpora is one introduced by the combination *that which*, as in:

(10) How with mine honour may I give him *that*
Which I have given to you?

(*Twelfth Night*, 3.4.116-7)

Rissanen lists the possible forms of nominal relative clause in Early Modern English: “In addition to *who* and (*that*) *which*, [nominal clauses] can be introduced by *that*...and by *what*...”

(1999, 300), making clear the fact that *that which* introduces a nominal clause. Correspondingly, and in line with Ryden (1966, 72), I assume here that *that which* is paraphrasable on all occasions as *what*, the latter form being an alternative to *that which*, and an increasing variant through the course of the seventeenth century, as explained by Rissanen (1999, 300) and Kemp (1979, 185). These cases have therefore been identified for the present purposes simply as ‘nominal clauses’ (since the analytical categories employed will not include non-relative nominal clauses), and the pronoun *that* is treated as the antecedent, (with Ryden 1966, 72). Occasionally this construction has an intervention between *that* and *which*, such as

- (11) Are ye fantastical, or *that indeed*
Which outwardly ye show

(*Macbeth*, 1.3. 57-8)

In these cases the relative clause is still seen as nominal, since the intervening word is not a noun phrase.

Adnominal clauses are, as Rissanen suggests, by far the most numerous, and seem to represent the canonical type of relative clause. As will be seen in section 3.3.2.3, where the distance between the antecedent and *which* is the focus of attention, in order to count the words intervening between antecedent and relative it was necessary to make a policy decision concerning what to identify as the antecedent per se. Frequently it is a noun phrase containing several words, but it was decided that, in order to make measurement possible, in the present research the antecedent would be identified as just the head of the phrase, either noun or pronoun, and for clarity in the process of counting and analysis, antecedents reducible to a single word will be termed ‘single word antecedents’ throughout this thesis. An occurrence

from *Twelfth Night* exemplifies this kind, where “that heart” is the antecedent noun phrase, but the head of the phrase (“heart”) is deemed the antecedent for counting purposes. This is also an example of the canonical type of adnominal clause, where the antecedent is immediately to the left of the relative:

- (12) Yet come again; for thou perhaps mayst move
That heart, *which* now abhors, to like his love.

(*Twelfth Night*, 3.1.129-30)

Sometimes there is more than one noun as adnominal antecedent, as in the case of *tricks* and *ceremonies* in this speech from *Titus Andronicus*:

- (13) Yet, for I know thou art religious
And hast a thing within thee called conscience,
With twenty popish tricks and ceremonies
Which I have seen thee careful to observe

(5.1.78-81)

As it seemed possible that multiple antecedents could be associated with syntactic complexity, they were counted separately in this study. In such instances the last of the nouns is assigned as the antecedent for counting purposes, also the case with antecedent nouns in apposition, and those where an intensive verb creates a complement, a semantic parallel between noun phrases. In all these cases it is, of course, also necessary to consider semantic relationships and use discretion in individual occurrences. The following is an example of a complement construction where the second noun phrase is not taken as the antecedent:

- (14) Those scraps are good deeds past, which are devour'd
As fast as they are made

(*Troilus and Cressida*, 3.3.157-8)

The main verb of the relative clause here clearly relates more closely to “scraps” than the complement “deeds”, so the head word of the first noun phrase is taken to be the antecedent in this case.

The final type of relative clause is sentential, where a verb phrase or whole clause is the antecedent of the relative, as exemplified in (15) and (16) below. Sentential relative clauses are potentially significant for the present analysis, since as Rissanen suggests “the link [between relative and antecedent] seems to be particularly loose when the antecedent is a clause or a VP” (1984, 420), and looseness of link is likely to be connected with syntactic complexity or processing difficulty.

- (15) Do it, then, that we may account thee a whore-master and a knave; which notwithstanding, thou shalt be no less esteemed.

(Timon of Athens, 2.2.107)

Here it is not “whore-master” or “knave” that are “notwithstanding”, but the fact “that we may account thee a whore-master and a knave”. The antecedent is not reducible to a single word.

Similarly, in the following case the antecedent is a non-finite clause, “to bottom it on me”:

- (16) Therefore, as you unwind her love from him,
Lest it should ravel and be good to none,
You must provide to bottom it on me,;
Which must be done by praising me as much
As you in worth dispraise Sir Valentine.

(The Two Gentlemen of Verona, 3.2.54-8)

3.3.2.2 The animacy parameter

Although in Present Day English *which* is only used for inanimate antecedents, a different situation pertained in Early Modern English, as mentioned in section 3.2 above.

Adamson explains:

[I]t was originally a gender-neutral form, occurring with both animate and inanimate antecedents... *Who* gradually became established as a generally available option in the sixteenth century, thus providing the formal possibility for the expression of the personal/non-personal contrast we have today. But in 1600, usage was still unregulated and it was not uncommon to find collocations of the type ‘the man which’ and ‘the thing who’.

(Adamson 2007, 3.2).

We find the following in *Macbeth*, for example:

(17) This disease is beyond my practice: yet I have known those *which* have walked in their sleep who have died holily in their beds”

(5.1.29-30)

and this in *Antony and Cleopatra*:

(18) less noble mind
Than she *which* by her death our Caesar tells
“I am conqueror of myself”

(4.12.74-76)

It is perhaps conceivable that, for Present Day English auditors, the use of an animate antecedent could contribute to difficulty in processing a *which* clause, or to its perceived complexity, but the principal reason for counting *which* antecedents in terms of the animacy parameter was that this criterion is so integral to the nature of the antecedent, and such a widely discussed issue in Early Modern linguistics, that it seemed too important to exclude. The diagram below represents in graphic form the relationships between the categories discussed so far.

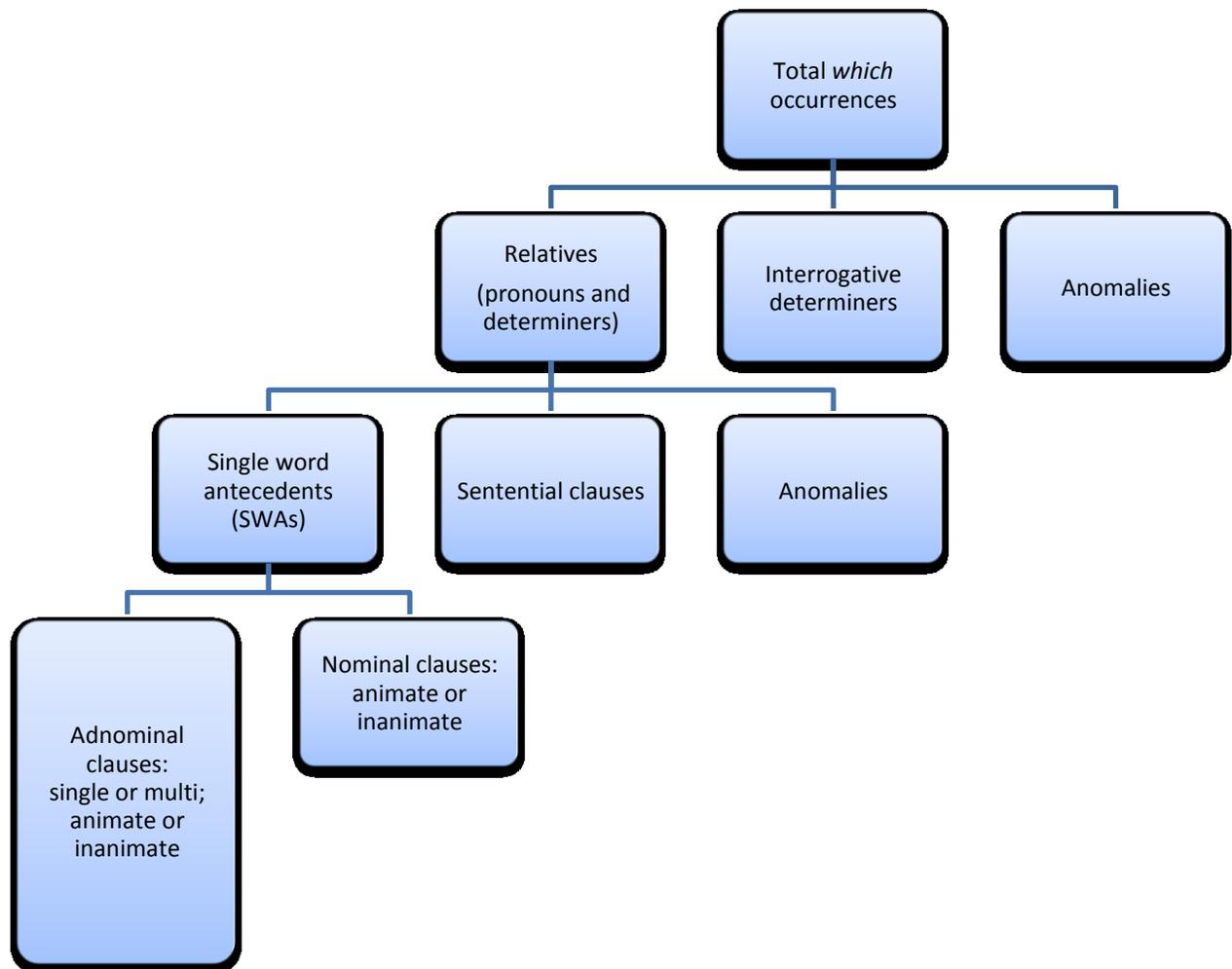


Figure 3.1: Analysis of *which* clause antecedent type

3.3.2.3 Distance from the relative

Ryden suggests that “[i]t is an observation of long standing that the distance between the antecedent and the relative clause plays a part in the selection of [the relative], with distancing favouring the wh- relatives” (1983, 215). Distance, or the number of intervening words, is one of the principal factors involved in the loosening of the link between antecedent

and relative. In his 1984 study of Early Modern relativisation, Rissanen argues that “[o]f the purely syntactic factors relevant to the tightness or looseness of the link, the separation of the relative clause from its antecedent is the most obvious” (1984, 420). Given then the probable connection between looseness of link, potential for ambiguity and difficulty of processing, the criterion of distance between antecedent and relative was felt to be an important one in charting the development of Shakespeare’s use of *which* across the canon. Some occurrences, such as those containing sentential relative clauses, are identified as having a relationship between relative and foregoing clause that is “particularly loose”, as seen in 3.3.2.1 above (Rissanen 1984, 421), but in this section we will be exclusively concerned with cases where there is a single word antecedent (adnominal and nominal clauses only), and so a precisely measurable distance between antecedent and relative, along with the possibility of analysis of the intervening linguistic material.

If the canonical single word antecedent immediately to the left of *which* (as in (12) above) is the least ambiguous and easiest to process, on other occasions things are more difficult. Early Modern English seems to have permitted its writers considerable freedom and flexibility in the distance between antecedent and relative, and in the nature and length of the words, phrases, even clauses, that intervene between them. In Present Day English we process anaphoric relative pronouns according to what George Dillon describes as “[t]he traditional stylistic precept – that the first eligible noun phrase to the left of the pronoun should be the antecedent” (1978, 69). In Shakespeare’s plays, although (as will be seen from the results of this analysis in Chapter 4) there is still an overall preference for antecedents to be placed immediately to the left of the relative (as in example (13) above), we find many occurrences that challenge this norm and would be unacceptable today. This speech is from *King Henry VI*

Part 2, where Warwick uses chilling rhetoric to persuade the king that Gloucester, whose body lies before them on stage, did not die naturally but was murdered:

- (19) See how the blood is settled in his face.
Oft have I seen a timely-parted ghost,
Of ashy semblance, meagre, pale and bloodless,
Being all descended to the labouring heart,
Who, in the conflict that it holds with death,
Attracts the same, for aidance 'gainst the enemy;
Which, with the heart there cools, and ne'er returneth
To blush and beautify the cheek again.
But see, his face is black and full of blood...

(3. 2. 60-168)

Employing Dillon's "first eligible noun phrase to the left" principle, there are numerous red herrings here, that seem to invite a reading as the antecedent, such as "enemy", "death", "heart" and "ghost", all syntactically feasible noun phrase antecedents that yet offer no semantic resolution. In order to find the antecedent that makes sense, we have to look back as far as the preceding sentence, to "blood". Admittedly, we are guided to that noun by the rhetorical thrust of the passage, with its repeated focus on the topic of blood, and more particularly perhaps by the fact that *which* is the third in a sequence of anaphoric references to the word "blood", after "being" (which implies "blood" as its subject), and "the same". Nonetheless, this speech stretches and tests the relationship between relative and antecedent, in a way that many modern auditors, even readers, would probably find challenging to process. In the closing scene of *Love's Labour's Lost*, one of the canon's most extreme examples of a relative remote from its antecedent occurs:

- (20) Your beauty, ladies,
Hath much deformed us, fashioning our humours
Even to the opposed end of our intents;
And what in us hath seemed ridiculous, –
As love is full of unbefitting strains,

All wanton as a child, skipping and vain,
Formed by the eye and therefore, like the eye,
Full of strange shapes, of habits and of forms,
Varying in subjects as the eye doth roll
To every varied object in his glance;
Which parti-coloured presence of loose love
Put on by us...

(5.2.740-51)

Close analysis here reveals that the distant nominal clause “what in us hath seemed ridiculous” is the antecedent recapitulated by *which* in line 750, and the intervening lines comprise a series of adjuncts. Craig signals the start of the intervening digression by adding both a comma and a dash after “ridiculous”, but is not alone in punctuating this point heavily. Henry Woudhuysen, editor of the 1998 Arden edition of the play, has a dash here, and the First Folio gives a colon after “ridiculous” and after “glance”, performing a comparable function (Hinman 1996, 161). This of course only suggests that other readers (editors, compositors, printers, perhaps) have sensed the need for explanatory punctuation in this speech, and have identified the semantic link between antecedent and relative. It does not necessarily mean that Shakespeare did so. But I think it can be seen here that there was considerable licence in Shakespeare’s Early Modern English as to the distance between antecedent and relative, and also that syntactic opacity can certainly result from such distancing, rendering it an appropriate factor to measure in the comparative analysis of the two corpora.

It was decided that two approaches would be used in the measurement of this distance. Eva Berlage’s (2008) study of noun phrase complexity, which examines certain Present Day English linguistic variations with the hypothesis that syntactic complexity causes them, surveys methods of measuring complexity, thus offering a useful point of reference for this

analysis. She compares the straightforward counting of words with counting of phrasal units (or nodes), and finds that both are appropriate gauges of complexity, with some variational features responding to one, and others to the other. It was therefore felt appropriate to employ both approaches here for the analysis of the linguistic matter intervening between antecedents and *which*.

The method employed for counting involved starting at the antecedent and counting the next word to the right as ‘1’, until *which* was reached, and results in the number of words that are neither antecedent nor *which*, that intervene between the two. As a demonstration, this method gives a distance of three words between the antecedent “fortunes” and “which”, in the following case from *The Winter’s Tale*:

(21) For myself, I’ll put
 My fortunes *to your service*, which are here
 By this discovery lost.

(1.2.511-13)

In a similar way, there are 43 words between “blood” and “which” in (19) above, counting the hyphenated words “timely-parted” as one (a policy employed throughout this study). Multiple antecedents arise on occasions, such as those italicised here:

(22) Besides *the applause and approbation*
 The which, most mighty for thy place and sway,
 And thou most reverend for thy stretch’d-out life,
 I give to both your speeches

(*Troilus and Cressida*, 1.3.62-5)

In such cases, a similar policy is adopted to that described earlier³⁹, and the final noun head word is taken as the antecedent from which counting begins, giving a one word intervention in this instance.

3.3.2.4 Word groups intervening to the left

The second approach involved identifying the number and nature of the word groups (phrases and clauses) intervening to the left of *which*, between antecedent and relative. Here we are not only concerned with the size of the intervention but its syntactic complexity, with the hypothesis that there may be a connection between the number and/or nature of intervening phrases or clauses and the degree of complexity or opacity created.

The identification of individual word groups, however, is by no means straightforward or given. Jackson, for example, uses phrase length as an authorship determinant, and for him “length of phrase is defined as the number of words intervening between any of the following marks of punctuation: full stop, question mark, exclamation mark, colon, semi-colon, parentheses...dash, and comma” (1997, 76). However, there is such variation in punctuation between editions of the plays that this method is likely to prove unreliable. Craig, for example, “thoroughly revised” (1894, vi) the punctuation of the edition used in this analysis, very generously adding to the number of commas and semi-colons in each text, as can be seen later in this study⁴⁰. There is no authoritative source for punctuation in the Shakespeare canon, due to the freedom with which scribes and composers intervened in the production of the texts

³⁹ See section 3.3.2.1.

⁴⁰ See for example 3.3.4.6.

(see for example Jowett 2007, 57 and Ryden 1966, xlii). Jackson even acknowledges this situation himself a few years later:

[H]owever one chooses to define a sentence, the punctuation of most English Renaissance dramatic texts is so patently not authorial, and editors' approaches to their repunctuation are so diverse, that tests based on the position of words within the sentence cannot reasonably be used on this kind of material.

(Jackson 2003, 110).

Even if punctuation is discarded as a guide, Shakespeare's syntax is often dense and convoluted, and (as in Present Day English) phrases and clauses are frequently embedded at different levels, often resulting in a lack of clarity as to where word groups start and finish, and meaning that some limitations needed to be acknowledged in order to render the counting practicable. A 'top level only' approach was therefore chosen in this analysis, which will be made clear through exemplification below. It did however seem important to distinguish between different types of word group intervening, (something that simple counting of phrase boundaries would not have offered), so a set of categories was assigned, and occurrences of each were counted separately. It seemed likely that verbal⁴¹ elements intervening may increase both syntactic complexity and the looseness of link between the antecedent and the relative, more so than phrases. Berlage adopts the same hypothesis (in her 2008 analysis of the relative complexity of noun phrases mentioned earlier), distinguishing incrementally between "simple NPs", "more complex but non-sentential NPs" and "NPs containing a VP" (2008, 4). In this part of the analysis, conjunctions, prepositions, determiners and the like (such as *and*, *the*, *to*, etc.) appearing alone between the antecedent and *which*, without any other items, were not

⁴¹ The sense intended is 'pertaining to verbs' rather than simply 'word-related'.

counted as intervening word groups, since they were seen as structurally fundamental to the clause. The following analytical categories were employed:

Phrase

This was defined as a group of words without a verb (rather than simply a NP, for example), and was most typically a prepositional phrase, although single lexical words such as adjectives or adverbs were also counted in this category. Examples include:

(23) it lies as coldly in him as fire *in a flint*, which will not show without knocking

(*Troilus and Cressida*, 3.3.267)

(24) with terms *unsquared*
Which from the tongue of roaring Typhon dropped

(*Troilus and Cressida*, 1,3,159)

Non-finite clause

An example of a non-finite clause intervening to the left, between *which* and its antecedent, is the following from *The Winter's Tale*:

(25) His cup-bearer,- whom I from meaner form
Have bench'd and rear'd to worship, who mayst see
Plainly, as heaven sees earth, and earth sees heaven,
How I am galled – mightst bespice a cup,
To give mine enemy a lasting wink,
Which draught to me were cordial.

(1.2.365-70)

Here, the antecedent “cup” is separated from *which* by the non-finite clause “to give mine enemy a lasting wink”. This category also includes verbs, or verb phrases with accompanying noun phrases, such as:

- (26) You have beguiled me with a counterfeit
Resembling majesty, which being touched and tried
Proves valueless.

(*King John*, 3.1.103-5)

The decision to conflate groups in this way was taken in order to simplify what is already a complex set of features, and to focus on the syntactic qualities most salient to the inquiry in hand. In this general category (word groups intervening), the principal issue is the distinction between intervening elements containing simply phrases, on the one hand, and verbal⁴² material (characteristic of what Rissanen calls “heavy elements” separating antecedent and relative (1984, 424)) on the other. As far as finiteness and non-finiteness are concerned, it seemed possible that intervening elements containing finite verbs have the potential to create an additional sense of semantic separation between antecedent and relative, so these were counted separately.

It is important to stress that the labels ascribed to the word groups intervening between antecedent and *which* are not intended as descriptions of their form in the sentence as a whole. The important criterion employed here is that it is strictly only the material intervening between the single word antecedent and the relative *which* that is counted, so in case (27) below, the auxiliary verb “have” (which of course renders the main clause finite when linked to “set down”) is outside the intervention, and only the non-finite (phrasal verb) component and its adverb are inside it. In order to distinguish this type of intervention from one where a whole (for example relative) clause intervenes between the antecedent and *which*, it is counted as a non-finite clause intervening:

⁴² As above, meaning ‘pertaining to verbs’.

- (27) I have a letter *guessingly set down*,
Which came from one that's of a neutral heart,

(*King Lear*, 3.7.46-7)

Partial finite clause

- (28) Oh, when degree *is shak'd*,
Which is the ladder to all high designs

(*Troilus and Cressida*, 1.3.104-5)

This construction is widespread in Shakespeare. “Degree” is here the antecedent of the relative clause, but (in a way that is unlikely in Present Day English) it is followed by its own predicate (and in some other cases an object as well), constituting a clause in its own right, before the relative appears. The most common occurrence of this type, as in the case of the example above, comprises a copular verb (passive “is” in this instance) with a complement intervening between antecedent and *which*. Such cases have been identified as partial finite clauses in order to distinguish them from other cases in the canon, where a whole (usually relative) clause intervenes between antecedent and *which* (see “Finite clause” below), with a subject of its own that is not the antecedent.

Finite clause

These are cases where another whole clause intervenes between antecedent and relative clause, as in the following two examples:

- (33) This is Worcester, *malevolent to you in all aspects*,
Which makes him prune himself, and bristle up
The crest of youth against your dignity.

(*Henry IV Part I*, 1.1.98-100)

The principal criterion for deciding whether an intervention is a single or multiple phrase was semantic separateness, namely whether one could be removed and leave the other in the clause making full sense. This is not a completely watertight distinction, but it was found to work sufficiently well here to implement it. In (32) above, “In name of lendings” and “for your highness soldiers” are counted as two phrases, as are “malevolent to you” and “in all aspects” in (33), despite the fact that in this case the removal only leaves a semantically sound phrase if the first (“malevolent to you”) is retained, and not the other way round.

A few occurrences have more than one intervening clause, or a combination of partial finite clause(s) and clause(s), and sometimes even one or more phrases as well. Because they are fairly rare, and the multiple predicates seemed to be the factor that may suggest a potential for complexity, the decision was taken to count all such cases together, labelling them as ‘multi clause’ interventions, and in these cases the additional presence of one or more phrase was not noted. The following are examples:

- (34) Peace *be to you*,
As I pursue this war, which shall be then
Beyond further requiring.

(*The Two Noble Kinsmen*, 1.3.24-26)

The antecedent here is “peace” (although this is another example of a loose and ambiguous link), followed by a partial finite clause, “be to you”, and then a complete clause, “As I pursue this war”. This occurrence from *The Tempest* is similar:

- (35) For I *am all the subjects that you have*,
Which first was mine own king.

(1.2.406-7)

“I” is the (again) rather distant antecedent with a loose link to the relative, followed by a verb phrase “am all the subjects” and a complete relative clause “that you have”. A more complex case is this from *Cymbeline*:

- (36) I have given him that
Which, if he take, shall quite unpeople her
Of leigers for her sweet, and which she after,
Except she bend her humour, shall be assured
To taste of too

(*Cymbeline*, 1.5.91-5)

Here, looking from the perspective of the relative “which” in line 93, which heralds the second coordinated relative clause, and taking “that” in the first line to be its antecedent, we see a subjunctive clause “which if he take” embedded in the first relative clause “which...shall quite unpeople her”, and followed by two phrases, “of leigers” and then “for her sweet”. This is an example of a multi-clause intervention.

More frequent in the canon, and counted as a distinct category, are intervening combinations of one single clause (or verb phrase) and one or more phrases, an example of which can be seen here, in *The Two Gentlemen of Verona*:

- (37) But think upon my grief, a lady’s grief
And on the justice *of my flying hence*,
To keep me from a most unholy match,
Which heaven and fortune still rewards with plagues.

(4.3.33-6)

Between “justice”, the single word antecedent here, and “which”, two phrases intervene first, namely “of my flying” and then the adverbial phrase “hence”, followed by the non-finite

clause “[t]o keep me from a most unholy match”. This section can be represented diagrammatically as Figure 3.2, below.

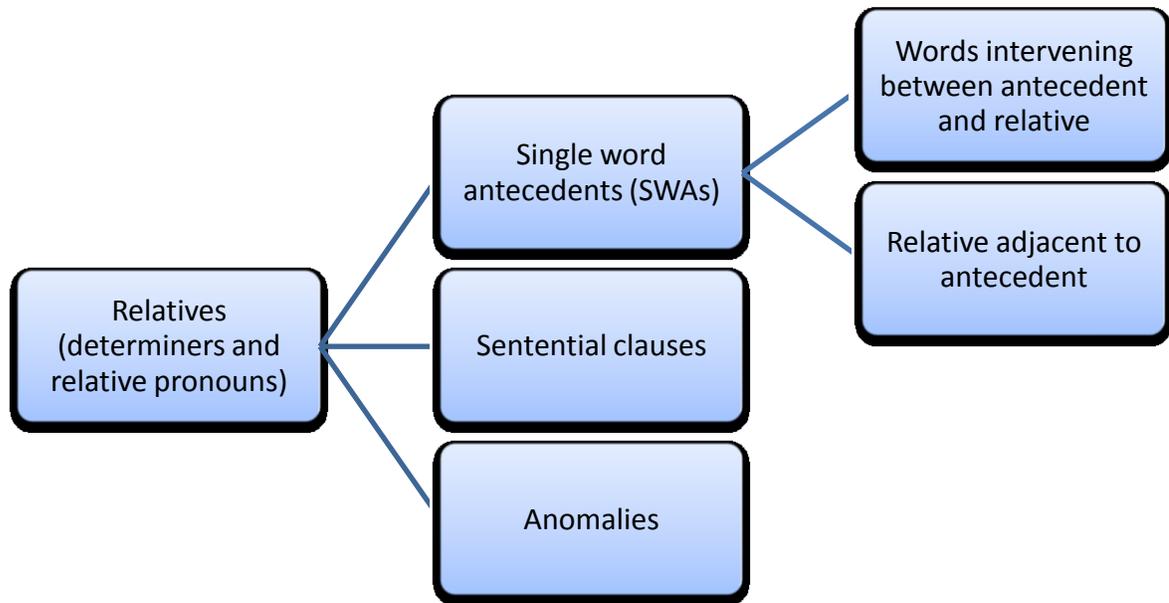


Figure 3.2. Single word antecedents

3.3.3 The information parameter

3.3.3.1 Restrictive and non-restrictive clauses

There are other factors involved in the degree of looseness of the antecedent/relative link, such as what Fischer terms the “information parameter” (1992, 286), or the distinction between restrictive and non-restrictive clauses. Here we are concerned with a more explicitly semantic relationship than in the previous section, and one which therefore requires careful definition in order to measure accurately. Discussing this distinction, Ryden suggests that “[t]he identification of semantic dependencies is a matter of interpretation, but this

interpretation may be based on formal criteria” (1983, 212). He describes the difference between these two types of relative clause as follows:

A non-restrictive clause does not affect the reference of the antecedent; it only gives additional or supplementary information about it. The antecedent may be an entire clause; this is impossible in the case of restrictive clauses

(Ryden 1966, xlvii).

Huddleston and Pullum use different labels for these clause types⁴⁴ in their useful definition of the pairing in Present Day English, although the appropriateness of their account to restrictives and non-restrictives in Early Modern English usage seems to suggest that there has been no significant change since the Early Modern period in the semantic concepts involved in this distinction:

The terms integrated [restrictive] and supplementary [non-restrictive] indicate the key difference between them: an integrated relative is tightly integrated into the matrix construction in terms of prosody, syntax, and meaning, whereas a supplementary relative clause is related only loosely to the surrounding structure.

(Huddleston and Pullum 2002, 1058)

They continue later with an additional description of the semantic operation of these clauses, which proves a useful test in discriminating between them. They demonstrate that the removal of the relative clause in integrated (restrictive) cases “would drastically change the meaning”, whereas “supplementary relatives ... by contrast, can be omitted without affecting the meaning of the remainder” (2002, 1059). Jespersen also includes this test in his definition of “[n]on-restrictive or loose clauses, which might be discarded without serious injury to the precise

⁴⁴ Elsewhere restrictive and non-restrictive clauses are also known as “attributive” and “appositive” respectively, for example in Traugott, E. (1972) *A History of English Syntax: a Transformational Approach to the History of English Sentence Structure*. New York: Holt, Rinehart and Winston. Mats Ryden (1966) and Tauno Mustanoja (1960) both refer to them as “defining” and “non-defining”.

understanding of the sentence as a whole” (1949, 82). The following are occurrences of *which* in restrictive clauses, from the Shakespeare corpora:

- (38) And my estate deserves an heir more rais'd
Than one *which holds a trencher*

(Timon of Athens, 1.1.141-2)

- (39) Is not parchment made of sheep skins?
Ay, my lord, and of calf skins too.
They are sheep and calves *which seek out assurance in that.*

(Hamlet, 5.1.42-4)

And these are examples of non-restrictives:

- (40) This youth that you see here
I snatched one half out of the jaws of death,
Relieved him with such sanctity of love,
And to his image, *which methought did promise*
Most venerable worth, did I devotion.

(Twelfth Night, 3.4.203-7)

- (41) Sweet are the uses of adversity,
Which, like the toad, ugly and venomous,
Wears yet a precious jewel in his head

(As You Like It, 2.1.14-16)

In (38) and (39) the removal of the (italicised) relative clause would render the meaning very unclear and the sentences unworkable, whereas in both (40) and (41) it leaves a perfectly acceptable sentence.

The importance of the information parameter can be seen from a number of comments made by Early Modern linguists. For Ryden, “the character of the relative clause with respect to the antecedent and the clause containing the antecedent” (essentially whether it is restrictive or non-restrictive) is another conditioning factor which has “ particularly decisive

distributional effect” in the use of Early Modern relatives (1966, xx). Since the earliest use of *wh-* forms, in Middle English, the restrictive/non-restrictive dichotomy has been a significant factor in the choice of relative, as we can see from Rissanen’s overview of the development of *wh-* relatives:

At the end of the Middle English period, *that* was the most common adnominal relative link...although there was a tendency to prefer *which* in non-restrictive clauses. The inflected forms *whom*, *whose* were common with personal antecedents in non-restrictive clauses ... As early as the sixteenth century *wh-* pronouns are well established in all types of non-restrictive clauses

(Rissanen 1999, 293).

Among other examples of this connection, *which* as a determiner only occurs in non-restrictive clauses⁴⁵, and “the zero link is confined to restrictive relative clauses” (Rissanen 1999, 298). Clearly then, this distinction needed to be part of the set of categories in the current analysis, particularly since, as seen above in the Huddleston and Pullum definitions for example, non-restrictive clauses are associated with a looser link between antecedent and relative, because the information they add to the antecedent clause is less directly related to it. As we have also seen, there is a likely connection between looseness of link and the potential for syntactic complexity, so non-restrictive clauses, associated with *wh-* forms and looseness of link, tend to be seen as more syntactically complex than restrictive clauses. Dekeyser and Ingels articulate this idea as follows:

Given the fact that non-restrictive clauses only provide additional information about the antecedent NP, and that less complex language, such as spontaneous speech, tends to present this in a paratactic or non-embedded structure, we expect [non-restrictive clauses] to be a mark of the more complex registers (1988, 32).

⁴⁵ See section 3.3.1.

3.3.3.2 Continuative clauses

There is more to the situation than that dichotomy, however, since there is another clause type, a subset of non-restrictives, which was very much in evidence in Early Modern English writing, representing an even further step in the direction of looseness of link and potential for semantic complexity, namely the continuative clause. In cases such as these, there seems to be something more occurring, both syntactically and semantically, than in a straightforward non-restrictive clause:

- (42) Whereupon I will show you a chamber and a bed; *which bed, because it shall not speak of your pretty encounters, press it to death*

(Troilus and Cressida, 3.2.148-9)

- (43) I am not prone to weeping, as our sex
Commonly are; *the want of which vain dew
Perchance shall dry your pities*

(The Winter's Tale, 2.1.133-5)

- (44) Do not embrace me till each circumstance
Of place, time, fortune, do cohere and jump
That I am Viola; *which to confirm,
I'll bring you to a captain in this town*

(Twelfth Night, 5.1.229-32)

Because precise definition of the continuative clause presents a number of problems, it will be useful to begin discussion with some of the principal descriptions of this type. Huddleston and Pullum, giving an account of twentieth century English usage, exemplify and describe what they call “[t]he continuative use of supplementary relatives” as follows:

I gave it to John, *who passed it on to Mary*, and she gave it back to me...

Illustrate[s] a use of juxtaposed supplementary relatives in narrative contexts that is traditionally referred to as **continuative**: they serve to continue, to develop, the narrative. The effect is like that of *and* + non-relative anaphoric expression: I gave it

to John and he passed it on to Mary... Whereas elsewhere the information conveyed in a supplementary relative is somewhat backgrounded relative to that conveyed in the clause containing the anchor, the continuative relative has equality of informational status, presenting a further event in a narrative chain

(Huddleston and Pullum 2002, 1064)

Quirk (1985), also concerned with Present Day English, does not define the continuative clause as such, but exemplifies it in a footnote with similar cases to Huddleston and Pullum, such as: “Tom lent the book to Sue, who lent it to Pat, who returned it to Tom” (1985, 1120). Jespersen briefly refers to the continuative clause thus: “A subdivision [of non-restrictive clauses] is formed by continuative clauses, in which *who* and *which* might be replaced by *and he*, *and she*, *and it*, or *and they*” (1934, 82). These commentaries describe a type of clause found in narrative contexts, seemingly straightforwardly delimited by their identification as coordinating rather than subordinating clauses (as is also suggested by the test for them, which involves their replacement by the coordinating conjunction *and*).

The situation is more complex, however, in Early Modern English, where this type of clause is not only much more common than in Present Day English, but is not by any means restricted to such narrative contexts, as can be seen from the examples given at the start of this section, and others from the Shakespeare corpus, such as (20) above. Considering example (42), for instance, there is no overtly narrative development with the continuative clause, certainly not in the same way as the modern English examples, where typically a material process was reported in the (narrative) simple past tense. Clearly, though, a new idea is introduced with the predicate of the relative clause, “press it to death”. *Which* is a relative determiner in this case, with the noun phrase *bed* recapitulating the antecedent in the foregoing clause. Relative determiners frequently appear in continuative clauses (as observed by

Rissanen (1999, 296)). The clause does have features in common with the modern cases, however, as the relative could be substituted by *and that* without any change of meaning (as suggested by Huddleston and Pullum and Jespersen), demonstrating the coordinating relationship between the two clauses linked by *which*.

Like the Present Day English grammarians, Rissanen sees the coordinating function as definitive for Early Modern continuatives: “[I]t has proved useful to distinguish a special type of non-restrictive clause called ‘continuative’. In this type the two clauses stand in coordinating rather than subordinating relationship” (1999, 293), and he offers the following examples:

in somme places they mowe it, *the whiche* is not soo good to the hosbandes profyte...

‘How now Perrott’ (quoth the Kinge) ‘what is the Matter that you make this great Moane?’ *To whom* Sir John Perrott answered [...]

(Rissanen 1999, 292).

The second example here is more like the Present Day English cases cited, with a stronger sense of narrative progression. But Rissanen argues elsewhere that the Early Modern continuative clause involves other pivotal features, for example: “a non-restrictive clause is continuative when the link between it and the main clause is very loose” (2008, personal communication).

Ryden’s 1966 monograph-length study of early sixteenth century relative constructions gives this description⁴⁶, again focusing on looseness of syntactic link:

There are, particularly in Early Modern English, relative clauses which exhibit a high degree of syntactic autonomy. In German terminology, they are said to be connected with the preceding clause by ‘relativische Anknüpfung’, or ‘relativische Anschluss’.

⁴⁶ Ryden, and also at times Rissanen, call continuative clauses “progressive”, but see the terms as interchangeable (e.g. Ryden 1966, xlvi and Rissanen 1984, 422).

Although such structures stand in loose syntactic relation to what precedes, the linking together of the two predications is still effected by means of a relative connective. Hence, from a formal point of view, they are to be regarded as relative clauses. Such loosely appended clauses are often preceded by a full stop or other marks of heavy punctuation. Considering, however, the inconsistent use of graphic symbols in the texts examined [early sixteenth century prose texts], we are not justified in setting up a special category on the basis of punctuation. These clauses are all non-restrictive[...] The antecedent may be a word (word-group) or an entire clause. The relative clause, which is invariably appended to (never included in) the clause containing or constituting the antecedent, is often placed at a considerable distance from the antecedent.

(Ryden 1966, xlii-xliii).

Ryden's description here is very pertinent to the continuative clauses found in the Shakespeare corpora. As can be seen from the Shakespeare examples given so far, the autonomy of the relative clause is manifested at least partly by its semantic departure from the antecedent clause: it introduces a new idea, and in that way progresses the discourse, if not strictly in a narrative sense. This autonomy is often expressed in the punctuation, as they frequently occur after a semi-colon or full stop (even occasionally after a change of speaker), although as Ryden says, punctuation is an unreliable gauge in the texts in question, and this is very much the case in Shakespeare.⁴⁷ Rather than the narrative and coordinating functions, then, the emphasis of definition has now shifted to the looseness of relative link, and Ryden notices an important additional formal feature of continuatives, namely the fact that they never interrupt the foregoing clause⁴⁸.

Nevalainen highlights the syntactic hiatus between antecedent and relative clauses in his definition of the Early Modern continuative clause:

⁴⁷ There are numerous accounts of the unreliable provenance of Shakespeare's punctuation, but see for example Jowett (2007, 57).

⁴⁸ For an example of a non-restrictive relative clause that does interrupt the antecedent clause, and therefore cannot be continuative, see (40).

Modelled on Latin, continuative relative clauses could begin a sentence, or even a new Paragraph...and were intended to improve the cohesion of the text. They are very frequent especially in the sixteenth century.

The example he then provides, from Francis Bacon, illustrates this point well:

For it is one thing to set forth what ground lyeth vnmanured; and another thing to correct ill husbandry in that which is manured. In the handling & vndertaking of **which** worke, I am not ignorant, what it is, that I doe now mooue and attempt, nor insensible of mine own weakenes, to susteine my purpose [...]

(Nevalainen 2006, 86).

Additionally, his related comment that “*Demonstrative* pronouns and determiners (this, these, that, those) could be used for similar purposes” (2006, 86) concurs somewhat with the modern grammarians’ definition, in the replacement of *which* with a non-relative anaphoric pronoun.

Romaine, in her account of relative clauses in Middle Scots from a socio-historical perspective, describes “resumptive or continuative” clauses as “advanc[ing] the discourse by adding new information”, saying that they “in many cases give information about an antecedent which is further back in the discourse, i.e. not next to the relative marker” (1982, 83), and so agrees with Rissanen in many respects. Interestingly however, especially given the fact that she discusses coordination in relative clauses at some length in the same chapter, she does not offer the coordinating function as a defining feature of continuatives. It is useful to note that she articulates the antecedent distance criterion in non-absolute terms, thereby acknowledging that there are some cases of continuatives where the antecedent is close to the relative.

Finally, Reuter’s article ‘On continuative relative clauses in English’ is devoted to the characterisation and historical mapping of continuative clauses. According to Reuter,

The high tide of our construction is not ... reached until the 16th century. ... In the 16th century this syntactical feature has so thoroughly saturated the language that it occurs abundantly everywhere in original writings, often in rather intricate relations to other clauses, relative concatenation being very popular⁴⁹ (1938, 39).

His description of this clause type is also very useful in its contribution to definition, so will be worth citing at some length:

On reading English prose of the 14th-17th centuries, and especially of the Elizabethan period, among syntactical constructions which seem more or less antiquated or strange to a modern linguistic feeling there is one in particular which arrests one's notice because of its frequency and very often its awkwardness. This is the construction by which many sentences, though giving fresh information such as would generally require a greater pause after the previous sentence, are introduced by a relative pronoun or a relative adverb intended to denote the close relation to the preceding context;

(Reuter 1938, 1).

What he rightly calls the “awkwardness” or “antiquated” appearance of Early Modern continuatives highlights the difference between their usage then and now. Examples (20) and (43) above, for instance, are typical examples of what Rissanen, Ryden and Reuter envisage as continuatives, with their widely separated antecedents and relatives striking oddly to modern readers, who might at times struggle to make the link between relative and antecedent. But this awkwardness is one reason why continuatives need to be identified in the present study. Not only are they a dominant feature of Early Modern syntax, as Reuter suggests above, but they may well constitute a source of linguistic opacity. Reuter continues:

There the relative pronouns introduce sentences which further develop the narrative or trend of thought by adding fresh and important information, and the relative pronoun links together two successive sentences by referring to a substantive expression or the contents of the preceding passage. To denote the introduction of such new elements, personal or demonstrative pronouns, after a more clearly marked pause, in most cases a full stop, would generally be used in unsophisticated English, as e.g. >>*He* – was well

⁴⁹ For discussion in this study of what Reuter calls “relative concatenation” or multiple clauses embedded within the relative clause, see the section on pushdown clauses in 3.3.4.5.

beloved by rich and poor>> [*he* instead of *who* or perhaps *which*] >>When John perceived *this*, he said ->>; [instead of ‘Which when John perceived, he said’] [...] When continuative relative clauses as described above are employed, the antecedent is

frequently not the noun or noun equivalent nearest to the respective relative pronoun, but often has its place further back in the previous sentence⁵⁰.

(Reuter 1938, 2-3).

Reuter’s description encompasses the modern grammarians’ definitions, incorporating both the elements of semantic departure, and the potential replacement of the continuative relative by personal or demonstrative pronouns, and also includes Rissanen’s “looseness of link”, in the distance from the antecedent.

However, even given these extensive characterisations, problems with definition remain. Reuter may have devoted an entire study to the construction without mentioning such difficulties, but Rissanen is doubtful about the possibility of pinning down continuatives, declaring that “[b]ecause of problems of classification ... there are no reliable criteria for determining whether a relative clause is progressive or non-progressive” (1984, 422). And Ryden’s observation, cited earlier, is nowhere more relevant than in the case of continuative clauses: “It is one thing to characterise and label type examples, real or invented. Another and much more difficult task is to identify various specimens as they occur in actual texts” (1983, 212). The examples which follow typify some of these problems.

Firstly, although, as Reuter and others point out, the relative is frequently some distance from its antecedent in continuatives, this is not always the case, as can be seen from examples already cited, such as (42) above:

⁵⁰ Reuter’s punctuation symbols are reproduced as accurately as possible. Square bracket explanations are added in order to clarify his exemplification.

Whereupon I will show you a chamber and a bed; *which bed, because it shall not speak of your pretty encounters, press it to death*

Here there is an antecedent immediately to the left of *which*, in the canonical antecedent position. The looseness of link here seems to be created, not by the distance of the relative from the antecedent, but by the change of idea in the relative clause, where the new idea also involves semantic development of only the second of the two noun phrases (“chamber” and “bed”). The determiner function, with the recapitulated noun “bed”, is used here as a means of clarifying which antecedent is referred to by the relative, as is often the case with determiner *which* (as both Mustanoja (1966, 195) and Rissanen (1999, 296) point out). The use of the determiner is perhaps in itself an indication of a syntactic and semantic hiatus of some kind, since it usually represents an implied need for clarification, so even if the antecedent is close at hand, there can still be a loose link. Further confirmation of the status of this case as a continuative clause is achieved through Huddleston and Pullum’s test, replacing the relative with *and*⁵¹ and a non-anaphoric pronoun, which works in this case, resulting in the following semantically satisfying paraphrase: “Whereupon I will show you a chamber and a bed; *and that* bed, because it shall not speak...” This test confirms that the relative clause is in a coordinating relationship with the foregoing, and has “equality of information status” (Huddleston and Pullum 2002, 1064).

Another case that does not fit all the descriptive criteria is this, from the confrontation between Guiderius and Cloten in the fourth act of *Cymbeline*. Here we have what may seem

⁵¹ *And* is used here in its Present Day English function as a coordinating conjunction, as opposed to the wider spectrum of possible functions it could perform in Early Modern English. (See, for example, Rissanen 1999, 281 for an account of these).

like a short, straightforward, non-restrictive relative clause, with an antecedent again immediately to the left:

(45) Thou villain base,
 Know'st me not by my clothes?

 No, nor thy tailor, rascal,
 Who is thy grandfather. He made those clothes,
 Which, as it seems, make thee.

(4.1.105-9).

However, the replacement test would result in “He made those clothes, *and*, as it seems, *they* make thee”, surely a semantically satisfactory paraphrase, suggesting that there is a coordinating rather than subordinating relationship between the clauses, and therefore it is probably continuative. The ironic joke contained in the pun on “make” arguably constitutes a new idea, a semantic departure from the foregoing clause, in a way that would not operate if the relative clause were simply “which you are wearing”, for example. Other criteria seem to be unfulfilled, such as the looseness of link, and the distance of relative from the antecedent, although as we have seen above, the distance criterion is not definitive, only typical. It may well be that a semantic departure such as we see here functions in a similar way to distance between antecedent and relative. Perhaps various kinds of hiatus can serve. Rissanen did not accept this case as continuative, suggesting that the “link with the antecedent is pretty obvious and [the relative is] not far removed from the antecedent” (2008, personal communication), but if Rissanen’s own earlier characterisation of continuatives as “stand[ing] in coordinating rather than subordinating relationship” (1999, 293) is prioritised, and the pronoun replacement test taken seriously, we can be confident that this is a continuative clause. The importance and effectiveness of the test in this instance become more salient if viewed in contrast to an

‘ordinary’ non-restrictive clause where the relationship is uncontroversially subordinate, not coordinate. Example (41) above from *Twelfth Night*:

Sweet are the uses of adversity,
Which, like the toad, ugly and venomous,
Wears yet a precious jewel in his head

would become “Sweet are the uses of adversity, and, like the toad, ugly and venomous, he wears yet a precious jewel in his head.” Although the result is syntactically viable, there is a distinct change in the semantic content of the clause, which does not occur with (45). It was therefore decided that the success of the replacement test and the semantic shift involved in the joke in (45) were sufficiently convincing evidence for the ascription of continuative function to this clause, but it certainly demonstrates the difficulty sometimes involved in the decision. Confirmation of the importance of the subtle semantic difference between a subordinating non-restrictive and coordinating continuative clause is provided by Quirk:

To take one English example, it is frequently said that many non-restrictive relative clauses in English are substitutable for co-ordinate structures, and it is difficult to find an informant who will deny it in a given case. Yet the substitution involves consequential differences in tone and pattern, stress-distribution, and relative prominence, and careful examination of contexts in which either form occurs naturally indicates that a significant choice has been made – that at some level of analysis they *mean* something different.

(Quirk 1958, 40).

Another example of interesting results from the replacement test is the case of the relative clause included in these lines from *Titus Andronicus*, which might be considered a continuative candidate, if perhaps on the borderline:

- (46) Let him that thinks of me so abjectly
Know that this gold must coin a stratagem,
Which, cunningly effected, will beget
A very excellent piece of villainy.

(2.3.6-9)

If the semantic content of the relative clause were seen as a sufficient departure from the antecedent clause to merit further investigation, we might try the replacement test here for confirmation. This would yield “know that this gold must coin a stratagem, *and*, cunningly effected, *it* will beget a very excellent piece of villainy”. The replacement is clearly unacceptable since it shifts the antecedent from “stratagem” to “gold”. The original relative clause is subordinate, and cannot survive the paraphrase into a coordinated relationship.

A case from *Richard III* further demonstrates the borderline between ordinary non-restrictive, and continuative clauses. The relative clause here seems to offer a new idea to some extent, to progress the idea of the antecedent clause, although the close proximity of the relative to the antecedent lessens the sense of continuativity. The replacement test is therefore pivotal to a decision again:

- (47) Did Julius Caesar build that place, my lord?

He did, my gracious lord, begin that place,
Which, since, succeeding ages have re-edified.

(3.1.73-75)

The presence of *that* NP *which* seems to give the impression of a tight link between relative and antecedent, almost to the point where one suspects a restrictive clause, and has to carry out the kind of check described later here⁵² with example (119) (although, as with (119), we discover that the reference of “that” is anaphoric, so it is non-restrictive). It seems odd, then,

⁵² See section 4.3.1.

that it could be considered as a candidate for a continuative clause, if so close to being restrictive. However, the result of replacement with *and* and non-anaphoric pronoun would be: “He did, my gracious lord, begin that place, and, since, succeeding ages have re-edified it”, a satisfactory paraphrase, so probably the clause is indeed continuative.

Intuitively, and with some support from the aforementioned writers, it seemed that continuative clauses tend to create an impression of linguistic complexity and opacity. It therefore seemed worthwhile attempting to formulate a workable definition for this type, so that counting could take place. For the purposes of the present analysis, the key features of all the above definitions were taken and amalgamated into a set of criteria for the positive identification of continuative clauses, as follows:

1. A semantically satisfying paraphrase must be possible, where *which* (or other relative pronoun) is removed and a coordinating conjunction and non-relative anaphoric (such as a personal pronoun, or *it* or *that*) substituted. The possibility of inserting *and* (or *but*) in this way confirms the fact that it is a co-ordinating rather than subordinating clause, having what Huddleston and Pullum call “equality of information status” with the foregoing main clause (2002, 1064). The distinction between the appropriateness and inappropriateness of co-ordination with *and* is a fine one, but very important to this test. The relative clause must introduce enough of a new idea to warrant *and*. This is a sense of a new direction with the relative clause, not just additional information to the foregoing one.
2. The preceding antecedent clause must be a complete free clause, and the relative clause must not interrupt it.

3. A clause is very unlikely to be continuative if it is in a structurally parallel position (coordinated) with another.
4. Frequently, although not definitively, the antecedent will be distant from the relative. Features such as determiner *which*, and/or heavy punctuation preceding the relative clause, may contribute to the confirmation of the clause as continuative, although are secondary to criteria 1 and 2.

It will also be important to subdivide continuative clauses into adnominal and sentential categories, since it is possible that the combination of “contexts of loose relative link – [namely] continuative and sentential relative clauses” (Rissanen 1999, 295) could intensify the effect of syntactic opacity or complexity.

Before starting analysis and counting of these clauses, because there is an element of doubt amongst current writers on Early Modern continuative clauses about the possibility of defining them, it seemed prudent to seek substantiation of this definition method. Two independent raters were therefore engaged to undertake analysis of a sample of clauses from the corpora. The raters were doctoral research students of Language in the English department of the University of Birmingham. They were trained in the identification of continuative clauses according to the method described above, and provided with copies of the two corpora, with full access to the wider textual contexts, so that any semantic ambiguities could be dealt with. They also had reference materials to support understanding of the Shakespeare texts. They were asked to analyse 200 of the 2,207 *which* clauses, comprising two randomly selected sections of 100 concordance lines, one from each corpus. It was important not to prejudice the raters’ findings by preselecting clauses in any way, so the sample was left as it appeared in the

concordance, with a range of clause types represented. They worked completely independently, without any communication with each other⁵³.

As will be seen below, on every occasion at least one of the raters agreed with the initial ascription. 93.5% of clauses found both raters agreeing with the category ascription made in the main analysis, representing a substantial endorsement of the set of criteria established here. Table 3.10 below shows the results for the 12 clauses where difference of opinion occurred, with the ascription made initially by this writer, according to the criteria set out in this thesis, and then the ascriptions of each of the two raters. This is accompanied by a list of the clauses in question, numbered 1-12, including some surrounding linguistic context, and with the *which* clauses in question italicised. Each case is discussed in turn.

Table 3.10 Rating results: divergent ascriptions

R= restrictive
 N= non-restrictive
 C= continuative

Clause	Initial ascription	Rater 1	Rater 2
1	N	N	R
2	C	C	N
3	C	C	N
4	R	N	R
5	N	C	N
6	N	N	C
7	N	N	C
8	R	N	R
9	N	N/C	N/C
10	N	N	C
11	N	N	C
12	N	C	N

⁵³ The complete list of clauses analysed by the raters can be found in Appendix 8.

1. and when he speaks,
'Tis like a chime a mending; with terms unsquar'd,
Which, from the tongue of roaring Typhon dropp'd,
Would seem hyperboles.

(Troilus and Cressida, 1.3.164-7)

Rater 2 identifies this *which* clause as restrictive, despite the fact that the antecedent (“terms”) can stand quite satisfactorily on its own, without the supplementary qualifying clause that follows it. This seems to be simply an error on the part of the rater, since there is no real case for the relative clause being definitive of the antecedent, as would be necessary for a restrictive clause.

2. [T]ake me by the hand and say, ‘Harry of England, I am thine:’ *which word thou shalt no sooner bless mine ear withal, but I will tell thee aloud – England is thine, Ireland is thine, France is thine, and Henry Plantagenet is thine.*

(King Henry V, 5.2.130-2)

Rater 2 sees this as non-restrictive rather than continuative. It seems however to fulfil all the criteria presented above for continuatives, and is in several respects a classic determiner-*which* continuative clause. The replacement test produces a fully satisfactory semantic outcome, as “and that word thou shalt no sooner bless mine ear withal”, particularly so since the relative clause introduces a new narrative idea.

3. By mine honour, in true English I love thee, Kate: *by which honour I dare not swear thou lovest me:*

(King Henry V, 5.2.125)

This case is somewhat more borderline than the previous one. The decision seems to hinge on the criterion of the novelty of the idea represented by the relative clause. A semantically satisfying paraphrase can be achieved if “which” is replaced with the coordinating conjunction *but* (instead of *and*), producing “but by that honour I dare not swear

thou lovest me”, which is the reason for the initial ascription of it as continuative. The disjunction seems to offer sufficient novelty of idea for this to be acceptable, and the heavy punctuation (Wells and Taylor (2005) even use a full stop) is also suggestive of continuativity.

4. Vouchsafe to those that have not read the story,
That I may prompt them: and of such as have,
I humbly pray them to admit the excuse
Of time, of numbers, and of due course of things
*Which cannot in their huge and proper life
Be here presented.*

(King Henry V, 5.0.2-7)

This seems a fairly uncontroversial case of a restrictive relative clause, since “things” is defined by what follows. “[T]hose” could be inserted comfortably to produce “due course of those things which”.

5. [O]f princes in this number,
And nobles bearing banners, there lie dead
One hundred twenty-six: added to these,
Of knights, esquires and gallant gentlemen,
Eight thousand and four hundred; *of the which
Five hundred were but yesterday dubb'd knights*

(King Henry V, 4.8.81-6)

Although technically it is possible to paraphrase this clause with “and five hundred of them were but yesterday dubbed knights”, it seems a case where the shift to explicit coordination causes a subtle semantic infelicity with the original, such as that discussed by Quirk (1958, 40), cited earlier⁵⁴, so it is not seen as continuative.

⁵⁴ See page 107.

6. A many of our bodies shall no doubt
Find native graves; *upon the which, I trust,*
Shall witness live in brass of this day's work

(*King Henry V*, 4.3.103-5)

This is another borderline case, which again seems to depend for its ascription on the criterion of the novelty and independence of the idea introduced by the relative clause. There seems, however, to be a sufficiently close relationship between antecedent and relative clauses for it to be seen as non-restrictive. The fact that Wells and Taylor (2005) choose to insert a semi-colon between the clauses, and that the First Folio has a colon in the same position (Hinman 1996, 441), suggest support for Rater 2, perhaps.

7. There is no bar
To make against your highness' claim to France
But this, *which they produce from Pharamond,*
In terram Salicam mulieres ne succedant

(*King Henry V*, 1.2.40-43)

In this case, however, the criteria for continuativity are not met at all satisfactorily. “Which” here is not in any sense replaceable by *and* or another coordinating conjunction. If anything, the contest would more plausibly be with restrictiveness, given the presence of the deictic demonstrative pronoun “this” immediately before the relative.

8. My lord, I'll tell you; that self bill is urg'd,
Which in th'eleveth year of the last king's reign
Was like, and had indeed against us pass'd,
But that the scrambling and unquiet time
Did push it out of further question.

(*King Henry V*, 1.1.3-7)

This case seems to be simply a slip on the part of the rater, since there is a tight referential link between “that self” and the relative “which”, so that “which” is clearly defining “bill”, and the clause must be restrictive.

9. Peace be to you,
As I pursue this war, *which shall be then*
Beyond further requiring.

(The Two Noble Kinsmen, 1.3.24-26)

Neither of the raters was confident of an ascription in this case, seeing it as too borderline to determine. The relative clause, which has “peace” as its rather distant antecedent, does not seem to offer sufficient semantic novelty or syntactic independence to warrant ascription as continuative, and seems rather added as a subordinate afterthought than as a new departure. This is borne out by the replacement test, which results in “Peace be to you, as I pursue this war, and it will then be beyond further requiring”. As with clause 5, there is a slight semantic shift caused by this paraphrase.

10. Cousin, I charge you,
Budge not from Athens; we shall be returning
Ere you can end this feast, *of which, I pray you,*
Make no abatement.

(The Two Noble Kinsmen, 1.1.222-5)

Rater 2 saw this as a continuative clause, although it does not seem to satisfactorily fulfil the first criterion for continuativity. Replacement of “which” with *and* causes a similar semantic shift to cases 5 and 9 above. The semantic content of the relative clause is very closely linked to that of the antecedent clause and is, I would argue, a subsidiary afterthought to it.

11. All my services
You have paid home; but that you have vouchsaf'd,
With your crown'd brother and these your contracted
Heirs of your kingdoms, my poor house to visit,
It is a surplus of your grace, *which never*
My life may last to answer.

(The Winter's Tale, 5.3.6-11)

Again, Rater 2 saw this as continuative, and for similar reasons to those given for clause 10 above, I would argue that it cannot be seen as such.

12. One of these two must be necessities,
Which then will speak, that you must change this purpose,
Or I my life.

(The Winter's Tale, 4.4.45-7)

Rater 1 saw this clause as continuative. It is possible that they mistook the sentence section that follows “speak” for part of the relative clause. In fact, the sentence as a whole only makes sense if “these” in line 45 refers cataphorically to that section, to the two propositions, “you must change your purpose” and “I [must change] my life”. In this light it becomes clear that the relative clause is non-restrictive, not least since it interrupts another clause, and thus contravenes part of criterion 2.

To summarise the rest of the findings, according to the initial ascriptions the sample comprised 108 non-restrictives, 52 restrictives, 36 continuatives, 3 interrogatives and 1 non-relative *which* (the anomaly from *The Winter's Tale*, discussed in 4.1). The interrogatives and anomaly were agreed by the raters. In 187 of the total 200 clauses tested there was consensus as to the ascription, representing a 93.5% agreement with the initial ascriptions. 9 of the 12 clauses where divergence occurred were cases of disputed continuativity, and in all of these one of the raters agreed with the initial ascription. All those 9 were debated as being either

non-restrictive or continuative. Only two of the clauses provoking different verdicts are, I would argue, genuine borderline cases, namely numbers 3 and 6. This represents 1% of the sample analysed. Of the 36 clauses initially ascribed as continuatives, and of the 108 ascribed as non-restrictives, 34 (94.4%) and 100 (92.5%) respectively were agreed unanimously. I would argue that, although absolute certainty has not been achieved, these results demonstrate fully the viability of the set of criteria offered above, and therefore the possibility of successfully identifying and counting continuative clauses in Early Modern texts. Figure 3.3 is a graphic portrayal of the relationships between functions discussed in this section⁵⁵.

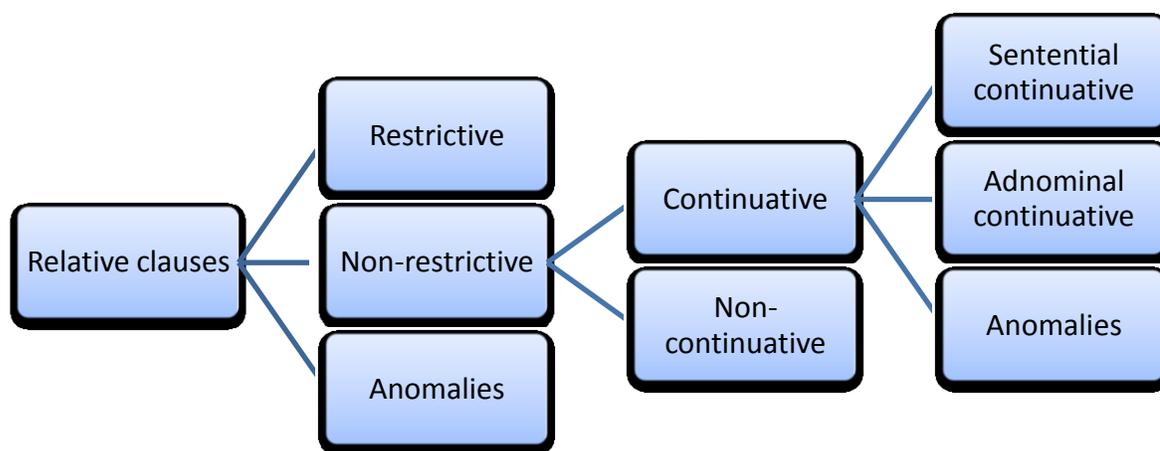


Figure 3.3 Analysis of the information parameter

⁵⁵ Although non-continuative non-restrictive clauses can also be subdivided into adnominal and sentential types, they are not shown in Figure 3.3 because they are not counted separately in the present analysis.

3.3.4 The role of *which* in the relative clause

3.3.4.1 Nominal roles

The focus now shifts to the syntagmatic relationships to the right of the relative, to an analysis of the relative clause itself. This means the grammatical role performed in the clause by *which*, principally and most commonly the question of whether it has a nominal function as subject or object. The work of Keenan and Comrie (1977), referred to above in section 3.2, posits a hierarchy of syntactic positions in the relative clause, usually ordered as subject, object, adverbial, oblique, genitive, comparison. They go on to demonstrate a universal principle of relativisation, whereby “[t]he frequency with which people relativize in discourse conforms to the [Case Hierarchy], subjects being the most frequent, then direct objects, etc.” (Keenan, 1975, 139), and also, significantly for this study, “[a]uthors who are reliably judged to use syntactically simple sentences will present a greater proportion of [relative clauses] near the high end of the CH than authors independently judged to use syntactically complex sentences” (Keenan 1975, 141). Theirs was a synchronic study based on Present Day English, and they “only consider definite restrictive [relative clauses]” (1977, 64). However, as mentioned earlier, the work has been taken up by historical linguists such as Suzanne Romaine (e.g.1980 and 1982) and Xavier Dekeyser (e.g.1984), and applied to Early Modern English, and they have extended it in a number of respects, one of which is the inclusion of non-restrictive relative clauses.

Keenan and Comrie established their categories for the purpose of comparing a large number of different languages, and several different relatives were analysed. In this study the focus is on one language only, and on one relative, so some of the case categories used by Keenan and Comrie are not relevant. Adverbials are typically temporal or locative (*when*,

where, or *wherein*, etc) and the genitive form is *whose*, none of which are at issue here. The oblique case in this study is only relevant to prepositional complement *which*, which is counted elsewhere as a distinct category. There were no examples of the comparison case found in the Shakespeare corpora, which accords with the findings of Dekeyser and Ingels in their set of samples (1988, 28). Indirect object and prepositional complement were also conflated, in accordance with the practices of Dekeyser and Ingels (1988, e.g.28) and Romaine (1982, 149). As a result the categories employed here are subject, object and prepositional complement.

Examples from the Shakespeare corpora follow, firstly with *which* as a subject in the relative clause, shown in an adnominal then a sentential clause:

- (48) O! When degree is shaken,
Which is the ladder to all high designs,
The enterprise is sick.

(Troilus and Cressida, 1.3.104-6)

- (49) And he will smile upon her, *which will now be so unsuitable to her disposition...*

(Twelfth Night, 2.5.91)

The following demonstrate the object position, again with adnominal and sentential antecedents respectively:

- (50) The pound of flesh *which I demand of him*
Is dearly bought

(The Merchant of Venice, 4.1.103-4)

- (51) Our business is not unknown to the senate; they have had inkling this fortnight what we intend to do, *which now we'll show 'em in deeds.*

(Coriolanus, 1.1.28-29)

Occurrences of *which* as a relative determiner, such as (7) and (8) above⁵⁶, were treated as having a nominal role in the relative clause (as either subject or object), despite the presence of the noun that they modify, which recapitulates the antecedent.

3.3.4.2 Which as prepositional complement

The following exemplify occurrences of *which* as prepositional complements, firstly with the preposition in the shifted, then the stranded, position:

- (52) If I be left behind
 A moth of peace, and he go to the war,
 The rites *for which I love him* are bereft me

(*Othello*, 1.3.273-5)

- (53) If you grant not
 My sister her petition in that force,
 With that celerity and nature, *which*
 She makes it in...

(*The Two Noble Kinsmen*, 1.1.199-202)

3.3.4.3 Relativiser-only which

Not all occurrences of *which* have a grammatical role in the relative clause, or function as prepositional complements, however. There is a type of clause that Jespersen describes as follows:

Not infrequently a relative clause, which has been begun in the ordinary way, is continued irregularly as if the power of the relative clause were exhausted.

(Jespersen 1949,107).

⁵⁶ See p. 75.

Blake describes this as a feature of Shakespearean English that differs markedly from modern usage:

In [Present Day English] the relative fulfils two functions: it signals that the following clause is subordinate and it also acts as a grammatical element, whether that is subject, object or whatever, within the subordinate clause. In [Shakespearean English] the relative may fulfil only one of these two functions.

(Blake 2002, 61-2)

He suggests several examples of this phenomenon, the most appropriate of which is this from *As You Like It*, (extended for clarity, numbered, and cited in modernised spelling, for ease of future reference here):

(54) Here we feel but the penalty of Adam,
The seasons' difference: as, the icy fang
And churlish chiding of the winter's wind,
Which, when it bites and blows upon my body
Even till I shrink with cold, I smile

(2.1.7-11)

As Blake explains, in this case “[w]hich merely indicates subordination, because grammatically, the subject of *smile* is I” (2002, 62) and “smile” is intransitive. More typically in the Shakespeare canon, however, this type of *which* usage occurs in a continuative clause, such as the following from Bolingbroke’s speech after his arrival back in England, declaring his intention to demand from

Richard the repeal of his banishment and the restoration of his lands:

- (55) If not, I'll use the advantage of my power,
And lay the summer's dust with showers of blood
Rained from the wounds of slaughter'd Englishmen:
The *which*, how far off from the mind of Bolingbroke
It is, such crimson tempest should bedrench
The fresh green lap of fair King Richard's land,
My stooping duty tenderly shall show.

(*Richard II*, 3.3.47-53)

The relative clause here could be paraphrased as “and my stooping duty will tenderly show how far off from the mind of Bolingbroke it is that such crimson tempest should bedrench the fresh green lap of Richard's land”. As can be seen from the fact that neither “which” nor an antecedent of “which” is required for a satisfactory paraphrase, “which” has no grammatical function in the relative clause. In common with Blake's example from *As You Like It*, the nominal roles are taken by other grammatical items in the clause. “My stooping duty” is the noun phrase subject of the clause, “shall show” is the main verb, and the preceding clauses (from “how far” through to “land”) are the object. The replacement test for continuativity (discussed in 3.3.3.2 above) is fulfilled here, as shown by the fact that “which” is satisfactorily replaced by the conjunction *and*, (although *but* might fit the semantic prosody of the clause better). The crucial fact that no anaphoric non-relative pronoun is required along with it, though, means that this is a case of single relative function, as argued by Blake above. It cannot quite be seen as “merely indicating subordination” as he suggests, though, since the relationship between the clauses is coordination, not subordination. The more general term ‘relativiser-only’ will therefore be used in this analysis to refer to such cases.

Blake also argues (2002, 62) that there are cases in Shakespeare where the relative performs only the other of the two functions, so would not be a relativiser at all, only playing a

nominal role in the relative clause. However, the example he gives is the following, from a dialogue between the Fool and the King in *King Lear*⁵⁷:

- (56) Lear: For by the marks of sovereignty, knowledge and reason, I should be false persuaded I had daughters.
Fool: *Which they, will make an obedient father.*

(*The History of King Lear*, scene 4, 228-9)

According to Blake, “[i]n this exchange the Fool adds to Lear’s sentence by adding what appears to be a relative clause, though *Which* fulfils only a grammatical function as object anticipating *father*” (2002, 62). This seems an unsatisfactory account of the case. It is unclear what Blake means by “anticipating”, since *which* is not operating cataphorically. Surely what we have here is another example of a continuative clause with a relativiser-only, as in (55), discussed above. “They” refers to Lear’s daughters, and “Which” here could be replaced by *and*, without the addition of an anaphoric non-relative pronoun, resulting in “And they will make an obedient father”. No other examples have been found in the corpora which fit Blake’s category of “only a grammatical function”. Moessner argues for a similar reading of the syntax of this example from *King Lear*, using it as a defining case for this type of relativiser-only relative clause:

Relative constructions of this type ... function as postmodifiers of substantival nuclei (here *daughters*), which are usually called their antecedents, and they are introduced by a relative pronoun (here *which*). In contrast to [Present Day English] constructions, the relative pronoun has only **one** syntactic function; it indicates that what follows is to be understood as a relative clause, i.e. it modifies a substantival nucleus.

(Moessner 1992, 336-7)

The double function of *which* and other relatives, found in typical Present Day English usage, and elsewhere in Early Modern English, he calls “functional amalgamation.” And he concurs

⁵⁷ Again, the example is numbered for future reference, although this is a citation from *The History of King Lear*.

with the reading offered by this research, identifying the kind of *which* usage found in the *Lear* example in terms of its lack of functional amalgamation, its lack of grammatical role in the relative clause. The role is instead played by the personal pronoun (in this case “they”):

In ordinary [Present Day English] constructions, one of the functions of the relative pronoun is that of a nominal constituent of the relative clause, e.g. subject, object ... As there is no functional amalgamation in EModE relative constructions of [the type in question], it is the personal pronoun pronominalizing the antecedent which realizes the function of a nominal constituent of the relative clause (here: the subject function). (1992, 337)

Although Romaine is primarily concerned with the development of relatives in Middle Scots in a socio-historic context, and although she does not for some reason distinguish coordination from subordination in relative clause relationships, she describes in some detail the phenomenon of relativiser-only *which* usage. She contrasts a Present Day English example, where *which* is performing its usual double function (Moessner’s functional amalgamation), with a relativiser-only occurrence:

In this [Present Day English] case, *which* appears to function in part as a conjunction (i.e. as a subordinator and means of linkage between two propositions), and also in part as a pronoun in relating parts/constituents of two sentences ... When the pronominal function is not in evidence by virtue of the syntax of the *which* clause (i.e. when *which* does not have an NP role in the clause it introduces), one can assume that the pronominal function has become opaque or lost and it is being used primarily as a conjunction.

(Romaine 1982, 86).

Some of Shakespeare’s relative clauses seem almost to defy syntactic analysis, but begin to fall into place better when we acknowledge the relativiser-only category. This case from *The Winter’s Tale* is a case in point:

- (57) Is't not the tenor of his Oracle,
That King Leontes shall not have an heir,
Till his lost child be found? *which that it shall,*
Is all as monstrous to our human reason
As my Antigonus to break his grave
And come again to me; who, on my life,
Did perish with the infant.

(5.1.47-53)

The first section of the *which* clause in line 49 here can seem baffling to a modern reader, and we struggle to make sense of it by attributing a function to “which” in the relative clause. However, once we accept that “which” is only introducing the relative clause and not participating in it syntactically, both the syntax and the sense become clear.

3.3.4.4 Cataphoric *which*

There are two closely related types of *which* function in the relative clause in Shakespeare's canon which are not accounted for by the foregoing descriptions. Neither of them is common, but they both require the establishment of additional categories. Firstly there are cases such as these:

- (58) The chief perfections of that lovely dame -
Had I sufficient skill to utter them -
Would make a volume of enticing lines,
Able to ravish any dull conceit;
And, *which is more*, she is not so divine

(*I Henry VI*, 5.7.14-18)

- (59) I saw him arrested; saw him carried away; and, *which is more*, within these
three days his head to be chopped off.

(*Measure for Measure*, 1.2.37-38)

In these occurrences, *which* seems to function as part of a fixed collocation, (perhaps synonymous with *moreover* in Present Day English), and does not perform an anaphoric role with regard to a foregoing clause. There is an element of cataphoric reference, fulfilled by the main clause that follows, although there is only minimal semantic content in that cataphoric link. Norman Blake suggests that “[w]hich may appear as a kind of indefinite pronoun especially in the phrase *which is more*” (2002, 63), although Quirk (1985, 634) categorises them as reinforcing conjuncts.

There are also instances of cataphoric *which* where the relative has more semantic substance than in the case of *which is more*, and a stronger cataphoric link with what follows, as can be seen in the following examples:

- (60) Fortune is blind: and she is painted also with a wheel, to signify to you, *which is the moral of it*, that she is turning and inconstant

(*Henry V*, 3.6.21-2)

- (61) He gave you all the duties of a man,
Trimmed up your praises with a princely tongue ...
And, *which became him like a prince indeed*,
He made a blushing cital of himself

(*Henry IV Part II*, 5.2.60-66)

Both types are included here in the category ‘cataphoric *which*.’

3.3.4.5 Pushdown constructions

Next there are cases where the role of *which* in the relative clause is less straightforward, where, although relative *which* does have a grammatical role in the clause, it is not simply subject, object or prepositional complement of just the main part of the clause, as is the norm. Reuter’s category comparable to this is what he calls “relative concatenation”

(e.g. 1938, 46) although he includes in the category cases where an additional clause is simply embedded in the relative clause, without the relative having a role in it, as can be seen from the example he offers from the Wycliffe Bible (italics added):

Go ye, telle ye to my britheren, that thei go in to Galilee; there thei schulen se me. The whiche *whanne thei hadden gon*, loo! summe of the keperis camen (1938, 46).

Which here is in fact a relativiser-only, as discussed above, since neither the main relative clause nor the embedded clause engages the relative “which” in a grammatical role. It is the presence of the embedded temporal clause “whanne thei hadden gon” that causes Reuter to label it a case of concatenation, or linking of clauses. However, because this analysis focuses principally on the functions of *which*, and therefore seeks categories that clarify such functions, cases of simple embedding are dealt with in the next section, in the category of phrases and clauses intervening to the right, between *which* and the subject or main verb of the clause. Another reason for rejecting Reuter’s organisational principles on this occasion is his prioritisation of semantic over syntactic definitions of “concatenated” relative clauses. He subdivides “relative concatenation” cases into “temporal clauses”, “conditional clauses”, “concessive clauses”, “clauses of finality”, “clauses of causality” and “clauses of modality” (1938, 46-47), whereas other writers, in particular Lilo Moessner, offer a syntactic approach that allows greater precision in the present analysis.

Moessner focuses on producing a tightly articulated taxonomy of complex relative constructions, in order to argue against the “hypothesis of Latin loan syntax” (Moessner 1992, 336). Although the historical provenance of relatives is not the concern of this thesis, the clarity and syntactic precision of Moessner’s analysis provides a very useful framework for understanding Shakespeare’s relative clauses. Of the five types established by Moessner, there

are two in particular which are critically important here.⁵⁸ Moessner identifies complex relative clauses as possessing a “lower” and a “higher element”, with the main clause of the relative clause seen as the higher, and the lower element being an additional, subordinate clause. Citing Quirk for his definition,⁵⁹ Moessner articulates the first of these two types thus: “The common feature of all these constructions is that the [relative] pronoun realizes the function of a lower-level constituent. Such elements which are “embedded further down in the constituent structure of the sentence” are called “push-down elements”” (1992, 340). He specifies that the type he identifies should be known as “relative pushdown elements”, presumably to distinguish them from other, non-relative, syntactic phenomena discussed by Quirk, although that distinction will not be necessary in the present thesis, and they will be termed simply ‘pushdown constructions’.

Moessner’s discussion is seminal for the present thesis, although on the following occasion there seems to be a mismatch between the example he uses and the category with which he associates it. The example that Moessner gives of the type where the relative realises the function in only a lower element in the relative clause is this from Shakespeare’s *The Tempest*⁶⁰:

- (62) For you, most wicked sir, *whom to call brother*
Would even infect my mouth
- (5.1.140-41)

The non-finite clause “to call brother” is the lower, pushdown element here, and “would almost infect my mouth” the higher. The relative, in this case “whom”, has a grammatical role

⁵⁸ Moessner’s other three types are reflected in features already discussed in this section.

⁵⁹ Quirk et al 1972, *Grammar of Contemporary English*. Harlow: Longman.

⁶⁰ As earlier, the example is included in the numbered sequence for ease of future reference.

in the lower element, and not, on its own, in the higher, although the whole lower element then becomes the subject of the higher. Moessner's cites it, without further comment, as an example of Type IV. Yet the non-finite, lower element of the relative clause does have a nominal role (in this case as subject) in the higher element. In Moessner's Type IV constructions this is not the case. Rather, this is an example of what Moessner identifies as Type V. In this thesis Moessner's Types IV and V will be known as single and double pushdown constructions respectively. The following is a typical case of a single pushdown construction. The higher element of the clause has nominal components separate from the lower element, as in example (44) from *Twelfth Night*, below:

Do not embrace me till each circumstance
Of place, time, fortune, do cohere and jump
That I am Viola; *which to confirm,*
I'll bring you to a captain in this town

(*Twelfth Night*, 5.1.229-32)

The relative has a grammatical role only in the lower element, in this case as object of the non-finite clause "which to confirm". The same analysis applies even to cases where *which* is a prepositional complement, such as this from *The Two Gentlemen of Verona*:

(63) My riches are these poor habiliments,
Of which if you should here disfurnish me,
You take the sum and substance that I have.

(4.1.16-18)

The role of "which" here is as a prepositional complement in the conditional lower element clause "of which if you should here disfurnish me", and again it has no role in the higher element.

Double pushdown constructions, on the other hand, are more syntactically complex than single pushdowns.

- (64) Thence it came
That she whom all men prais'd, and *whom myself*,
Since I have lost, have lov'd

(*All's Well That Ends Well*, 5.3.64-66)

As explained earlier, in Present Day English, and in the majority of Early Modern English relative clauses, relatives such as *which* perform two functions, in what Moessner and others call “functional amalgamation” (e.g. 1992, 337). One exception in Early Modern English, as we have seen, is the relativiser-only construction, where the relative performs only one function. The other exception is this, Moessner’s final type, which he exemplifies with the case from *All's Well That Ends Well* above and explains thus:

In this type, it has even more than two functions. The relative pronoun *whom* [in the example from *All's Well*] indicates that what follows is a relative clause, and in this relative clause it realizes the object function in the main and in the subordinate clause” (1992, 341).

In these constructions the relative has a role in two clauses, both the lower and the higher elements.

- (65) It is true, and I will give you comfort,
To give your dead lords graves, the *which to do*
Must make some work with Creon

(*The Two Noble Kinsmen*, 1.1.147-9)

“Which” here is both the object of the non-finite lower element “which to do”, and at the same time the subject of the higher element “must make some work with Creon”. Since the double pushdown construction seems to have a particular potential for complexity or opacity, it is counted separately from the single type in this analysis.

3.3.4.6 Recursions

There is another kind of occurrence in the canon, presenting a different role for *which*, and one that is also notably complex syntactically. In these cases, *which* is part of a rank-shifted clause, and the whole clause then functions as a nominal component of the top level relative clause, as can be seen in this example from *Richard III*:

- (66) The right and fortune of his happy stars;
Which God defend that I should wring from him!

(3.7.177-8)

Here, “which” is the object of the subordinate “that” clause, which itself is embedded inside the main relative clause, where “God” is the subject and “defend” the predicate. A comparable construction occurs in this occurrence from *King Lear*:

- (67) Our father he hath writ, so hath our sister,
Of differences, *which I best thought it fit*
To answer from our home;

(2.1.132-4)

In this case it may even be possible to construe the role of “which” as being in a recursion within a recursion, a double rank-shifting. “I” is the subject of the relative clause, ‘best’ is a straightforward phrasal intervention, and then the subordinate clause following the predicate “thought” behaves exactly like a *that* clause, reproducible as “[that] it [was] fit to answer [which] from our home”, in which case “which” is the object of the non-finite clause “to answer from our home”. Most cases of this type result from subordinate *that* clauses after verbs of mental process (such as “thought” in the example from *King Lear* above), and the majority are clear enough to identify.

A problem arises for categorisation, however, with those where complementiser *that* has been deleted⁶¹, and the top level subject and predicate (the mental process verb) can be confused with an interjection. Shakespeare uses interjections frequently, such as this in *Much Ado About Nothing*:

(68) You break jests as braggarts do their swords, which, *God be thanked*, hurt not.

(5.1.154)

In this case it seems clear that the interjection is just that, and not the subject and predicate of the main clause, precipitating a clausal object⁶². But on occasions like the following, the matter is rather different:

(69) I only have made a mouth of his eye, by adding a tongue *which I know will not lie*.

(*Love's Labour's Lost*, 2.1.261)

In this case and others like it, the punctuation could turn the syntactic relationships either way. If commas were added on either side of “I know” it would be turned into an interjection. If not, ‘that [which] will not lie’ becomes a subordinate nominal clause, (the object of the top level clause “I know [...]”), and “which” is its subject. Interestingly, although Craig is normally zealous in his addition of commas, he refrains in this instance, supporting the case for its being a recursion. Even modern editors have a habit of preferring to add commas to this kind of construction, and in this case from *Measure for Measure* that is what usually happens:

(70) and now she professes a hot-house, which, I think, is a very ill house too.

(2.1.57)

⁶¹ See, for example, Romaine (1980, 221), where this practice is discussed with regard to syntactic complexity.

⁶² Such cases are categorised as finite clauses intervening to the right of *which*, in 3.3.5.2 and the corresponding sections in subsequent chapters.

Craig inserts these two commas around “I think”, but in the First Folio it appears as “which, I thinke is a very ill house too” (Hinman 1996, 83), suggesting the recursive syntax rather than the interjection. The provenance of the punctuation in the First Folio is variable and unreliable, perhaps no more ‘authentic’ than Craig’s interventions. But this is not a one-off case. In Craig’s edition of *Coriolanus* we find the following:

(71) Nor did you think it folly
To keep your great pretences veil’d till when
They needs must show themselves, which in the hatching,
It seem’d, appear’d to Rome.

(1.2.24-27)

Wells and Taylor (2005) and Evans (1997) both follow Craig in adding two commas either side of “It seem’d”, but there are none at all in the First Folio. In a later scene of the same play we see yet another case, with a similar editorial history. This is Craig’s punctuation:

(72) Thou art their soldier, and being bred in broils
Hast not the soft way *which, thou dost confess,*
Were fit for thee to use, as they to claim

(3.2.101-3)

In the First Folio the same passage appears as:

Thou art their Souldier, and being bred in broyles,
Hast not the soft way, which thou do’st confesse
Were fit for thee to use, as they to clayme

(Hinman 1996, 634)

It seems likely that modern editors opt for the clarifying commas as a means of simplifying Shakespeare’s syntax, since the interjection is the more accessible modern alternative. But given the fact that this research is focusing on identifying syntactic complexity, and given the repeated omission of commas in the First Folio, it was felt that where there was a genuine

dilemma between the two versions, the recursive construction would be assigned. If such cases were read as the modern editors punctuate them, they would of course still be examples of complex syntax, having a finite clause intervening,⁶³ but recursions are more complex still, so it is important to count them separately where they occur.

Figure 3.4 below shows an overview of the relationships between the analytical categories presented for *which* function in this section.

⁶³ See 3.3.5.2.

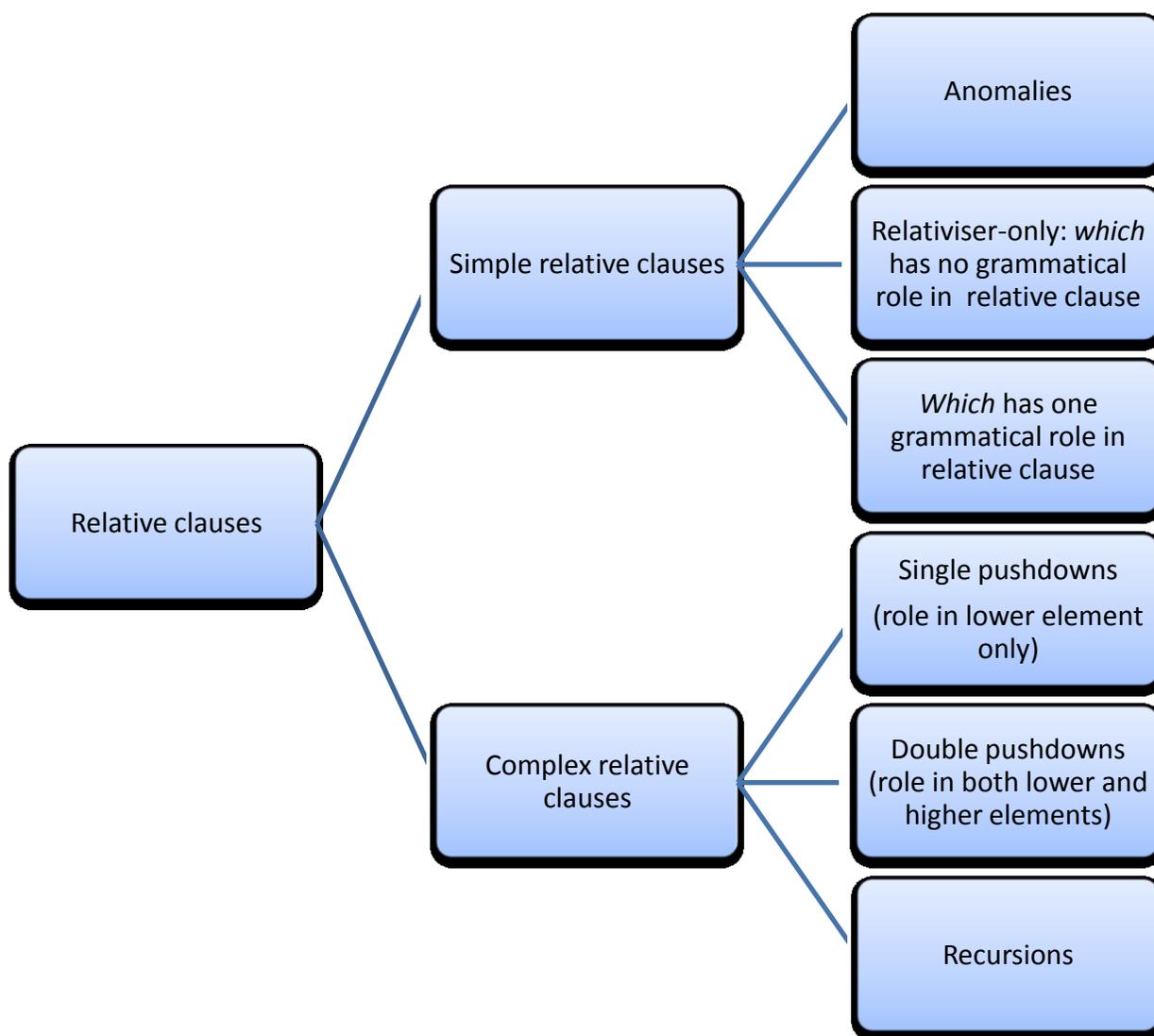


Figure 3.4. Analysis of *which* function in the relative clause

3.3.5 Interventions to the right of *which*

3.3.5.1 The number of words intervening to the right

The interrelationships between the categories described in this section are complex, so they are clarified further in the following chapter, in section 4.5, alongside the statistical data for this set of features.

Ryden identifies the “distance between the relative (as subject) and the finite verb in the relative clause” as being of “distributional significance” in early sixteenth century relative clauses (1966, xx). It therefore seems likely that this will be a worthwhile criterion to adopt in the present analysis, especially since an increase in this distance may perhaps contribute to a sense of complexity or opacity in the relative clause in a comparable way to the distance between antecedent and relative. The definition of such interventions employed in this analysis is the linguistic material intervening between *which* and the next principal component of the relative clause to the right, where ‘principal component’ means subject, object or main verb. It should be stressed that interventions to the right of *which* are not the norm. A sentence is already complex if it has a relative clause appended to it. In the cases counted in this section, however, Shakespeare has added further syntactic complexity, with even more syntactic boundaries involved.

In general the same counting principle was employed here as in 3.3.4.1 above, including the discounting of single function words (such as *and*, *if*, *to*, etc.⁶⁴) where interventions to the left of *which* were involved. However one change was made to the criteria for identifying interventions to the right. Where, in the analysis of antecedent distances, the noun phrase of the antecedent was reduced to the head word for counting purposes, in this part of the analysis it was felt that this was inappropriate, since in English (Early Modern as Present Day) the bulk of noun modification occurs to the left of the head word. This meant that most of the modification of the antecedent was automatically excluded from the analysis of interventions to the left of *which*, since *which* always comes to the right of the antecedent noun phrase. Now, however, with the focus on interventions to the right of *which*, a different

⁶⁴ See 3.3.2.4.

situation pertains. Although, as Ryden implies above, the typical syntax of the relative clause finds *which* functioning as subject, followed by the main verb, there are of course numerous occasions in the Shakespeare canon where this is not the case, and where the next principal item in the clause is either subject or object. In such instances, modification for that noun phrase would appear in the area under scrutiny, to the left of the subject or object, between *which* and the head word. It was felt that syntactic complexity would be more likely to be successfully addressed by including only interventions that are not part of the modification components of the subject or object, and so modification elements for the noun phrase to the left were not included in the count. Examples of this may clarify the point:

(73) the ceremonious vows of love
 And ample interchange of sweet discourse,
 Which so long sunder'd friends should dwell upon

(*Richard III*, 5.3.109-11)

Here, the next principal component to the right is the the subject of the relative clause, the noun phrase “so long sunder’d friends”. If the headword were taken as the point up to which counting should happen, the modifying phrase “so long sunder’d” would count as an intervention. The decision has been taken here to include such modification as part of the noun phrase to which it belongs, so in this case there is no intervention. In the same way, in cases of determiner *which*, where a noun follows immediately to the right of the relative, that noun is not counted as an intervention.

(74) And so, good Capulet, *which name* I tender
 As dearly as my own, be satisfied

(*Romeo and Juliet*, 3.1.45-6)

In this example, where “which” is the object of the relative clause, and the subject “I” is the next principal component to the right, the noun “name” is not counted as an intervention, since it ‘belongs’ to determiner “which”, as part of the object.

Since relativiser-only *which* has no syntactic relationship with the relative clause, and therefore the items intervening would have no potential for disruption of the processing of the syntax, these cases were not included in the count. An example from the Epilogue of *Henry IV Part 2* demonstrates this point:

- (75) I was lately here in the end of a displeasing play, to pray your patience for it, and to promise you a better. I meant indeed to pay you with this; *which if like an ill venture it come unluckily home, I break, and you, my gentle creditors, lose.*

(Epilogue, 8-13)

The conditional clause “if, like an ill venture, it come unluckily home” comes between “which” and “I break”, suggesting perhaps at first glance that “which” functions as an object in the relative clause. Semantically, however, this of course makes no sense. “[W]hich” here has no grammatical role in the relative clause: the subject is “I”, and the main, intransitive verb is “break”. There is a subsequent coordinated clause (“and you, my gentle creditors, lose”) that has its own subject and main verb. “[W]hich” is simply a signal that a (loosely appended, continuative) relative clause is to follow, and is replaceable by *and* alone, with no anaphoric pronoun required, because the relative clause has its own full set of grammatical components. There is therefore no sense in which the conditional clause postpones the syntactic realisation of the relative clause. In fact, the conditional clause is semantically essential in the position in which it occurs, as can be seen if the replacement is made: “and if like an ill venture it come unluckily home, I break, and you, my gentle creditors, lose”. This

case is typical of relativiser-only occurrences, and demonstrates why clauses that seem at first glance to intervene syntactically do not.

Recursions are not counted for interventions to the right, since the rankshifted syntax would make the gauging of the next principal component extremely convoluted. As elsewhere, pushdown constructions present some additional challenges for counting of interventions to the right, requiring clarification of procedures employed. Double pushdowns are included, and (on the whole) it is the lower element that forms the intervention to the right, as can be seen here:

- (76) Scars and bare weeds
 The gain o' the martialist who did propound
 To his bold ends honour and golden ingots,
 Which though he won, he had not.

(The Two Noble Kinsmen, 1.2.17-20)

In this fairly typical case, “though he won” is construed as the intervention between “which” and the next principal component to the right, namely “he”, the subject of the (higher element of the) relative clause.

Single pushdown constructions, on the other hand, can manifest interventions, but only inside the lower element of the relative clause, because *which* has no syntactic relations outside that lower element, and there would be no disruption or complication of the syntax of *which*. This means that in the most typical kind of single pushdown construction, there is no intervention counted, as can be seen in the following case:

- (77) Call him a slanderous coward and a villain,
 Which to maintain I would allow him odds

(Richard II, 1.1.65- 6)

Here, “Which to maintain” is the lower element of the relative clause, where “which” is related only to the non-finite verb phrase “to maintain”, so there is no intervention. More problematic cases are considered on an individual basis in the next chapter, alongside the statistical results.

3.3.5.2 Word groups intervening to the right

As was the case in the analysis of interventions to the left, the issue of phrasal or clausal material intervening between *which* and the next principal element of the clause pertains to a different kind of syntactic complexity from that of distance. In the case of distance, we assume a direct relationship between the number of words intervening and the syntactic complexity that results. On the other hand, it is possible for a substantial intervention to be made in one or two words, as can perhaps be seen in this from the first scene of *Romeo and Juliet*:

(78) What sadness lengthens Romeo’s hours?

Not having that, *which having, makes them short.*

(1.1.151-2)

“[H]aving” is an intervention of just one word, to the right of *which* in this case, (the lower, non-finite element of the double pushdown construction), but it adds considerable density and complexity, both syntactically and semantically, principally because it is a verbal element rather than a phrase. In common with 3.3.2.4 above, this part of the analysis is concerned with the quality, rather than quantity, of the words intervening. As earlier, ‘clause’ refers to both verb phrases and fully fledged clauses, and is only distinguished on the basis of finiteness or non-finiteness. On this occasion, though, there is no need for the separate category of “partial

finite clause”, since that was established simply in order to identify a specific linguistic and stylistic phenomenon to the left of *which*. *Which* with object function in the clause is included, in addition to the subject function mentioned by Ryden above⁶⁵, but where *which* is a prepositional complement it is not possible to pin down interventions. The syntactic relationships are much more fluid in such cases, and the position of the prepositional complement in the clause much more variable. The feature was measured in syntactic terms, by counting separately whole intervening phrases, multiple phrases, and non-finite and finite clauses, in much the same way, and for similar reasons, as in 3.3.2.4 above. Cases were counted separately where a combination of phrase and single clause elements intervenes, and where more than one clausal element intervenes. Again as in 3.3.4.2, multiple intervening clauses with additional phrases were not counted separately, since they are very rare, and the multiple clause intervention was seen to be more important.

Phrases intervene the most frequently in the relative clause, as in the opening speech of *Antony and Cleopatra*:

(79) his captain’s heart,
 Which *in the scuffles of great fights* hath burst
 The buckles on his breast

(1.1.8-10)

The phrase “in the scuffles of great fights”⁶⁶ intervenes here between the relative as subject, and the main verb “hath burst”. The term ‘phrase’ is understood here in the same way as 3.3.2.4. vi) above, with the same criteria used in deciding on single or multiple assignment.

⁶⁵ 1966, xx.

⁶⁶ A ‘top level’ only approach is used in this category, as elsewhere.

The following example from *Twelfth Night* is a case of multiple phrases intervening, in this case three, namely “to a stranger”, “unguided and unfriended”, and “often”:

(80) what might befall your travel,
Being skillless in these parts; which *to a stranger*,
Unguided and unfriended, often prove
Rough and inhospitable

(3.3.10-13)

“[I]n the scuffles of great fights” (79) might seem to present an argument for being two prepositional phrases, but the clause would not make sense with one or other phrase removed. On the other hand, in (80), both “unguided and unfriended” and “often” could arguably be removed and still leave a semantically satisfying clause behind, so these are seen as three intervening phrases.

In a comparable way, the following occurrence from *Othello* is only counted as having one finite clause intervening:

(81) Tis yet to know,
Which when I know that boasting is an honour,
I shall promulgate

(1.2.23-5)

The intervention comes between “which” (as object of the relative clause) and the next component, the subject “I” (in “I shall promulgate”). There are two clauses in the intervention, but neither “when I know that” or “boasting is an honour” can be left behind without the other, so this is a case of a single clause intervening.

As seen above in the case of (80), single words can also be counted as phrases, and adverbs in particular are often found intervening between *which* and the next component in the relative clause. Further examples include the following:

- (82) To bear the golden yoke of sov'reignty
Which *fondly* you would here impose on me

(*Richard III*, 3.7.151-2)

- (83) Tamora, the empress of my soul,
Which *never* hopes more heaven than rests in thee.

(*Titus Andronicus*, 2.3.43-44)

In cases of mixed intervention, where there is a clause and a phrase (or more than one phrase), the same criteria are applied as above, namely that if the phrase is separable in this way from the intervening clause, it is counted as a clause and a phrase intervening. In the following example, the finite clause and adverbial phrase can clearly be separated:

- (84) Be satisfied, dear God, with our true blood,
Which, *as thou know'st, unjustly* must be spilt.

(*King Richard II*, 3.2.23-4)

On other occasions, however, they are too closely related to be separated:

- (85) 'tis your passion
That thus mistakes, the which, *to you being enemy*
Cannot to me be kind

(*The Two Noble Kinsmen*, 3.1.49-51)

Here, the prepositional phrase “to you” is too closely related to “being enemy” to be separated, and so this is only counted as one clause intervening.

Cases of more than one clause intervening are rare, and highly complex, and when they do occur they are usually double pushdown constructions, like this one from *The Winter's Tale*:

- (86) I loved him as in honour he required,
With such a kind of love as might become
A lady like me...
Which *not to have done, I think* had been in me
Both disobedience and ingratitude

(3.2.60-65)

3.3.6 Clausal length

The final category established for this analysis is clausal length. Dekeyser and Ingels see it as a “measure of syntactic complexity” (1988, 33), an observation derived from comparison of usage of relatives in different registers,⁶⁷ so it was seen as worth including in the present study. They do not specify their criteria for determining what a relative clause is, but the term needs clarification in the present analysis in order for counting to be possible. The aim here is to employ a ‘top level’ approach to clause measurement, in accordance with earlier sections. Clauses become longer and more complex (the issue at stake here) by the inclusion of additional phrasal or clausal elements. If we dissect them to their ‘lowest level’, we risk losing sight of their potential complexity. There are, however, limits to how inclusive one can be. Examples will demonstrate some of the issues involved. The case of occurrence (35), for instance, is straightforward. There is no boundary problem, since the two consecutive relative clauses (italicised) are distinct, in a coordinated relationship, and syntactically linked only by the antecedent they have in common:

⁶⁷ Dekeyser and Ingels allude to Quirk, R. (1968) “Relative clauses in educated spoken English”, in *Essays on the English Language, Medieval and Modern*. Longmans’ Linguistic Library. London: Longmans, pp. 94-108.

I have given him that
Which, if he take, shall quite unpeople her
Of leigers for her sweet, and which she after,
Except she bend her humour, shall be assured
To taste of too."

(*Cymbeline*, 1.5.91-5)

These are, uncontroversially, two separate *which* clauses. Similarly, there is no problem in cases such as this, also from *Cymbeline*, which typifies those clauses whose immediately succeeding elements relate back to the antecedent clause:

(87) and shall make your lord
That *which he is*, new o'er

(1.6.188-9)

Only the italicised words comprise the relative clause here, because "new o'er" is an adjunct qualifying the foregoing clause "shall make your lord that", not "which he is".

Some cases require careful attention from a semantic perspective, such as this relative clause from *Coriolanus*:

(88) To brag unto them, thus I did, and thus;
Show them the unaching scars *which I should hide*,
As if I had received them for the hire
Of their breath only!

(2.2.146-9)

The adverbial clause of the final two lines here could, at first glance, be thought to qualify "hide", since that is the verb immediately preceding it, and in which case it would need to be included in the count. Closer reading reveals that it qualifies "show", and so the adverbial clause is not related to the relative clause, and the relative clause ends at "hide".

Greater problems for definition of the boundaries of relative clauses are caused by cases where one relative clause is followed by another which qualifies part of the first, as in:

- (89) Be not angry,
Most mighty princess, that I have adventured
To try your taking of a false report, *which hath*
Honoured with confirmation your great judgement
In the election of a sir so rare,
Which you know cannot err.

(*Cymbeline*, 1.6.196-201)

The question here is whether the non-restrictive relative clause “which you know cannot err” forms part of the relative clause beginning “which hath honoured”. And in this next case, if the non-restrictive clause interrupts the first relative clause, would it be appropriate to remove the non-restrictive clause from the continuative clause in which it is embedded, when counting the length of the continuative clause?

- (90) They confess
Toward thee forgetfulness too general, gross;
Which now the public body, *which doth seldom*
Play the recanter, feeling in itself
A lack of Timon’s aid, hath sense withal
Of its own fail

(*Timon of Athens*, 5.1.136-41)

If the second clause were restrictive, as in the following case from *Richard III*, the second relative clause (here “which you have pill’d from me”) would surely be seen as integral to the first, making the whole italicised section one unit:

- (91) Hear me, you wrangling pirates, *that fall out*
In sharing that which you have pill’d from me!

(1.3.163-4)

At the other end of the spectrum, it might seem unreasonable to include continuative clauses in this way, since the syntactic link between clauses in such cases is much weaker, and the continuative relative clause is characterized by a certain autonomy.

Clearly, a formula is needed in order to make possible the counting of occurrences in the corpora. Given that this thesis argues for continuative clauses to be defined as coordinating, not subordinating⁶⁸, it seems sensible to suggest that, for the purposes of the clausal length category, the entity “relative clause” is the main relative clause, plus any subordinate to it that are either restrictive or non-restrictive, but not continuative, and plus any additional phrases that function as adjuncts to it. Junctions characterised by the coordination of what are essentially two main clauses (as is the case with continuatives) are deemed to represent the point where separate syntactic entities meet. Succeeding clauses that share a nominal feature with the preceding relative clause are deemed to be included, as in this case from *Romeo and Juliet*:

(92) True, I talk of dreams,
 Which are the children of an idle brain,
 Begot of nothing but vain fantasy;
 Which is as thin of substance as the air,
 And more inconstant than the wind, who woos
 Even now the frozen bosom of the north,
 And, being anger'd, puffs away from thence,
 Turning his face to the dew-dropping south.

(1.4.105-112)

The initial relative clause here, comprising lines 106 and 107, has two further adnominal relative clauses dependent on it (starting with “[w]hich” in line 108, and “who” in line 109), and then also a verb phrase (“puffs”) in the final two lines, coordinated with “woos” in line 109, and sharing a subject with the “who” clause. This approach means that in (89) above the italicised section is counted as one clause, because the second “which” clause is non restrictive and subordinate. The second “which” clause in (89) is also counted as an occurrence in its

⁶⁸ See 3.3.3.2.

own right. So the first “which” clause is 21 words long, and the second 5. That is the policy followed in the measuring of clausal length in this analysis.

The following case from *Henry V* shows another implementation of the policy, with regard to conjunctions:

- (93) They would have me as familiar with men’s pockets as their gloves or their handkerchers: *which makes much against my manhood if I should take from another’s pocket to put into mine; for it is plain pocketing up of wrongs.*

(3.2.18-19)

The clause to be counted is the italicised section, including the subordinated “if” and “for” clauses⁶⁹. This next occurrence demonstrates a coordinated non-finite clause (italicised) following the main part of the relative clause, which is also included in the count because it shares clausal components with the preceding section (namely “making that idiot, laughter”):

- (94) Or if that surly spirit, melancholy,
Had baked thy blood and made it heavy thick,
Which else runs tickling up and down the veins,
Making that idiot, laughter, keep men’s eyes
*And strain their cheeks to idle merriment,
A passion hateful to my purposes.*

(*King John*, 3.3.45-50)

In this next case, however, the italicised clause following the *which* clause is excluded from the count, because it is a main clause in its own right, coordinated with what precedes it:

- (95) or say to them,
Thou art their soldier, and being bred in broils,
Hast not the soft way which thou dost confess
Were fit for thee to use as they to claim,
In asking their good loves; *but thou wilt frame*

⁶⁹ *For* is understood in this case as a subordinating conjunction, based on Quirk’s comment that it “can reasonably be classed as a peripheral subordinator”, despite its medial presence on the “scale relating coordinating and subordinating conjunctions” (1985, 90). However, in this study each occurrence of the word in the canon is judged separately, according to its semantic effect.

*Thyself forsooth hereafter theirs, so far
As thou hast power and person.*

(Coriolanus, 3.2.100-106)

Questions may also arise at times as a result of other kinds of additional elements following the relative clause.

- (96) Sir, I invite your highness and your train
To my poor cell, where you shall take your rest
For this one night; *which - part of it - I'll waste*
With such discourse as, I doubt not, shall make it
Go quick away; the story of my life
And the particular accidents gone by
Since I came to this isle. And in the morn...

(The Tempest, 5.1. 334-10)

In cases such as this, the two noun phrases “the story of my life” and “the particular accidents gone by since I came to this isle” are viewed as included since they are syntactically, if asyndetically, related to the relative clause, not the foregoing antecedent clause. They function as coordinated noun phrases, semantic parallels to “such discourse” in line 337. It is not until “And in the morn” that a new syntactic entity truly begins. So in this case, the relative clause is counted as having 37 words.

Relativiser-only clauses also cause difficulties, as can be seen from example (90) again, extended here for clarity:

They confess
 Toward thee forgetfulness too general, gross;
 Which now the public body, which doth seldom
 Play the recanter, feeling in itself
 A lack of Timon's aid, hath sense withal
 Of its own fail, restraining aid to Timon,
 And send forth us, to make their sorrow'd render,
 Together with a recompense more fruitful
 Than their offence can weigh down by the dram;

(*Timon of Athens*, 5.1.136-41)

The problem is that relativiser-only *which* (here the first “which” in line 138) functions only as a coordinating conjunction, as is the norm when it starts a continuative clause. When the relative has a linking function as weak as this, with a minimal antecedent relationship and no grammatical role in the relative clause it heralds, it is difficult to identify anywhere a relative clause per se to which it belongs. In order to resolve the matter, to enable the count to be made in cases such as this, the main clause following the *which* in question is designated as the relative clause. So here, “the public body” is the subject of the relative clause, “hath sense” is the predicate or main verb phrase. The clause ends at line 141 because “[a]nd send forth us” is not a main clause, since it shares a subject with the preceding relative clause.

In the case of both single and double pushdown constructions, the whole construction is counted, since the lower and higher elements are both components of the relative clause. Where *which* is a prepositional complement, and the preposition in question precedes the relative, the preposition has been included in the count. So, the following clause is counted as having seven words in it.

(97) It is enough my hearing shall be punished
 With what shall happen, *gainst the which there is*
No deafing

(*The Two Noble Kinsmen*, 5,5,7-9)

Occasionally, the relative clause extends further to the left beyond *which*, again because of a prepositional complement construction, as example (43) above, from *The Winter's Tale*, repeated here:

I am not prone to weeping, as our sex
Commonly are; *the want of which vain dew*
Perchance shall dry your pities;

(*The Winter's Tale*, 2.1.133-5)

In this case, as can be seen from the italicisation, the relative clause extends leftwards to include “the want of which” and so is counted as having eleven words in it. The same principle is followed when the relative is *the which*. Hyphenated words (such as “weather-fends” (*The Tempest*, 5.1.15) have been treated as one word, not two, throughout this analysis⁷⁰. Contractions are, as pointed out in section 3.1, of unreliable provenance, and have been counted as single words in this category, but contractions of a pronoun or noun and a verb (such as *I'll*) are counted as two words throughout, because they often conflate two major elements of a clause, so need to be isolated for analytical purposes.

Finally, it is perhaps inevitable that occasionally in Shakespeare's dramatic language there are interrupted clauses such as the following:

(98) A book of prayers on their pillow lay;
'Which once', quoth Forrest, 'almost changed my mind;
But O! the devil' – there the villain stopped

(*Richard III*, 4.3.16-18)

⁷⁰ Editorial variation makes hyphenation an unreliable feature, but for the sake of consistency in the count, this analysis accepts the hyphenation as presented in the electronic text used.

In this instance the relative clause is taken as the whole line, “‘Which once’, quoth Forrest, ‘almost changed my mind;’ (stopping before the conjunction “but”), since, for most purposes, the speech tag “quoth Forrest” is not essentially different from an embedded clause.

CHAPTER 4

RESULTS AND DISCUSSION

4.0 General remarks

This chapter presents the results of the analysis of all the *which* occurrences in the Shakespeare canon, comparing those in the group of six late plays with those in the preceding thirty-two. As explained earlier,⁷¹ we are not now concerned with the relative frequency of *which* occurrence in the corpora: rather we are mapping developments in Shakespeare's syntactic use of *which*. Many of the syntactic features identified here are multi-word occurrences, so do not lend themselves to calculation in terms of proportion of total tokens or rate of use relative to total tokens. We already know that *which* relativisation increases in the late plays. But what we do not yet know is, within that body of relative *which* clauses, how (relatively) often Shakespeare uses one or more of the features of syntactic complexity we have identified. Most of the findings are therefore expressed in terms of the proportion of all relative *which* clauses in the corpus in which the given feature is found⁷². Where appropriate, rate of use will also be expressed in terms of '1 in x' relative *which* clauses⁷³, for clarity,

⁷¹ See p. 71.

⁷² Variations in the detail of this practice occur at times, when necessary, for example when analysing the number of words intervening between *which* and its antecedent. In this case, data is only presented as a proportion of all clauses with single word antecedents, since sentential clauses were obviously not countable in this way.

⁷³ '1 in x' single word antecedents will also be used, where appropriate.

although as explained in section 3.0, this measurement is not as precise as the relative frequency percentage.

The findings will be presented in the sequence in which the topics were discussed in the previous chapter, for ease of cross-reference. The results comprise a network of sometimes complex interrelating categories and their subsets, which required careful, repeated checking and cross-totalling in order to ensure that they were error-free. As part of this process, anomalies needed to be identified wherever they occurred, and accounted for in the statistics. These are presented and discussed in the relevant sections below, along with any occurrences that are unusual or difficult to categorise. Percentages in all cases are rounded to two decimal places. Statistical significance test results are presented, as before, at the foot of each table.

4.1 Main functions of *which*

There are two anomalies in this section. The first of these is a characteristically comic linguistic exception from Don Armado's letter, in the first scene of *Love's Labours Lost* (all italicised in the Craig text):

(99) *So much for the time when. Now for the ground which; which, I mean, I walked upon.*

(1.1.206)

When *which* is reiterated after the semi-colon here, it clearly functions as a relative pronoun, but in the first instance it seems best acknowledged as a metalinguistic oddity. Another anomaly appears in *The Winter's Tale*, in Hermione's trial scene:

(100) More than mistress of
Which comes to me in name of fault, I must not
At all acknowledge.

(3.2.56-8)

Here *which* is used in a rather untypical way, with no anaphoric function or antecedent, and as a pronoun introducing a nominal clause, paraphrasable as “what comes to me in name of fault”. It is interesting to speculate on why Shakespeare did not use *what*, which may have been clearer semantically, and more appropriate syntactically. There are numerous nominal clauses starting with *what* in the canon, but perhaps a more likely construction for Shakespeare in this case, found frequently elsewhere in the plays, and an earlier variant than *what*, might have been *that which*, which becomes commoner during the course of the seventeenth century (Kemp 1979, 187). Perhaps in this case he simply omitted *that* for metrical reasons, again. The boundaries between these different constructions were still fluid at the time, as Rissanen demonstrates with a case of *that* being used by Latimer in a similar way to *that which*:

Let vs not inclyne our selues vnto the precepts and tradycyons of oure fathers, nor let
vs do *that* semes right in our eyes

(Rissanen 1999, 300)

Whatever the cause, however, the outcome in the occurrence from *The Winter's Tale* is that *which* is not functioning as an anaphoric relative in the same way as elsewhere here, so is counted as an anomaly.

As could be seen from the discussion of paradigmatic matters in 3.2, the percentage of *which* occurrences that are relatives increases from 92.5% in the first thirty-two plays to 97.3% in the late group. Table 4.1 shows the wider picture.

Table 4.1 Main functions of *which*

<i>Which</i> function	First 32 plays	Late plays
Interrogative	131 (7.43%)	12 (2.70%)
Relative	1631 (92.51%)	431 (97.07%)
Pronoun	1562 (88.60%)	419 (94.37%)
Determiner	69 (3.91%)	12 (2.70%)
Total <i>which</i> occurrences	1,763 (incl. 1 anomaly (0.06%))	444 (incl. 1 anomaly (0.2%))

Statistical significance: risk level 0.001

We see here a significant drop in the proportion of *which* types that function as interrogatives, from 7.43% (or 1 in 13) of all occurrences in the earlier canon, to only 2.70% (1 in 37) in the late plays. Increased overall density of *which* occurrences in the corpus, and within that, increased frequency of the relative function, indicate a substantial increase in density of *which* relativisation in the late plays. Shakespeare also uses *which* as a determiner less frequently in the late group. Rissanen (1999, 296), Reuter (1938, 18), Romaine (1980, 227) and Nevalainen (2006, 86) discuss the determiner function of *which*, but do not identify it as a waning variant in Early Modern English, so the decline of this function in later Shakespeare may be an individual stylistic feature, although analysis of other, contemporary corpora might elucidate the situation more satisfactorily.

It is not easy to draw conclusions about the proportional reduction in interrogative usage in the late plays corpus. It seems unlikely that Shakespeare's characters are less frequently in questioning or decision-making situations in the plots of the late plays, so

perhaps the drop should be seen, once again, in terms of the significant increase in relativisation.

4.2 Antecedents

4.2.1 Types of relative clause according to antecedent

There are a few cases in the Shakespeare canon where the antecedent is not a clause or verb phrase, but cannot be reduced to a single word, as in the following lines from *A Midsummer Night's Dream*:

(101) A play there is, my lord, some ten words long
Which is as brief as I have known a play

(5.1.66-7)

Although it would fit syntactically, the antecedent here is not in fact “play”. “Which” refers anaphorically to the idea of the play’s brevity, expressed in the adjunct “some ten words long”, and so the antecedent cannot be reduced to a single word. Blake identifies a comparable phenomenon in an example from *Venus and Adonis*, where *which* is a determiner:

But *which* together with a single noun can refer to rather wider antecedents without repeating any word among the antecedents:

*So thou wilt buy, and pay, and use good dealing, Which
purchase if thou make*

(Blake 2002, 63).

Similarly, in this example from *Henry the Fifth* Shakespeare uses sleight of hand in linking *which* as a determiner to the noun “marriage”, to give the impression of anaphoric reference, despite having only presented the merest outline of such an idea in the foregoing clauses:

- (102) Also, King Lewis the Tenth
Who was sole heir to the usurper Capet,
Could not keep quiet in his conscience,
Wearing the crown of France, till satisfied
That fair Queen Isabel, his grandmother,
Was lineal of the Lady Ermengare,
Daughter to Charles, the foresaid Duke of Lorraine:
By the *which marriage* the line of Charles the Great
Was reunited to the crown of France.

(1.2.82-90)

There is clearly no single word antecedent here, despite the impression given of an adnominal relative clause through the use of determiner *which* and an adjacent noun. All cases where there is no single word antecedent have been assigned to the sentential category, because they seem to have more in common with sentential than adnominal clauses, and may be comparably complex. This was also necessary from a practical perspective, because this analysis aims to measure and describe the words intervening between antecedents and relatives, so the cases that cannot be measured, such as these, need to be separated out.

In some cases, as seen in 3.3.2.3, the antecedent is difficult to identify, and some discussion is needed here of the more difficult individual instances.

- (103) Call it a travel that thou tak'st for pleasure.

My heart will sigh when I miscall it so,
Which finds it an inforced pilgrimage.

(*Richard II*, 1.3.266-8)

- (104) Well, if any man in Italy have a fairer table *which* doth offer to swear upon a book, I shall have good fortune.

(*The Merchant of Venice* 2.2.57)

to lack a main verb and is itself rather syntactically opaque. There is no real syntactic resolution possible in this case. In this study, however, given the aim of statistical analysis, decisions have to be made where possible, and so in this case “all” was taken as the antecedent, since it seems to give the best semantic fit.

Two further cases need to be discussed in this context, since they present particularly complex problems for analysis. The first appears in one of Shakespeare’s more convoluted sentences, from *Henry VIII*:

(107) ‘If’, quoth he, ‘I for this had been committed,
As, to the Tower, I thought, I would have played
The part my father meant to act upon
Th’usurper Richard: who being at Salisbury,
Made suit to come in’s presence; *which* if granted,
As he made semblance of his duty, would
Have put his knife into him.’

(1.2.224-30)

The first antecedent that seems appropriate is “suit”, since the conditional clause element of the relative clause (“if granted”) seems to suggest a semantic link between “granted” and “suit”. However, “which” is the only component of the relative clause here that can serve as the subject for “would have put his knife in him” in line 230, so it is very unlikely to have “suit” as its antecedent. It transpires that “which” is not, after all, the object of “if granted”, but is co-referent with “he” in line 229. There are four agents involved in this extract: the surveyor (speaking, but reporting the alleged words of Buckingham), Buckingham (the surveyor’s employer), Buckingham’s father, and King Richard II. “[W]ho” in line 227 refers to Richard; Buckingham’s father is the subject of “made suit” in line 228, and “which” also refers to him. The surveyor is saying that Buckingham (said that he) would have put his knife

into the king (Henry VIII), as his father had intended to do to his own king, Richard. Surprisingly, this gives us “father” as the animate antecedent of “which”.

The second case is this from *Troilus and Cressida*, where there is a sense of an anaphoric adnominal link, but a more precise identification of the antecedent is difficult. An extensive section of preceding text is offered here in order to clarify the point:

(108) what the declin'd is
He shall as soon read in the eyes of others
As feel in his own fall; for men, like butterflies,
Show not their mealy wings but to the summer,
And not a man, for being simply man,
Hath any honour, but honour for those honours
That are without him, as places, riches, and favour,
Prizes of accident as oft as merit:
*Which when they fall, as being slippery standers,
The love that leaned on them as slippery too,
Doth one pluck down another, and together
Die in the fall.*

(3.3.81-92)

Achilles discourses on a subject that combines two well-known medieval and Renaissance topoi: the fall of princes, and the vanity of worldly renown. Present Day English linguistic habits would suggest that the antecedent should be the noun phrase closest to the relative. But (assuming that “which” and “they” in line 89 are coreferent), neither “merit” nor “prizes” gives a very satisfactory semantic fit with “when they fall”. There does seem to be a series of parallel noun phrases preceding “which”, more than one of which could be a candidate for antecedent: “men, in line 78, “honours” in line 81, the coordinated series “place, riches and favour” and then “prizes”. Ultimately, it seems that “which” must be the subject of “Doth one pluck down another” in line 91, (and also therefore of “die” in line 92), so the likely antecedent is “men” in line 78.

There are also a few anomalous occurrences in this category, where relative *which* does not have an antecedent. As can be seen in Table 4.2, there are 6 in the earlier plays. Firstly there is one occurrence that cannot be assigned a category because an aposiopesis follows it, breaking off the rest of the clause:

- (109) Do not plunge thyself too far in anger, lest thou hasten thy trial: *which* if—Lord have mercy on thee for a hen!

(*All's Well That Ends Well*, 2.3.187)

Which is clearly a relative of some kind here, but more cannot be determined without further substance, so this case is an anomaly, in this section and some others. There are also five cases of relativiser-only *which*⁷⁴ where there is no discernible antecedent, examples of which follow. A wider than usual textual context is provided, for clarity.

- (110) But from the inward motion to deliver
Sweet, sweet poison for the age's tooth;
Which, though I will not practise to deceive,
Yet, to avoid deceit, I mean to learn

(*King John*, 1. 1. 212-15)

Despite first impressions of a possible antecedent link with “poison”, “age” or “tooth”, it is clear that none of these makes satisfactory sense, because “deceive” is intransitive here. [*W*]hich heralds the relative clause, but is in fact replaceable by *and*, and has no antecedent in the foregoing clause. Similarly:

⁷⁴ See section 3.3.4.3.

- (111) The senators with one consent of love
Entreat thee back to Athens: who have thought
On special dignities, which vacant lie
For thy best use and wearing.

They confess
Toward thee forgetfulness too general, gross;
Which now the public body, which doth seldom
Play the recanter, feeling in itself
A lack of Timon's aid, hath sense withal
Of its own fail, restraining aid to Timon

(*Timon of Athens*, 5.1.132-41)

“Which” in line 137 here is a typical relativiser-only⁷⁵, replaceable simply by *and*, but it also has no anaphoric referent in the preceding clause, so needs to be treated as anomalous in this category.

As Table 4.2 shows, the changes in the late plays in this feature are very small, to the extent that they are not statistically significant. Even the slightly greater increase in sentential antecedents is not a statistically significant result.

⁷⁵ See 3.3.4.3.

Table 4.2 Types of relative clause according to antecedent

	First 32 plays (% of relative <i>which</i> clauses)	Late plays (% of relative <i>which</i> clauses)
Adnominal clauses (single NP antecedent)	1176 (72.10%)	306 (71.00%)
Adnominal clauses (multiple NP antecedent)	57 (3.49%)	17 (3.94%)
Total adnominal clauses	1,233 (75.60%)	323 (74.94%)
Nominal clauses (beginning <i>that which</i>)	149 (9.13%)	40 (9.28%)
Total single word⁷⁶ antecedents	1,382 (84.73%)	363 (84.22%)
Sentential antecedents	243 (14.90%)	68 (15.78%)
Total relative <i>which</i>	1631 (incl. 6 anomalies)	431

4.2.2 The animacy parameter

In this category all the single word antecedents in the corpora (see Table 4.2) are counted for animacy or inanimacy. As discussed in section 3.2, the animacy parameter was in the process of being established during Shakespeare's lifetime, so there was a steady decrease in the use of *which* for animate antecedents. Dekeyser tabulates this trend:

⁷⁶ As explained in the previous chapter, all noun phrase antecedents are reduced to single (head) words in this study for the purpose of counting, and the term 'single word antecedent' refers to those, as distinct from sentential antecedents.

Table 4.3 “The de-humanization of (THE) WHICH” (Dekeyser 1984, 71)

	+H	-H	Totals
1520-1560	4,402 (33.22%)	8,851 (66.78%)	13,253
1560-1599	371 (18.18%)	1,670 (81.82%)	2,041
1600-1649	97 (10.34%)	841 (89.60%)	938
Totals	4,870	11,362	16,232

Shakespeare’s animacy percentages (Table 4.4) are lower in both corpora than the lowest (the 1600-1649 section) in Dekeyser’s table, which may perhaps be due to generic or register differences in the texts sampled (many of them are discursive prose), but it is clear that the trend in Shakespeare is in the other direction. Shakespeare’s late plays use animate antecedents more frequently than the earlier plays. In the earlier canon 7.02% (or 1 in 14) of *which* clauses with a single word antecedent has an animate antecedent, whereas in the late plays that figure rises significantly to 9.64% (or 1 in 10). Despite the unfamiliarity of this usage to Present Day English ears, it is doubtful that animate occurrences of *which* could contribute either to syntactic or semantic complexity, but it is an interesting finding in itself.

Table 4.4 The animacy parameter

	First 32 plays (% of single word antecedents)	Late plays (% of single word antecedents)
Animate antecedent	97 (7.02%)	35 (9.64%)
Inanimate antecedent	1,285 (92.98%)	328 (90.36%)
Total clauses with single word antecedents	1,382	363

Statistical significance: risk level 0.10

4.2.3 Distance from the relative

Once again, this part of the analysis pertains to single word antecedents only, and records the number of words intervening between antecedent and relative⁷⁷. Decisions needed to be made concerning the significance of the different numbers of words intervening, both for the methods of measurement and the interpretation of the results. Should all words intervening be counted? At what level does the intervention make the antecedent significantly distant? Occurrences with ten words intervening would in most cases be very strong candidates for increased complexity, but what about five words? There is necessarily an element of arbitrariness in the assignation of such boundaries, since the number of words intervening is only one factor among several potentially contributing to complexity. However, it seemed appropriate to begin the count at two words distant, since antecedents with one word intervening are both numerous and unlikely to represent any challenge to syntactic processing purely on the basis of distance⁷⁸. The point where the intervention would be seen as significantly distant, with a potential for further complexity, was taken as five words intervening. In this table, and in all those subsequent tables dealing with interventions either to the left or the right of *which*, the total number of interventions is also given, along with the proportion of that total represented by each individual figure. This was felt to be important since it is quite possible that, even if the quantity of interventions did not increase, the interior proportions of the body of interventions might change, either towards or away from the more complex syntactic features.

⁷⁷ See 3.3.2.3 for details.

⁷⁸ A one word intervention might present challenges to processing in other respects, though, for example if it contained clausal material, but this aspect is analysed in the following section.

In Early Modern relative clauses the immediately adjacent antecedent was the most typical feature, as in Present Day English, so it is no surprise that the raw frequencies for intervening words are small. The effect of this is that the changes between the corpora in the table as a whole are not sufficiently marked to be statistically significant. However, some of the individual comparisons are statistically significant, namely the drop in relative frequency of occurrences of antecedents both 2 and 4 words distant (risk levels 0.05 and 0.10 respectively), and the overall relative increase in five words or more intervening, (risk level 0.2). There is some evidence here of a shift from closer to more distant antecedents in the late plays. 25.63% (or 1 in 12) single word antecedent clauses in the earlier canon have distant antecedents, but this rises to 36.88% (or 1 in 9) in the late plays.

Table 4.5 Distance between antecedent and relative

	First 32 plays (% of single word antecedents)	Late plays (% of single word antecedents)
2 words distant	143 (10.35%)	24 (6.61%)
3 words	106 (7.67%)	30 (8.26%)
4 words	73 (5.28%)	11 (3.03%)
5 words	39 (2.82%)	12 (3.30%)
6 words	17 (1.23%)	6 (1.65%)
7 words	19 (1.37%)	7 (1.93%)
8 words	12 (0.87%)	5 (1.38%)
9 words	6 (0.43%)	3 (0.83%)
10 or more words	18 (1.30%)	5 (1.38%)
Total interventions with 5+ words	111 (8.03%) (25.63%)	38 (10.47%) (36.88%)
Total clauses with single word antecedents	1,382	363

4.2.4 Word groups intervening to the left

As in the previous chapter, the second part of this section involves the identification and counting of the phrases and clauses intervening between the antecedent and *which*, the results for which are contained in Table 4.6 below, but some problem cases require discussion before the results are considered. Although it has been made clear that a ‘top-level only’ approach was taken in analysing the syntactic components of interventions, occasionally slight exceptions were made, as in the case of this speech from *Twelfth Night*:

(112) I have your own letter *that induced me to the semblance I put on*; with the which I doubt not but to do myself much right, or you much shame

(5.1.274)

The antecedent of “which” here is “letter”, followed by a relative clause heralded by “that”, and if we were adhering strictly to the ‘top-level only’ rule, the intervention would be classified as a single finite clause. In this case, and a few others like it, the zero relative clause “I put on”, embedded in the noun phrase “the semblance I put on”, was noted, and the case counted as a multi clause intervention, since the presence of more than one intervening predicate was felt to be significant.

Finally, it was decided not to include in the count interventions comprising single grammatical words such as a preposition or determiner, as we see here:

(113) I say, there is no darkness but ignorance, *in* which thou art more puzzled than the Egyptians in their fog.

(*Twelfth Night*, 4.2.21)

Grammatical words such as *this* were felt to be tightly linked to the relative, so not to belong syntactically to the linguistic material of the intervention⁷⁹. *Which* as a prepositional complement was counted as a separate statistic⁸⁰. Combinations of preposition and determiner (such as *of the which*) or conjunction and determiner (*and the which*) were also discounted as far as their role as interventions was concerned, for the same reason.

Some borderline cases also require discussion. There are occasional instances where the antecedent does not have a full nominal role in the clause of which it is a part (as subject, object or complement) but is perhaps a prepositional complement, as in the case below from *Twelfth Night*:

(114) A bawbling vessel *was he captain of,*
For shallow draught and hulk unprizable
With which such scathful grapple did he make

(5.1.50-52)

Here, semantic emphasis on the noun phrase “a bawbling vessel” is achieved through thematic fronting, giving a comparable syntactic sense to OVS inversion, although the noun phrase is ‘only’ the complement of the preposition “of”. Although technically “of a bawbling vessel” is a prepositional phrase, and the clause “was he captain” intervenes here, (with the phrase “for shallow draught and hulk unprizable” following), this case was counted as an intervening partial finite clause and a phrase, since it follows so closely the construction exemplified in (28) above⁸¹, suggesting it belongs more to the category of partial finite clause than finite clause.

⁷⁹ See 3.3.2.4.

⁸⁰ See 3.3.4.2.

⁸¹ See 3.3.2.4.

The group of *which* clauses counted here, the results for which are shown in Table 4.6, is of course almost the same as for Table 4.5, except that in the earlier count only those interventions comprising two words or more were included. In this case, phrases and clauses intervening were counted, meaning that there would be some cases of one word interventions which were now included⁸² that had not been previously, and this accounts for the slight increase in the totals between Tables 4.5 and 4.6.

Data for all the categories is shown in the table, including those cases where one or more additional phrase intervenes. However, because the raw frequencies are too small to be statistically significant when the categories are subdivided in that way, a statistical significance test was carried out for just the emboldened categories (total with a phrase, total with a non-finite clause, etc.). This data did prove to be statistically significant (at risk level 0.05). There is some overall increase in interventions containing clausal elements, but the most significant changes are in phrases, and non-finite and finite clauses intervening. Phrasal interventions decrease proportionally (with statistical significance risk level 0.10) from 18.74% (1 in 5 clauses) to 14.32% (1 in 7), whereas non-finite clauses and finite clauses increase significantly (0.20 and 0.10 respectively). The direction of this shift can certainly be considered evidence of increased syntactic complexity⁸³.

⁸² Some one word interventions constitute a phrase, according to the criteria set out in 3.3.4.2.

⁸³ See 3.3.2.4, with reference to Eva Berlage (2008).

Table 4.6 Word groups intervening to the left.

Feature intervening between antecedent and relative	First 32 plays (% of single word antecedents)	Late plays (% of single word antecedents)
1 phrase	238 (17.22%)	47 (12.95%)
2 or more phrases	21 (1.52%)	5 (1.38%)
Total with a phrase	259 (18.74%)	52 (14.32%)
1 non-finite clause	23 (1.66%)	8 (2.20%)
1 non-finite clause and phrase(s)	11 (0.79%)	6 (1.65%)
Total with a non-finite clause	34 (2.46%)	14 (3.85%)
1 partial finite clause	72 (5.21%)	13 (3.58%)
1 partial finite clause and phrase(s)	32 (1.81%)	8 (2.20%)
Total with a partial finite clause	104 (7.52%)	21 (5.78%)
1 finite clause	35 (2.53%)	13 (3.58%)
1 finite clause and phrase(s)	14 (0.87%)	8 (2.20%)
Total with a finite clause	49 (3.54%)	21 (5.78%)
Multiple clauses	22 (1.59%)	5 (1.38%)
Total with clause	209 (15.12%)	61 (16.80%)
Total with intervention	468 (33.86%)	113 (31.13%)
Total clauses with single word antecedent	1,382	363

The drop in proportion of partial finite clauses, even though not statistically significant on its own, is interesting in the context of the increases in other clausal interventions in the late plays. These cases seem to have the quality of a fixed syntactic formation, and very frequently involve the construction ‘NP is complement’, as may become clear from the following list of examples:

- (115) *The feast is ready which* the careful Titus
Hath ordained

(*Titus Andronicus*, 5.3.24-5)

- (116) For now *my love is thawed*,
Which, like a waxen image

(*The Two Gentlemen of Verona*, 2.4.198-9)

- (117) For *that is not forgot*
Which ne'er I did remember

(*Richard II*, 2.3.41-2)

It seems possible that the decline in usage of partial finite clauses is the result of Shakespeare tending to leave behind a specific syntactic feature. However, the trend that can also be seen in the table, of an increase in occurrences with a clausal construction and an additional phrase, is sustained even in this case, since there is still a higher relative frequency of partial finite clauses with an additional phrase in the late plays than there was in the earlier corpus.

4.3 The information parameter

4.3.1 Overall statistics

This section is concerned with the wider group of relative *which* occurrences as a whole, rather than simply the single word antecedents, because sentential antecedents are now included. Some occurrences in the canon cause problems for categorisation, lying on the borderline between restrictive and non-restrictive. One such case is the famous lines from *Macbeth*:

(118) Is this a dagger *which I see before me,*
The handle toward my hand?

(2.1.44-5)

There is a strong sense in which the relative clause “which I see before me” defines and specifies the noun phrase “a dagger” in the antecedent clause, as is typical of restrictive clauses. However, arguably the presence of the demonstrative “this” renders the antecedent clause independent. Certainly, if the relative clause is removed, leaving ‘Is this a dagger, the handle towards my hand?’ it still makes sense, and using Huddleston and Pullum’s test⁸⁴, it does not “drastically change the meaning”. On that basis this occurrence was counted as a non-restrictive clause. Another example of the problem occurs in Olivia’s closing lines of act 3 scene 1 in *Twelfth Night*:

(119) Yet come again, for thou perhaps mayst move
That heart, *which now abhors*, to like his love.

(3.1.129-30)

Again, the semantic crux is the pronoun before the antecedent, in this case “that”. If we understand it as anaphoric, referring back to a heart previously mentioned, it is possible that the relative clause “which now abhors” is supplementary information linked loosely to the antecedent. If however “that” has no previous mention in the dialogue, the relative clause must be seen as restrictive, since “that” would then refer cataphorically to the relative clause, with “heart” defined by it. In fact, in the speech prior to Olivia’s at line 161, Viola declares:

(120) I have one heart, one bosom, and one truth,
And that no woman has...

(*Twelfth Night*, 3.1.124-5)

⁸⁴ See section 3.3.3.1.

And so the “that heart which now abhors” is Viola’s heart, which abhors Olivia, and is indeed an anaphoric reference. The relative clause offers supplementary information about an antecedent which has already been defined, and is therefore non-restrictive. These considerations demonstrate, I think, the fineness of such semantic and syntactic distinctions, but with careful analysis it is usually possible to assign a category.

Table 4.7 The information parameter

Clause type	First 32 plays (% of relative <i>which</i> clauses)	Late plays (% of relative <i>which</i> clauses)
Restrictive	440 (26.98%)	115 (26.68%)
Non-restrictive	1190 (72.96%)	316 (73.32%)
Continuative	244 (14.96%)	82 (19.02%)
Total relative <i>which</i> clauses	1631 (incl. 1 anomaly)	431

The anomaly in the corpus of earlier plays is the instance of aposiopesis from *All’s Well That Ends Well*,⁸⁵ where, although *which* is obviously a relative pronoun, the absence of part of the relative clause makes ascription of an information parameter type impossible.

Table 4.7 shows that the proportions of restrictive and non-restrictive *which* clauses remain almost completely stable across the two corpora, with no statistically significant change. What does change significantly, however, (with a statistical risk factor of 0.10), is the relative frequency of Shakespeare’s use of continuative clauses in the late plays. In the earlier

⁸⁵ See (109) in 4.2.1.

canon 1 in 7 relative clauses is continuative, increasing to 1 in 5 in the late plays. It was decided that this increase warranted further attention.

4.3.2 Continuative clauses

The substantial increase in the percentage of continuative clauses in the late plays group can now be assessed for its significance. Both Nevalainen (2006, 86) and Reuter (1938, 39) describe it as pre-eminently a feature of sixteenth century English syntax, and Romaine summarises the trend by saying that “[d]uring the seventeenth century, [the continuative construction] begins to decrease” (1982, 84). Shakespeare’s syntactic style is therefore once again developing in the opposite direction from the background linguistic environment. Given this interesting result, continuative clauses were investigated in more detail, principally to see whether any other features associated with syntactic complexity in relative clauses might also be linked to the increased use of continuatives in the late plays. The continuative clauses in both corpora were therefore analysed to ascertain their relative frequencies of the features listed in Table 4.8 below. Sentential antecedents are listed separately.

Caution needs to be exercised in interpreting these results, because the raw frequencies are on the whole very small, and the results of the table as a whole are not therefore statistically significant. There are, nonetheless, some points of interest here. There is a significant increase (risk level 0.10) in the late plays in the proportion of continuative clauses that have distant antecedents, from 13.52% to 18.60% (or from 1 in 7 to 1 in 5 continuative clauses). Thus two features of syntactic complexity co-occur in the late plays with increased relative frequency.

Another feature that linguists associate with syntactic complexity is sentential antecedents, and there are interesting results shown here for the proportions of continuative clauses with sentential antecedents. This proportion is substantially higher in both corpora than in *which* clauses overall. Sentential antecedents make up 14.90% and 15.78% of relative clauses in the earlier and late plays respectively⁸⁶: about 1 in 6 relative *which* clauses have them. Here, however, we can see that almost 1 in 2 continuative clauses (43.03% and 46.51%) have sentential antecedents. Even though the changes seen in this table are not distinct enough to be statistically significant, this is an important finding in itself, and clearly suggests a link between sentential antecedents and continuative clauses, both of which are strongly connected with looseness of the relative link. In the late plays, this link is slightly stronger than in the earlier plays.

The proportion of *which* relative clauses where *which* is a determiner decreases overall from 3.9% in the earlier canon to 2.7% in the late plays. We can see here that the decrease is greater in the case of just continuative clauses, going from 25.00% to 13.95%, and the data for determiner *which* here are statistically significant (risk level 0.05). This is perhaps a surprising statistic given the fact that, according to Reuter, “[i]t was certainly largely due to this use of *which* that the continuative relative clauses gained such a wide circulation” (1938, 18). What we can see, in some confirmation of Reuter’s view, though, is that all the 12 determiner *which* occurrences in the late plays are found in continuative clauses. They may be, as we have seen, a declining function of which in Shakespeare’s later work, but where they occur they are still very much associated with continuative clauses.

⁸⁶ See Table 4.2.

Table 4.8 Analysis of continuative clauses

	First 32 plays (% of continuatives)	Late plays (% of continuatives)
2 words distant	10 (4.10%)	2 (2.32%)
3 words	9 (3.68%)	5 (5.81%)
4 words	11 (4.51%)	1 (1.16%)
5 words	5 (2.05%)	1 (1.16%)
6 words	4 (1.64%)	3 (3.49%)
7 words	9 (3.69%)	5 (5.81%)
8 words	7 (2.87%)	1 (1.16%)
9 words	1 (0.04%)	3 (3.49%)
10 or more words distant	7 (2.87%)	3 (3.49%)
Total 2+ words distant	63 (25.82%)	24 (27.91%)
Total 5+ words distant	33 (13.52%)	16 (18.60%)
With determiner <i>which</i>	61 (25.00%)	12 (13.95%)
Sentential antecedent	105 (43.03%)	40 (46.51%)
Total continuative clauses	244	86
Total relative <i>which</i> clauses	1631	431

One very marked increase here is in the overall proportion of clauses that have antecedents 5 or more words distant. In Table 4.5 we saw that this was 25.63% of relative clauses in the earlier canon, and 36.88% in the late plays. In the case of continuative clauses the same figures are 52.38% and 66.67%. Clearly then, continuative clauses are not only associated with sentential antecedents, but even those with single word antecedents are much more likely to have 5 or more words distant between the antecedent and the relative. Across the whole canon, the majority of *which* continuative clauses with single word antecedents have

5 or more words between their antecedent and the relative. Furthermore, this tendency increases with the late plays.

Although there is only one 5 word intervention in the late plays, there do seem to be intensified results in the categories of more distant interventions, with increases in 6, 7, 9 and 10 or more words distant (all more marked than in the statistics for relatives as a whole). As in the overall data for relatives, there are more three word interventions than two in the late plays (although the significance of this is somewhat obscure, and the raw frequencies very small). What can be said is that relative clauses in the late plays are significantly more likely to be continuative than in the earlier canon, somewhat more likely to be continuative and sentential, and if they have adnominal antecedents, then they are somewhat more likely to be six or more words distant from that antecedent. Continuative clauses in the late plays are also more likely even than non-continuatives in those plays to have seven or more words intervening.

If, as Rissanen and others suggest, continuative clauses, sentential antecedents, and distance between the relative and its antecedent are manifestations of increased syntactic complexity, then the occurrence of these three features together in a relative clause should be seen as an indication of intensified complexity. As there is some evidence to suggest that this intensification of features increases in the late plays, it would seem clear that syntactic complexity is also on the increase.

4.4 The role of *which* in the relative clause

This part of the study involves analysing the total group of relatives again, and includes an attempt to ascertain whether there is a change in the distribution of *which* usage in

the late plays according to Keenan and Comrie's Case Hierarchy⁸⁷, whereby the function of the relative in the relative clause is seen as related to complexity of syntax.

Some unusual cases arise in this section. Firstly, there are a very few instances in the canon where the principal verb element of the relative clause is ellipped, and where it is therefore more difficult to ascertain the role of *which* in the relative clause. One such is from Volumnia's speech to Coriolanus, cited here. She exhorts him to demonstrate humility before the Roman people, using gestures (indicated by the frequent deictics, such as "thus" here) and energetic, fragmented syntax:

(121) I prithee now, my son,
Go to them, with this bonnet in thy hand;
And thus far having stretch'd it,—here be with them—
Thy knee bussing the stones,—for in such business
Action is eloquence, and the eyes of the ignorant
More learned than the ears,—waving thy head,
Which often, thus, correcting thy stout heart,

(*Coriolanus*, 3.2.92-8)

"Which often" can be paraphrased as 'and do that often', so "which" is understood as equivalent to *that*, and therefore counted in this case as functioning as an object in the relative clause.

Occasionally there are anomalous examples of prepositional complement usage in relative clauses. One such is this, from the 'box tree scene' in *Twelfth Night*:

(122) By your leave, wax. Soft! and the impressure her Lucrece, with which
she uses to seal:

(2.5.53)

⁸⁷ See section 3.2.

Present day usage would find the preposition “with” redundant here, and since the occurrence is in prose there are no metrical exigencies contributing to the choice, so it may simply be an example of the fluidity of relativisation syntax at the time. It seems that “which” in this case is both an object and a prepositional complement in the relative clause, but since a decision was required one way or another for the count in this analysis, it has been listed as a prepositional complement. Conversely, occasional anomalies of this type occur (interestingly also from *Twelfth Night*):

(123) A spirit I am indeed;
 But am in that gross dimension grossly clad
 Which from the womb I did participate.

(5.1.214-6)

Present Day collocational usage would prefer “participate” to take the preposition “in”, so would make the relative clause “in which from the womb I did participate”. Shakespeare’s “participate” is a transitive verb here though, with “which” as object. It is difficult to ascertain the cause of this anomaly, as with (121). Neither case is commented on by current editors, although frequently syntactic problems do go unremarked. Shakespeare’s usage may simply differ from ours, or it may be a compositor error, repeated from the First Folio (where it appears as above (Hinman 1996, 292)) through subsequent editions. (123) may be an elision for metrical purposes, although this seems unlikely, given that Shakespeare’s pentameter lines frequently have additional lightly stressed syllables, as “in” would be here. Whatever the provenance of the occurrence, (123) has been accepted in this analysis as a transitive usage of “participate”, and “which” therefore as an object in the relative clause.

An interesting, if very unusual, case of relativiser-only *which* occurs in *Coriolanus*:

- (124) Away, my disposition, and possess me
 Some harlot's spirit! My throat of war be turn'd,
Which quired with my drum, into a pipe
 Small as a eunuch, or the virgin voice
 That babies lulls asleep!

(3.2.138-42)

Here, the anaphoric role of “which” is clear enough, linking it to the antecedent “throat”, but there is no relative clause as such at all here. In this case, “which” simply functions as a coordinator linking the predicator of the antecedent clause (“be turn'd”) to an additional non-finite clause (“quered with my drum”). Paraphrased into Present Day English, “which” would disappear altogether here, and the coordinated verb phrases would be juxtaposed asyndetically. The fact that this case is so unusual in the canon perhaps suggests that the relative appears here for metrical reasons, making up the pentameter in that line, but if nothing else it again reveals the flexibility of *which* usage for Shakespeare, and the readiness with which he employs it as a simple coordinator.

Occasionally in Shakespeare the syntactic ‘strings’ that tie the relative to the surrounding word groups seem to become tangled, as can be seen in this example from *Cymbeline*. In a moment of anxiety at the slaying of Cloten by Guiderius, Belarius's syntax becomes somewhat disjointed as he struggles to determine Cloten's motives for being in Wales, trying to decide whether Cloten had servants or companions with him:

- (125) Although, perhaps,
 It may be heard in court that such as we
 Cave here, hunt here, are outlaws, and in time
 May make some stronger head; *the which he hearing*, –
As it is like him, – *might break out, and swear*
He'd fetch us in

(4.2.177-82)

The relative clause (italicised) is a double pushdown construction⁸⁸, with *which* performing a grammatical role in both the lower element (“he hearing”) and in the two parallel higher elements of the relative clause (“might break out” and “swear he’d fetch us in”). However, different antecedents for “which” are implied by these lower and higher elements. The lower element must have the sentential antecedent contained in the “that” clause of lines 178-9, namely “such as we cave here, hunt here, are outlaws”, but by the time the relative clause progresses to its higher elements the antecedent has become the personal pronoun for Cloten himself, namely the “he” of the lower element. As earlier, this could be Shakespeare using a kind of syntactic shorthand for dramatic purposes, or perhaps a mimetic display of linguistic panic and confusion. It does not seem necessary to attribute a cause to the phenomenon; rather to register it simply as an anomaly for counting purposes. It will only be counted as a double pushdown occurrence, despite the fact that it has more complex elements inside it.

Single pushdowns can sometimes present problems for ascription of the role of *which* in the relative clause. These constructions only have a grammatical role in the lower element of a two-part relative clause, as explained earlier⁸⁹. In some instances the lower elements are dependent (often conditional or concessive) clauses, in which cases it is a straightforward matter to identify the grammatical role played by *which*, as in this example from *Richard III* (the lower element is highlighted):

(126) Star’d on each other, and looked deadly pale.
Which when I saw, I reprehended them;

(3.7.28-9)

⁸⁸ See section 3.3.4.5.

⁸⁹ See section 3.3.4.5.

Here, *which* is clearly the object of the dependent lower element clause “which when I saw” (in an OSV inversion). Also, in the same way as in a normal relative clause, *which* can be a prepositional complement in the lower element of a single pushdown construction, as can be seen in (63):

My riches are these poor habiliments,
Of which if you should here disfurnish me,
You take the sum and substance that I have.

(The Two Gentlemen of Verona, 4.1.16-18)

On these two occasions, and all others like them, *which* is counted in the normal way, and included in the main category statistics. Where the lower element is a non-finite clause, subject and object function were also ascribed where possible. For example:

(127) Proteus, I thank thee for thine honest care,
Which to requite, command me while I live

(The Two Gentlemen of Verona, 3.1.24-5)

“Which” is clearly the object of the non-finite clause in this case. When the lower element is just a past participle, on the other hand, the solution is more problematic:

(128) What heart receives from hence the conquering part,
To steel a strong opinion to themselves?
Which entertain'd, limbs are his instruments...

(Troilus and Cressida, 1.3.360-2)

In occurrences such as these two, there is insufficient clausal material to determine satisfactorily whether *which* is the subject or object of the past participle, so no ascription was made. This means that some, but not all, single pushdown constructions are included in the statistics for subject, object and prepositional complement. Details of all these are included in Tables 4.9 and 4.10 below.

Very occasionally we come across a single pushdown construction with more than one lower element. Richard of Gloucester's breathtaking display of persuasive rhetoric in act 1 scene 2 of *Richard III*, as he seduces Anne alongside the corpse of her father in law (whom Richard himself murdered), seems to be paralleled by his control of complex syntax:

(129) Lo! here I lend thee this sharp-pointed sword;
Which if thou please to hide in this true breast,
And let the soul forth that adoreth thee,
I lay it naked to the deadly stroke,
And humbly beg the death upon my knee.

(1.2.185-9)

The lower element of the relative clause has two parts here, very unusually, where the dependent conditional clause “which if thou please to hide in this true breast” is followed by a co-ordinated non-finite⁹⁰ clause (which itself, extraordinarily, contains an embedded restrictive relative clause) “And let the soul forth that adoreth thee”. These two lines form two lower elements, both conditional dependent clauses, to the higher element (“I lay it naked” etc.) that follows. “Which”, however, only has a role in the first of the two lower element clauses, so, as far as categorization is concerned, it is simply counted as the object of the first conditional clause, of “hide”⁹¹. There is one other example in the canon of a single pushdown construction with a double lower element, from *Cymbeline*:

(130) he began
His mistress' picture; *which by his tongue being made,*
And then a mind put in 't, either our brags
Were crak'd of kitchen trulls, or his description
Prov'd us unspeaking sots.

(5.5.207-11)

⁹⁰ “[L]et” is probably non-finite because it parallels “hide”.

⁹¹ The occurrence is further complicated syntactically by the compound verb phrase “please to hide”, which is almost – but not quite – a recursive clause in itself.

In this case the speaker's prolixity is the dramatic point of the extended clauses, as Iachimo tantalizingly suspends the denouement narrative of the play's final scene. As in the previous instance, *which* does not have a role in the second, coordinated lower element, so the assignation is once again the role of object.

As discussed in 3.3.4.6 above, recursions present problems of categorization, because of the clausal embedding which defines them, and the role of *which* in the relative clause is one such issue, as can be seen here:

(131) Doubt not, the commoners, for whom we stand,
But they upon their ancient malice will
Forget with the least cause these his new honours,
Which that he'll give them, make I as little question
As he is proud to do't.

(*Coriolanus*, 1.2.141-5)

The subject of the relative clause is "I", and "which" is embedded in the rankshifted "that" clause, functioning there as the object, although not, of course, as the object of the top level clause, which is "as little question". Recursions, therefore, are simply listed as such in the table, rather than in terms of subject or object function.

Finally, statistical complexities arise in this section, which need to be addressed. Although this part of the analysis is principally intended to assess Shakespeare's *which* clauses in terms of the Case Hierarchy, that model is only concerned with clauses where the relative has a grammatical function, which in this context means the two nominal functions of subject and object, and prepositional complements, which usually form part of an adjunct. As we have seen however, there are other, less straightforward, types of *which* role in the relative clause in evidence in the Shakespeare canon, only some of which are encompassed by the grammatical

role categories. The interrelationships between the different roles listed in this section therefore require some further clarification before considering the results.

The categories of subject, object and prepositional complement are obviously mutually exclusive. Cataphoric constructions such as examples (58) to (61) are self-contained relative clauses (albeit without antecedents) and they are counted normally, in terms of the grammatical role of *which* in that clause, and are included in the statistics for subject and object in Table 4.8. There are then four types of relative construction where the role of *which* in the relative clause is more problematic for counting. Double pushdown constructions, firstly, are those where *which* has a role in both the lower and higher elements of the relative clause⁹², so in these cases it would be possible to count two roles for *which* instead of one. In an attempt to maintain clarity and simplicity where possible, it was decided to count only the role of *which* in the higher element in these cases. *Which* can be subject, object or prepositional complement in double pushdowns (although in fact only one case in the canon is a prepositional complement).

Secondly, relativiser-only constructions do not have a grammatical role in the relative clause at all, so are counted separately and are mutually exclusive with the main categories of subject and object. However, there are two cases of relativiser-only constructions in the canon that are also prepositional complements, shown here with the *which* constructions italicised:

- (132) She should this Angelo have married; was affianced to her oath, and the nuptial Appointed: *between which time* of the contract, and limit of the solemnity, her brother Frederick was wracked at sea

(*Measure for Measure*, 3.1.188-9)

⁹² See section 3.3.6.5 for explanation of this type.

- (133) To-night at Herne's Oak, just 'twixt twelve and one,
Must my sweet Nan present the Fairy Queen;
The purpose why is here: *in which disguise*,
While other jests are something rank on foot,
Her father hath commanded her to slip
Away with Slender...

(*The Merry Wives of Windsor*, 4.6.19-24)

Which has no grammatical role in any clause in either of these two instances, simply functioning (as is the norm with relativiser-only *which*) as a herald of a dependent clause to follow, with an anaphoric link back to the antecedent.⁹³ But since prepositional complements are counted in the case of ordinary *which* clauses, the fact that they are also prepositional complements needs to be recorded in the statistics. These cases are therefore counted both as relativiser-only cases and prepositional complements.

Statistical details of the overlapping category relationships discussed here are included in Table 4.9 below, in order to clarify matters further. Note that some features overlap: for example some prepositional complements are also single pushdown or relativiser-only clauses, and the single and double pushdown categories are additional, not alternative, to subject and object function. Similarly, some occurrences (recursions and some single pushdowns) do not have an identifiable role for *which*. As a result of this overlapping of categories, the figures do not total exactly 100% (of relatives) overall. Since the data is already complex, and the present table shows the key trends between the corpora, it was decided only to present the findings in this manner.

⁹³ See section 3.3.4.3.

Table 4.9 The role of *which* in the relative clause: principal functions

Role of <i>which</i>	First 32 plays (% of relative <i>which</i>)	Late plays (% of relative <i>which</i>)
Subject (incl. single and double pushdowns)	771 (47.27%)	232 (53.83%)
Object (incl. single and double pushdowns)	579 (35.50%)	131 (30.39%)
Prepositional complement (incl. single and double pushdowns)	216 (13.25%) (incl. 2 relativiser-only)	50 (11.60%)
Total relative <i>which</i>	1631 (incl. 1 anomaly ⁹⁴)	444

Statistical significance: risk level 0.05

The late plays show a shift towards *which* as a subject in the clause, from 47.27% to 53.83% of all *which* relatives, principally away from object, and to a lesser extent away from all the other functions shown here. As far as the Case Hierarchy is concerned, this seems to represent a move away from complexity and towards simplicity of syntax. The trend in the linguistic background was from *wh*- forms with subject function (where they were first established) towards increased use in the more complex functions (e.g. Dekeyser and Ingels 1988, 31), so a shift towards the more complex end of the Hierarchy might have been expected, given the evidence elsewhere for increased syntactic complexity in the late plays. What these results do have in common with others, though, is the fact that once again Shakespeare's syntactic usage seems to be retrogressive in relation to the background linguistic trends.

⁹⁴ Anomalies in this category are occurrences of *which* that are relatives, but do not function as any of the features listed here. There is only one in the canon, namely the aposiopesis from *All's Well That Ends Well*, for discussion of which see p.162.

Table 4.10 The role of *which* in the relative clause: additional constructions

Role of <i>which</i>	First 32 plays (% of relative <i>which</i>)	Late plays (% of relative <i>which</i>)
Relativiser-only	17 (1.04%) (incl. 2 prepositional complements)	3 (0.70%)
Single pushdown	56 (3.43%) (incl. 13 not assigned a role elsewhere ⁹⁵)	9 (2.09%)
Double pushdown	15 (0.92%)	17 (3.94%)
Recursion	37 (2.27%)	15 (3.48%)
Cataphoric <i>which</i>	15 (0.92%)	2 (0.46%)
Total relative <i>which</i>	1631 (incl. 1 anomaly, as Table 4.9 above)	431

Statistical significance: risk level 0.001

Table 4.10 shows the data for the additional constructions of *which* in the clause, and achieves the highest level of statistical significance. There are several important points to notice here.

Firstly, relativiser-only clauses are examples of minimal *which* involvement in the relative clause, and might intuitively seem to be an example of syntactic complexity, at least to a present day auditor or reader⁹⁶, yet they show a considerable drop in relative frequency (going from 1 in 96 *which* clauses to 1 in 142), along with cataphoric *which* (from 1 in 108 to 1 in 217). But interestingly, in cases where *which* is more involved in the syntax of the relative clause, we see a striking increase in occurrences. There is a 300% increase in the density of

⁹⁵ See p.176 for discussion of this type.

⁹⁶ See for instance examples (54) and (55) on pages 121 and 122.

double pushdown constructions in the late plays, rising from 1 in 108 to 1 in 25 clauses. There are more of them in the six late plays than in the entire corpus of thirty-two preceding. And recursions, the other highly complex formation that occurs with *which* relatives, increase from 2.27% to 3.48%, or 1 in 44 *which* clauses to 1 in 29.

The marked increase in Shakespeare's adoption of the double pushdown construction in the relative clauses of the late plays, further strengthened by the increase in recursive structures, is, I would argue, an unequivocal manifestation of increased syntactic complexity in relativisation.

4.5 Interventions to the right of *which*

4.5.1 The number of words intervening to the right

As explained in 3.3.5, the first part of this section of the analysis involves counting the number of words intervening to the right, between *which* and the next principal component of the relative clause, which can be the main verb, or the subject or object. The total number of *which* clauses with an intervention between the relative and the next principal component of the relative clause is shown here to decrease somewhat in the late plays, falling from 23.11% of *which* relatives to 21.11%. But where in the analysis of antecedent distance we saw a shift from smaller to larger interventions, here the opposite occurs. Surprisingly, (since in interventions to the left of *which* the trend was in the other direction), the proportion of relative clauses with interventions to the right of five or more words drops significantly (risk level 0.025), halving in the late plays. In the earlier canon 1 in 22 relative clauses has a five or more word intervention to the right of which, whereas in the late plays this drops to only 1 in 48).

Table 4.11 Number of words intervening to the right

	First 32 plays (% of relative <i>which</i>)	Late plays (% of relative <i>which</i>)
1 word	140 (8.58%)	35 (8.12%)
2 words	38 (2.33%)	23 (5.34%)
3 words	77 (4.72%)	17 (3.94%)
4 words	48 (2.94%)	7 (1.62%)
5 words	15 (0.92%)	4 (0.93%)
6 words	18 (1.10%)	2 (0.46%)
7 words	19 (1.16%)	2 (0.46%)
8 words	9 (0.55%)	0
9 words	6 (0.37%)	1 (0.23%)
10 or more words	7 (0.43%)	0
Total with intervention	377 (23.11%)	91 (21.11%)
Total interventions with 5+ words	74 (4.54%)	9 (2.09%)
Total relative <i>which</i> clauses	1631	444

Statistical significance: risk level 0.025

The other marked and unexpected change is in the proportion of two word interventions, where we see a significant increase in the late plays (risk level 0.01), from 10.79% of all interventions to 25.27%. They increase from 1 in every 9 clauses to 1 in every 4. This pattern of shifts requires further explanation, which the next section of the analysis, the description of the word groups that make up these interventions, may provide.

4.5.2 Word groups intervening to the right

The second part of this section looks at the same body of clauses as the previous one, but takes further the examination of interventions to the right, considering the nature of the words intervening. The details of the categories were discussed in 3.3.5.2, but essentially interventions with phrases and clauses (finite and non-finite) were counted separately. Cases

with both clausal and phrasal material were also counted separately, as were those with more than one clausal element intervening. The total number of occurrences with some kind of clausal intervention was also felt to be important information, so this is also presented as a separate figure.

This section also presents a number of problems for ascription of categories. The question of the separability of phrases arises again, as discussed in 3.3.2.4, in the following:

- (134) in her eye I find
 A wonder, or a wondrous miracle,
 The shadow of myself form'd in her eye;
 Which, *being but the shadow of your son*

 Becomes a sun, and makes your son a shadow

 (*King John*, 2.1.513-7)

In the intervention italicized here, the question arises as to whether “of your son” is to be counted separately from the rest of the non-finite clause. It is not counted separately because the definite article preceding “shadow” means that the removal of the phrase would not leave behind a semantically acceptable construction. Multiple phrases can also be intertwined, making the decision problematic, as seen here in *Henry V*:

- (135) ‘tis best to weigh
 The enemy more mighty than he seems:
 So the proportions of defence are fill'd;
 Which *of a weak and niggardly projection*
 Doth like a miser spoil his coat with scanting
 A little cloth

 (2.4.47-52)

The semantic rule established in 3.3.2.4 needs to be somewhat modified here. If following the principle of removability, “of a weak and niggardly projection” might require categorization as two phrases intervening. However, the logical consequence of applying that practice would

be that all adjectives or adverbs, no matter how integrated syntactically, would be deemed additional phrases. This was felt to be a step too far, so phrases were only seen as separate if they were not embedded in other phrases.

Difficulties for category assignation arise also on the border between phrasal and verbal elements in this case from *Richard II*:

(136) And for we think the eagle-winged pride
Of sky-aspiring and ambitions thoughts,
With rival-hating envy, set on you
To wake our peace, which in our country's cradle
Draws the sweet infant breath of gentle sleep;
Which *so rous'd up with boist'rous untun'd drums,*
With harsh-resounding trumpets' dreadful bray
And grating shock of wrathful iron arms
Might from our quiet confines fright fair peace

(1.3.133)

Shakespeare draws on a number of rhetorical effects for this pivotal dramatic moment, as Richard grandiloquently sweeps aside the chivalric solution to his nobles' argument, sowing the seeds of his own downfall, and years of national strife. He uses no less than seven epithets in the intervention to the right of *which* here, including the compound epithet "harsh-resounding" (these being associated, appropriately enough, with national pride, as Adamson explains (1999, 580)). But both "rous'd up" and "harsh-resounding" are verb-derived epithets, endowing this intervention with clausal weight, to the extent that it would seem inappropriate to label this as a four phrase intervention. It was decided finally that "rous'd up" was a non-finite (past participle) clausal element, so this case was identified as an intervention of one clause and three phrases.

Another unusual case occurs here, in *The Winter's Tale*, in the famous conversation between Polixenes and Perdita:

(137) and streak'd gillyvors,
Which some call Nature's bastards: of that kind
Our rustic garden's barren, and I care not
To get slips of them.

 Wherefore, gentle maiden,
Do you neglect them?

 For I have heard it said
There is an art *which in their piedness shares*
With great creating nature

(4.3.102-4)

There is more than one problem with the italicized “which” clause. Clearly the phrase “in their piedness” intervenes to the right of “which”, but the antecedent of “which” is very unclear, and there is a lack of agreement between the clausal elements. “[T]heir” is plural, yet the predicate is singular, suggesting that they do not refer to the same noun phrase. Although the semantic content of this instance is usually understood fairly clearly, there does not seem to be a satisfactory analytical account available for this clause, and certainly not the intervention to the right. In this analysis the intervention was labelled as one phrase.

The situation regarding interventions to the right in single pushdown constructions was explained in 3.3.5. Typically, they do not occur in these constructions, but there are exceptions. Although *which* does not have a role in the main or higher element of the clause, phrases or clauses were counted as interventions on the rare occasions when they come between *which* and the next principal element of the lower element clause, which could be its principal verb, subject or object. An example of this is (130), which was shown to be an unusual case of a single pushdown with a double lower element, which of course means that there is a clausal intervention to the right of *which*:

a sentential relative clause, where the conditional lower element only has an adverb, not a verb:

(140) God's is the quarrel; for God's substitute,
His deputy anointed in His sight,
Hath caused his death; the which *if wrongfully*,
Let heaven revenge,

(1.2.39-42)

Because of the sense of an ellipped verb, the lower element was counted as a non-finite clause intervening between *which* and the higher element main verb “let...revenge”. And another case of an ellipsis of a verb in a double pushdown construction also needs discussion:

(141) O! now, after
So many courses of the sun enthroned,
Still growing in a majesty and pomp, the which
To leave a thousand-fold more bitter than
'Tis sweet at first to acquire, after this process
To give her the avaunt!

(*Henry VIII*, 2.3.7-12)

Here, the lower element is clearly the non-finite verb phrase “to leave”, but the main verb of the higher element – probably an intensive verb, precipitating the adjectival complement “a thousand times more bitter...” – is ellipped. Effectively, however, the intervention to the right remains unaltered by the ellipsis, since the complement simply replaces the main verb as the next principal component to the right, and the intervention is identified as the non-finite clause “to leave”.

The data in Table 4.11 below obviously concerns the same group of clauses as was analysed in 4.5.1, so we see the same raw frequencies and proportions of total interventions. In addition to the general interest in features of complexity or otherwise, we are approaching this set of data in the hope of an explanation for the marked increase in

proportion of two word interventions, and the sharp decrease in proportion of interventions of 5 words or more, seen in Table 4.10.

Firstly, the statistical significance of the data in Table 4.11 is at the highest level, mainly as a result of the very marked, significant increase in relative frequency of non-finite clauses in the late plays, part of an overall significant increase (risk level 0.025) in the proportion of clausal interventions to the right, from 5.58% of *which* clauses to 8.82%. On the other hand, and forming a striking picture overall, there is also a significant drop (risk level 0.025) in the relative frequency of phrases intervening to the right. Occurrences with phrases intervening drop from over three quarters of total interventions to just over half, while at the same time the earlier canon relative frequency of non-finite clauses intervening increases by almost 150% in the late plays. The statistical significance of the increase in proportion of non-finite clauses (a chi square value of 61.17) is six times as great as it needs to be to reach the maximum threshold for the highest significance level of 0.001 risk. The occurrence of non-finite clausal interventions to the right of *which* increases from 1 in 35 *which* clauses in the earlier canon to 1 in 15 in the late plays.

Table 4.12 Word groups intervening to the right

Feature intervening to the right	First 32 plays (% of relative <i>which</i>)	Late plays (% of relative <i>which</i>)
1 phrase	246 (15.08%)	47 (10.90%)
2 or more phrases	40 (2.45%)	6 (1.39%)
Total with phrase(s)	286 (17.53%)	53 (12.30%)
1 non-finite clause	31(1.90%)	23 (5.34%)
1 non-finite clause and phrase(s)	16 (0.98%)	5 (1.16%)
Total with a non-finite clause	47 (2.88%)	28 (6.50%)
1 finite clause	20 (1.23%)	6 (1.39%)
1 finite clause and phrase(s)	10 (0.61%)	0
Total with a finite clause	30 (1.84%)	6 (1.39%)
Multiple clauses (of either or both kinds)	13 (0.80%)	4 (0.93%)
Total with clause(s)	90 (5.58%)	38 (8.82%)
Total with intervention	376 (23.11%)	91 (21.11%)
Total relative clauses	1631	444

Statistical significance: risk level 0.001

What we are seeing evidence of here is a very marked shift in Shakespeare's syntax in the late plays, where intervening phrases to the right of *which* are replaced by clauses. Phrasal interventions containing similes, such as the following:

(142) I will not praise thy wisdom,
Which, *like a bourn, a pale, a shore*, confines
Thy spacious and dilated parts:

(*Troilus and Cressida*, 2.3.184-6)

- (143) Upon his bloody finger he doth wear
A precious ring, that lightens all the hole,
Which, *like taper in some monument*,
Doth shine upon the dead man's earthy cheeks

(*Titus Andronicus*, 2.3.232-5)

or prepositional phrases such as these:

- (144) Spare thy Athenian cradle, and those kin
Which *in the bluster of thy wrath* must fall
With those that have offended:

(*Timon of Athens*, 5.4.48-50)

- (145) Gone to seek his dog; which, *tomorrow, by his master's*
command, he must carry for a present to his lady.

(*The Two Gentlemen of Verona*, 4.1.59)

are being replaced by non-finite clausal interventions such as we see here in (65), and elsewhere in the late plays:

It is true, and I will give you comfort,
To give your dead lords graves, the which *to do*
Must make some work with Creon

(*The Two Noble Kinsmen*, 1.1.147-9)

- (146) Thou art like the harpy,
Which, *to betray*, dost with thine angel's face,
Seize with thine eagles' talons.

(*Pericles*, 4.3.55-7)

- (147) and that 'twas dangerous for him
To ruminate on this so far, until
It forged him some design, which, *being believ'd*,
It was much like to do.

(*Henry VIII*, 1.2.206-9)

This most interesting finding accounts for the changes observed earlier in the number of words intervening to the right in the late plays. The phrasal interventions are typically longer in terms of word number, but less syntactically complex than the shorter, clausal interventions they tend to be replaced with in the late plays, which frequently seem to comprise only two words. There would seem to be some correlation between this tendency and the account of Shakespeare's late style given by, amongst others, Anne Barton, cited earlier in 1.2, where she describes the language of *The Tempest* as having "a compression, a stripped-down quality, more extreme than anything else in Shakespeare's previous work. Linguistically, the play is remarkably intense." (1968,13).

The picture that is emerging here needs also to take account of the results discussed in section 4.4, for the role of *which* in the relative clause. The very marked increase in double pushdown constructions seen in Table 4.8 is of course related to the results we see here, since a number of the non-finite clause interventions to the right (such as (65) cited above) are double pushdown constructions. The late plays' increase in recursions, on the other hand, is not recorded in the statistics for interventions to the right, but can perhaps be seen as contributing further to the strong shift towards syntactic complexity derived from clausal material. As discussed in section 3.3.4.6, recursions often present what looks like an intervening finite clause to the right of *which*, typically a subject and a finite verb of mental process, such as "I think". In all cases, recursions introduce at least a finite clause intervening to the right, and should therefore be seen as contributing to the picture of increased clausal intervention in the late plays.

Finally, the very marked increase in non-finite clauses intervening to the right begs the question of whether they are less complex than finite interventions. The distinction is being

made in this analysis between finite and non-finite clauses intervening (both to the left and right) based on the assumption that there may be more syntactic weight in the former than the latter. However, in practice, such assessments of relative weights of complexity are not straightforward. While it is quite likely that some finite interventions to the right of relative *which* do add considerably to complexity, it seems from present experience that the majority of this kind are discourse markers or phatic phrases, rather than fully-fledged semantic components of the clause, such as we see in these two cases:

- (148) Let but the commons hear this testament—
Which, *pardon me*, I do not mean to read—

(*Julius Caesar*, 3.2.110-11)

- (149) Sir, you shall understand what hath befallen,
Which, *as I think*, you know not.

(*Othello*, 5.2.129-30)

In these cases, although arguably increased syntactic complexity results from the presence of a finite verb in the intervention, the interventions may be quite easily processed semantically due to their being frequently used pragmatic collocations. If these are compared with the two-word, non-finite clauses exemplified above in (65), (146) and (147), there seems to be more syntactic density and complexity resulting from the non-finite than from this type of finite intervention.

4.6 Clausal length

The average length of *which* clauses in the earlier canon, assessed according to the criteria set out in 3.3.6, is 10.13 words, whereas in the late plays it drops slightly to 9.47 words. Since the spectrum of clausal lengths extends from two words to over a hundred, a

more detailed analysis of the data is required, to establish whether change takes place in this category in the late plays. As Figure 5 clearly shows, there is very little change in this feature in the late plays.

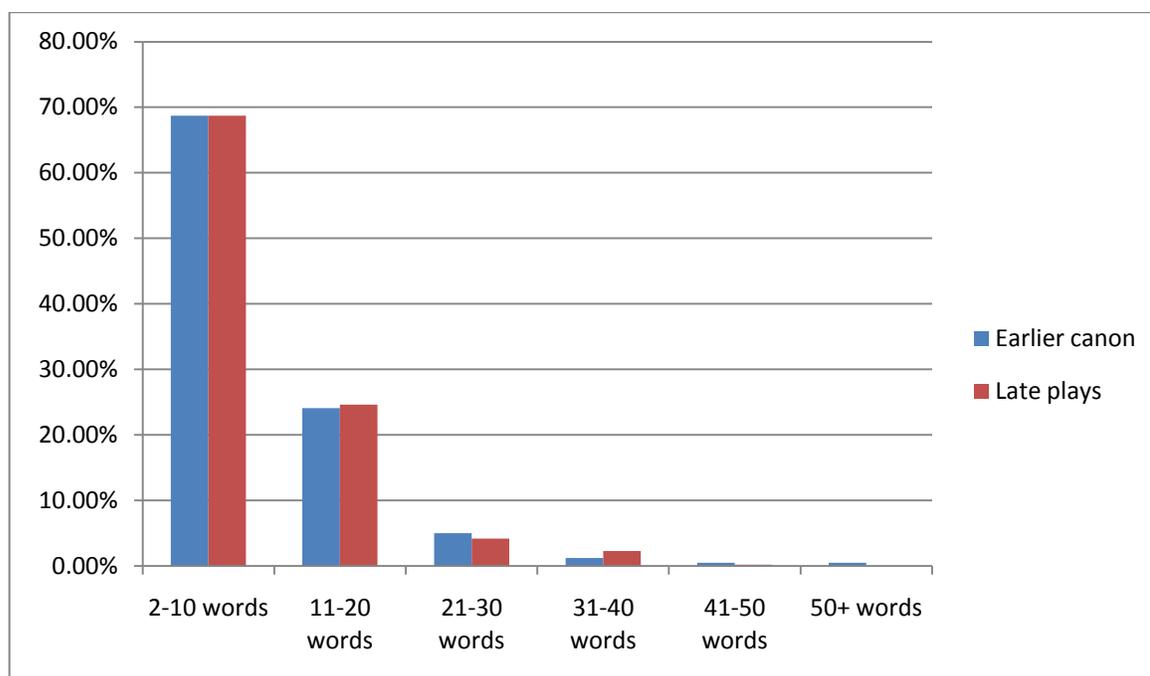


Figure 4.1 Clausal length

In both corpora, clauses of five words were the commonest, and the proportion of clauses between two and ten words in length is identical, at 68.7%. The very small variations that do occur appear in the increased numbers of clauses in the late plays that have between thirty-one and forty words, and in the decrease in numbers of late plays clauses with between twenty-one and thirty words, and over forty words. It is the latter two decreases that cause the overall slight drop in the average length with the late plays, but these changes do not seem to

constitute a significant trend of any kind, so this aspect of the analysis was not pursued any further.

4.7 Summary of findings for the late plays corpus

Analysis of the corpus of six late plays in comparison with the earlier thirty-two revealed a marked increase in the relative frequency of the type *which*, from 0.25% to 0.44% of running tokens, a highly significant finding, since relativisation is one of the principal sources of syntactic complexity. Furthermore, in the late plays the proportion of these *which* occurrences that function as relatives (as opposed to interrogatives) also increased, meaning an even more pronounced rise in the proportion of *which* relatives.

Relative *which* usage was then analysed syntagmatically, and a range of features with the potential to increase syntactic complexity were found to be more frequent in the late plays than the earlier corpus. The most salient of those are now summarised.

The use of *which* with animate antecedents increases in the late plays, from 7.02% (1 in 14) of all single word antecedents to 9.64% (1 in 10). Although probably not associated with syntactic complexity, this is an important finding, since Shakespeare's diachronic trend is in the opposite direction from that of the background language.

The distance between antecedent and relative *which* is generally recognised an important criterion for syntactic complexity, and some interesting shifts were observed between the earlier and late corpora. Shorter interventions (of 2 and 4 words) were found in significantly reduced proportion in the late plays, but all the categories from five words intervening and upwards showed very marked increases in their relative frequencies, with the total proportion of clauses with five or more words intervening rising from 25.63% (1 in 12)

of single word antecedents to 36.88% (1 in 9). The nature of these intervening word groups was then analysed. A significant decrease appeared in the late plays in the proportion of clauses with just phrasal material intervening to the left of *which*, along with an increase in clausal interventions. There was also an increase in the density of additional phrases with all types of clausal intervention in the late plays. These results suggest a shift from syntactically simple to complex interventions in the late plays. *Which* relative clauses in those plays are more likely to have a distant (five or more words) intervention between their antecedent and relative, to contain clausal material, and to contain additional phrases alongside those clauses.

Continuative clauses, the subset of non-restrictives identified by commentators as more syntactically complex than other kinds, were found to increase significantly in relative frequency. In the earlier canon 1 in 7 *which* clauses is continuative, but this rises to 1 in 5 in the late plays. It was also found that continuative clauses in the late plays are more likely to contain an intervention of five or more words between antecedent and *which*, and more likely to have a sentential antecedent than those in the earlier plays. We see here an intensification of potential complexity in late relative clauses with *which*, through the combination of several features all increasing in relative frequency in those six plays.

An analysis of the role of *which* in the relative clause revealed some increase in subject function for *which*, from 47.27% to 53.83% of *which* clauses. This seems to run counter to other findings, since Keenan and Comrie's Case Hierarchy⁹⁷ suggests that the subject function represents the greatest ease of syntactic processing, and is therefore associated with simplicity rather than complexity. Similarly, although seen by some commentators as a feature of clausal complexity, the length of the *which* clause was not found to increase in the late plays.

⁹⁷ See section 3.2.

Elsewhere, however, we see a demonstrable increase in syntactic complexity in the language of the late plays, with very marked increases in the proportion of double pushdown constructions and recursions. These two are rare constructions, and represent the most complex of all functions for *which* in the relative clause. Along with clausal interventions to the right, they are arguably the features that most intensify the convolutedness and complexity of the relative clause. There is an increase of over 300% in the proportion of double pushdowns, from 0.92% (1 in 108) of *which* clauses to 3.94% (1 in 25). There are more double pushdown constructions in the six late plays than in the entire previous canon.

Finally, the behaviour of *which* in the relative clause was looked at in terms of material coming between *which* and the next principal component of the clause to the right. In the earlier analysis of the behaviour to the left of *which*, we have already seen a pattern of diminishing proportions of phrasal interventions, alongside increasing clausal interventions, between *which* and its antecedents. To the right of *which*, the same pattern is found, but to a considerably greater degree, although some investigation was needed to explain an initial conundrum. The proportion of late plays clauses that have long, multi-word interventions to the right drops by half, from 4.54% to 2.09%, which seems to suggest a simplification of syntax, particularly as it seemed to be related to the very substantial rise in two word interventions, where the proportion more than doubles, from 2.33% to 5.34%. Only when the nature of the word groups intervening was considered did the picture become clear. In the late plays, there is a very pronounced development (with the highest statistical significance) in Shakespeare's syntactic habits in *which* clauses, from the use of long phrasal interventions to the right, to interventions that are shorter and clausal, particularly containing two-word, non-finite clauses. It seems possible that this is a manifestation of what a number of literary critics

have seen as characteristic of the late plays, namely a more syntactically and semantically condensed, as well as increasingly complex, style.

CHAPTER 5

ADDITIONAL PERSPECTIVES

5.1 A new late group?

We have seen so far, then, that a corpus comprising the six late plays (*Pericles*, *Cymbeline*, *The Winter's Tale*, *The Tempest*, *Henry VIII* and *The Two Noble Kinsmen*) has a very significantly higher density of *which* clauses than the preceding canon. It has also become clear that in that corpus the syntax of Shakespeare's *which* clauses changes significantly from the earlier work, in a number of respects. But if we now go back and look at the statistics for *which* frequency in all the individual plays of the canon, problems arise that threaten to undermine the premises of this analysis.

Table 5.1 shows the relative frequencies of occurrence of the type *which* in all thirty-eight plays, in chronological order (according to Wiggins et al, forthcoming). Results over 0.35% are emboldened for clarity. A concentration of high relative frequencies of *which* can be clearly seen at the bottom of the table, in the late plays. But two striking features appear here. Firstly, although it was included in the late plays corpus, *Pericles* has a considerably lower relative frequency of *which* than the other five members of that corpus, so seems not to belong to the group. Secondly, the high density of *which* begins earlier than previously assumed: it begins with *All's Well That Ends Well* in 1605, and includes *Macbeth*, *Antony and Cleopatra* and *Coriolanus*.

Table 5.1 *Which* frequency in the individual plays

Play	<i>Which</i> relative frequency (% of tokens)
<i>The Two Gentlemen of Verona</i> (1590*)	0.24%
<i>The Taming of the Shrew</i> (1590*)	0.13%
<i>Henry VI Part Two</i> (1591)	0.11%
<i>Henry VI Part Three</i> (1591*)	0.17%
<i>Henry VI Part One</i> (1592*)	0.20%
<i>Titus Andronicus</i> (1592)	0.15%
<i>The Comedy of Errors</i> (1592)	0.15%
<i>Richard III</i> (1594*)	0.27%
<i>Love's Labour's Lost</i> (1595*)	0.31%
<i>Richard II</i> (1595)	0.21%
<i>Romeo and Juliet</i> (1595)	0.26%
<i>A Midsummer Night's Dream</i> (1595)	0.23%
<i>King John</i> (1596)	0.30%
<i>The Merchant of Venice</i> (1596)	0.28%
<i>Henry IV Part One</i> (1597)	0.20%
<i>The Merry Wives of Windsor</i> (1597)	0.12%
<i>Henry IV Part Two</i> (1597-8)	0.32%
<i>Much Ado About Nothing</i> (1598)	0.28%
<i>Henry V</i> (1599)	0.27%
<i>Julius Caesar</i> (1599)	0.24%
<i>As You Like It</i> (1600)	0.27%
<i>Hamlet</i> (1600)	0.21%
<i>Twelfth Night</i> (1601)	0.12%
<i>Troilus and Cressida</i> (1602)	0.15%
<i>Measure for Measure</i> (1603-4)	0.27%
<i>Othello</i> (1603-4)	0.18%
<i>All's Well That Ends Well</i> (1605)	0.38%
<i>King Lear</i> (1605)	0.32%
<i>Macbeth</i> (1606)	0.48%
<i>Antony and Cleopatra</i> (1606)	0.45%
<i>Timon of Athens</i> (1607)	0.28%
<i>Pericles</i> (1607)	0.29%
<i>Coriolanus</i> (1608)	0.38%
<i>Cymbeline</i> (1610)	0.46%
<i>The Winter's Tale</i> (1611)	0.49%
<i>The Tempest</i> (1611)	0.44%
<i>Henry VIII</i> (1613*)	0.42%
<i>The Two Noble Kinsmen</i> (1613-14)	0.47%

*Asterisks denote dates still under consideration by Wiggins et al. (personal communication).

This second finding is extremely interesting, given the fact that a number of critics argue for the inclusion of *Macbeth*, *Antony and Cleopatra* and *Coriolanus* in the late group. Lyne (2007), Houston (1988), Kermode (2000) and McDonald (2006), for example, link one or more of *Macbeth*, *Coriolanus* and *Antony and Cleopatra* to the group of late plays, and all these three have *which* frequencies of 0.38% or higher, raising the question now of whether the boundaries of the group need to be redrawn.

Houston sees linguistic similarities between *Coriolanus*, *Antony and Cleopatra* and the late plays analysed earlier in this study, although it is principally their asyndetic syntax that he suggests they have in common (1988, e.g.182). Increased frequency of the relative pronoun *which* suggests the opposite: an increase in connective rather than asyndetic, and hypotactic rather than paratactic, syntax. Building on Houston, Russ McDonald discusses *Antony and Cleopatra*, *Coriolanus* and *Macbeth* in terms which construe them as exemplifying elements of what he sees as the late style. He identifies a seemingly paradoxical situation where both parataxis and hypotaxis increase in plays of this period:

The romances are notorious for their intricate syntax, some of it so clotted as to resist comprehension and to defy adequate paraphrase. Much of this complexity is attributable to the elliptical urge, in that the omission of non-essentials, particularly verbs and relative pronouns, markedly reduces the elements of the sentence and sometimes creates unfamiliar, puzzling grammatical relations. Yet a frequent source of syntactical difficulty is the poet's additive impulse, usually his continuation of an apparently completed sentence by appending non-restrictive relative clauses.

(McDonald 2006, 46).

McDonald may be referring in the final sentence here to a perceived increase in continuative clauses in the late plays, which is corroborated by the statistical findings of this thesis, but

since he does not distinguish between continuatives and non-restrictives⁹⁸ that point is unfortunately lost. What is more significant for the present analysis, however, is that he goes on to exemplify this “additive impulse” of the late style, not from one of the ‘romances’, but with part of a speech from *Macbeth*:

(150) Now, if you have a station in the file,
Not i'th'worst rank of manhood, say't;
And I will put that business in your bosoms,
Whose execution takes your enemy off,
Grapples you to the heart and love of us,
Who wear our health but sickly in his life,
Which in his death were perfect.

(3.1.112-8)

As McDonald rightly observes (2006, 46), this sentence includes an extended sequence of three relative clauses, heralded by the *wh*- relatives “whose”, “who” and “which”. I am not, however, aware of any writers who include *All's Well That Ends Well* in the late group in any respect, yet its relative frequency of *which* is as high as that of *Coriolanus*.

This new information presents a challenge to the premise of the analysis reported in Chapter 4 here, and points to a need for the investigation of a different configuration of late plays. Previously, the six late plays were selected on the somewhat tenuous basis of partial critical consensus⁹⁹, rather than any more empirical linguistic foundations, although Chapter 4 seemed to confirm that they were indeed linguistically distinct. If the present two sets of findings (high density of relative *which*, and marked increases in features of syntactic complexity) are indeed important elements of Shakespeare's late syntactic style, we need to analyse the nine plays with high relative frequency of *which* as a ‘new late’ corpus, to see

⁹⁸ “Usually the additional clauses that so distend the sentence are non-restrictive, or as linguists call them, “continuative clauses.”” (2006, 130).

⁹⁹ See page 10.

whether these nine are the ‘real’ late style plays. If this is indeed the case, we would expect to find that their high density of relative *which* is accompanied by marked features of syntactic complexity, as we did in the previous six.

This analysis was carried out as a second corpus comparison, in order to maintain the possibility of statistically significant results. If just the four new plays (*All’s Well That Ends Well*, *Antony and Cleopatra*, *Macbeth* and *Coriolanus*) had been selected for analysis, the low numbers of the raw frequencies involved may have prevented that.

The hypothesis upon which the new analysis is based, then, is that the nine plays in the new grouping are the ‘real’ late plays as far as syntactic style is concerned. If this is true, we would expect to see the features of syntactic complexity that were found in Chapter 4 sustained across the new group of nine plays. Specifically this would mean similar increases to those seen before in the density of *which* relativisation (as opposed to interrogative function), the density of key complexity features such as continuative clauses¹⁰⁰, antecedent distance¹⁰¹, and clausal intervention on both sides of *which*¹⁰², and also sustained reductions in the density of phrases intervening to the left and right of *which*¹⁰³. In fact, there are good reasons to anticipate even more marked increases and decreases in those features. Four plays with high density of the *which* type were originally in the corpus of the earlier canon, weighing against the original six late plays, and they have now been included. Similarly, *Pericles*, the play with *which* density of only 0.29%, has now been removed from the new late corpus.

¹⁰⁰ See 4.3.2.

¹⁰¹ See 4.2.3.

¹⁰² See 4.2.4 and 4.5.2.

¹⁰³ See 4.2.4 and 4.5.2.

So exactly the same analytical framework was used as in Chapter 4, in exactly the same sequence, but this time it was applied to the new groupings. The full set of results can be found in Appendix 9, but the most significant findings will now be discussed.

5.2 Findings for the ‘new late group’

What transpires from the new comparison is the opposite of the anticipated outcome. Where sustained or more marked results were expected, differences between the corpora are now attenuated, or even disappear altogether, in the new grouping.

Function of *which*¹⁰⁴

Although the nine plays now included in the corpus were selected because they all have markedly high relative frequencies of *which* (all above 0.35%), it transpires that the proportion of these that function as relatives is now slightly lower, at 95.72%, instead of 97.07% in the late plays. There is a corresponding smaller drop in the proportion of *which* with interrogative function, which is now at the higher level of 3.91% of *which* occurrences (1 in 21), instead of 2.70% (1 in 13) in the six late plays. Interrogative function is not, of course, associated with syntactic complexity.

Overall relativisation¹⁰⁵

Whereas in Chapter 4 we saw overall use of relatives increase in the six late plays to 1.31% of all tokens (1 in 76 running words was a relative), in the new corpus it only increases to 1.25% (1 in 80 running words). The difference between these results is statistically significant, with a risk level of 0.20.

¹⁰⁴ See Appendix 9, Table 5.

¹⁰⁵ See Appendix 9, Table 4.

The animacy parameter¹⁰⁶

In Chapter 4 animacy of antecedent was seen to increase from 7.02% (1 in 14) of single word antecedents in the earlier thirty-two plays to 9.64% (1 in 10) in the late plays. In the new analysis there is hardly any change in this feature at all. In the earlier plays with low *which* density, 7.55% (1 in 13) of single word antecedents of *which* are animate, and in the nine high *which* plays this figure is 7.58% (also 1 in 13). This outcome is unexpected, since the original six late plays (where the marked increase occurred) are still in the group under scrutiny. There must therefore be a considerably lower density of animate antecedents in the newly added four plays.

Antecedent distance¹⁰⁷

In Chapter 4 there were two key areas of antecedent distance where the late plays showed significant developments over the earlier canon, namely a drop in the proportion of *which* clauses with (single word) antecedents only 2 words distant from the relative, and an increase in those with antecedents 5 or more words away. There, the proportion that are 2 words distant fell from 10.35% (or 1 in 10) of earlier canon single word antecedents to 6.61% (1 in 15) in the late plays. In the new analysis, however, there is hardly any movement at all, with 9.83% (1 in 10) in the earlier (low *which*) corpus, and 9.13% (or 1 in 11) in the new group of nine (high *which*) later plays. The previous results were statistically significant, with a risk level of 0.005, but that significance is now completely lost.

A similar situation occurs with the significant increase that was found in distant antecedents (those with 5 or more words intervening) in the late plays. What was there an

¹⁰⁶ See Appendix 9 Table 7.

¹⁰⁷ See Appendix 9 Table 8.

increase from 8.03% (1 in 12) to 10.47% (1 in 9) of single word antecedents now dwindles to a negligible difference, from 8.28% (1 in 12) to 8.98% (1 in 11). Again, the statistical significance of the data comparing the two corpora is now lost. There is no statistically significant difference between the new low *which* and high *which* groups as far as antecedent distance is concerned. The introduction of the extra four plays has had the effect of removing any saliency in this respect.

Word groups intervening to the left¹⁰⁸

In the six late plays a significant change was observed in the nature and relative frequency of the word groups intervening between antecedent and relative in the six late plays. We saw there a drop in the relative frequency of phrases intervening, from 18.74% (or 1 in 5) of clauses with single word antecedents, to 14.32% (1 in 7), and this difference had a statistical significance risk level of 0.10. In the new comparison, however, the drop is only from 18.83% (1 in 5) to 16.10% (1 in 6), and the statistical significance falls to 0.20.

A similar pattern occurs in the case of non-finite clauses found intervening to the left of *which*. Previously this feature increased from 2.46% (1 in 41) of single word antecedents to 3.85% (1 in 26), and had a statistical significance risk level of 0.20. The new arrangement of plays does not produce a statistically significant change, with 2.55% (1 in 39) only rising to 3.10% (1 in 32).

The results for finite clauses intervening to the left also behave this way. In the previous analysis we saw an increase from 3.54% (1 in 28) in the earlier 32 plays to 5.78% (1 in 17) in the late plays, with a statistical significance of risk level 0.10. In the new analysis this

¹⁰⁸ See Appendix 9, Table 9.

statistical significance disappears, and the increase is only from 3.55% (1 in 28) to 4.80% (1 in 21).

The marked movement from phrasal to clausal intervention to the left of *which* has been lost. The newly-added four plays clearly do not have this feature of their *which* relativisation syntax.

Continuative clauses¹⁰⁹

The increase in relative frequency of continuative clauses was a prominent feature of the late plays found in the previous analysis. There we saw a rise from 14.96% (1 in 7) to 19.02% (1 in 5) of all relative clauses with *which*, with a statistical significance risk level of 0.10. In the new analysis this change is actually reversed, with a fall now from 15.97% (1 in 6) in the earlier, low *which* plays, to 15.53% (still 1 in 6) in the nine high *which* plays. The statistical significance of the new results is also lost.

Double pushdown constructions and recursions¹¹⁰

Double pushdown constructions are arguably the most complex feature of relativisation syntax in the canon, and their density was found to increase dramatically in the late plays from 0.92% (1 in 108) to 3.94% (1 in 25) of relative *which* clauses, results which have the highest level of statistical significance (0.001 risk). In the new configuration, however, the increase diminishes to one from 1.00% (1 in 100) to 2.24% (1 in 44), with a reduced statistical significance level of 0.05.

Recursions are also examples of unusually complex relativisation syntax, and (very unusually, given the picture that is emerging from this new corpus analysis of the nine plays)

¹⁰⁹ See Appendix 9, Table 10.

¹¹⁰ See Appendix 9, Table 13.

there seems to be evidence that their occurrence is distributed across the later part of the canon overall, not just the late six plays. In the original analysis they increased from 1 in 44 of *which* clauses in the first thirty-two plays, to 1 in 29 in the late six. Now, however, there is a bigger increase, from 1 in 48 to 1 in 30. Although the rate of occurrence is similar (and marginally lower) in the new, high *which* group, the difference is more statistically significant, with a risk level of 0.10 instead of 0.20.

Number of words intervening to the right

The original analysis in Chapter 4 saw a very significant doubling of the density of 2 word interventions to the right in the late plays, from 2.33% (1 in 43) of *which* clauses in the earlier canon, to 5.34% (1 in 19) in the late six, resulting (as it interestingly transpired) from a very marked rise in 2 word, non-finite clausal interventions to the right. In the new grouping, this difference diminishes to one from 2.38% (1 in 42) to 3.95% (1 in 25), and the statistical significance of the change is now only a risk level of 0.05, as opposed to 0.01 earlier, strongly suggesting that this feature is not found to the same extent in the four newly-added plays.

Word groups intervening to the right¹¹¹

Three aspects of interventions to the right of *which* were found to be significant in the analysis of the six late plays, namely the drop in proportion of clauses with phrases intervening, the very significant increase in non-finite clauses intervening, and the increase in proportion of those with any kind of clausal (as opposed to phrasal) intervention to the right.

The density of phrases intervening to the right decreased previously from 17.53% (1 in 6) to 12.30% (1 in 8) of *which* clauses, with a statistical significance of risk level 0.025. But

¹¹¹ See Appendix 9, Table 15.

once again the new configuration reveals attenuated results, with a smaller drop from 17.89% (1 in 6) to 13.95% (1 in 7), and the statistical significance falls to 0.05.

Non-finite clauses previously rose considerably from 2.88% (1 in 35) to 6.50% (1 in 15) of *which* clauses, with high statistical significance, at risk level 0.01. Yet again, in the new grouping this change is attenuated to an increase from 3.07% (1 in 32) to 4.60% (1 in 21), and the statistical significance is reduced (to risk level 0.10), supporting the evidence found in the previous section, where 2 word interventions also increased less markedly.

Finally, the overall proportion of *which* clauses that have clausal material intervening to the right of the relative went from 5.58% (1 in 18) to 8.82% (1 in 11) in the original late plays analysis, with statistical significance at risk level 0.025. Now, in the new late group there is almost no change in this feature at all. The low *which* plays have 6.14% (1 in 6), and the high *which* nine plays have 6.31% (still 1 in 6).

General comments

Clearly, the hypothesis behind the new corpus analysis¹¹² was wrong. This group of nine plays does not belong together in terms of their syntax. All of the most significant changes in the behaviour of *which* as a relative, found in the six late plays in Chapter 4, show attenuation in the new configuration. Since the original six plays are still in the corpus of nine high *which* plays, yet the changes between the two new corpora are so diminished, there can be no doubt that *All's Well That Ends Well*, *Antony and Cleopatra*, *Macbeth* and *Coriolanus* are quite different in their *which* syntax from the group of late plays written after 1609, namely *Cymbeline*, *The Winter's Tale*, *The Tempest*, *Henry VIII* and *The Two Noble Kinsmen*,

¹¹² See pp. 209-210.

and do not contain the marked densities of complexity features seen in that group of five. Quite what happens in the syntax of *Pericles* is as yet unclear.

The overall picture that seems to be emerging is that there are two separate phenomena at stake in the *which* syntax of the later plays in the canon. An increase in the density of *which* does occur in the later canon, beginning in 1605 with *All's Well That Ends Well*, and (despite an unexplained dip in 1607) this increase is sustained to the end of Shakespeare's output. But in addition to this, five years later in 1610, a further development occurs in his *which* relativisation, with the incorporation into his style of the series of features of increased syntactic complexity that we have seen in Chapter 4.

Only an examination of the individual plays in question will be able to flesh out the detail of this unexpected picture, and perhaps at the same time throw some light on the seemingly anomalous case of *Pericles*.

5.3 The individual plays

The data for the ten individual plays analysed is presented in full in Appendix 10. This ten comprises the original six (*Pericles*, *Cymbeline*, *The Winter's Tale*, *The Tempest*, *Henry VIII* and *The Two Noble Kinsmen*) plus the four newly-added plays (*All's Well That Ends Well*, *Antony and Cleopatra*, *Macbeth* and *Coriolanus*), which for the sake of clarity and brevity will be referred to here as the 'new four'. Salient features of the results are discussed here, in the same order as in section 5.1 above. Now that we are dealing with individual plays, the raw frequencies of the features under scrutiny are in most cases too small to give statistically significant results, so caution needs to be exercised in their interpretation. They do

however give us useful, more detailed accounts of the results for the groups of plays seen in Chapter 4 and section 5.1, and some explanation can now be forthcoming.

Function of *which*¹¹³

The drop in proportion of *which* types with relative function that we saw in 5.1 (p.213) seems to be the result of three of the new four having higher proportions of interrogatives than was the case in the original late six plays. This is especially so in *Macbeth*, where 8.75% (1 in 10) of *which* occurrences is an interrogative, as compared with a range between 0 and 4.17% (1 in 23) in the six original late plays. There is also an unusually high number of *who* interrogatives in *Macbeth*, perhaps suggesting a distinctive element in the dramaturgy of that play.

Overall relativisation¹¹⁴

Relativisation as a whole appears as a spectrum between a low point of 1.10% (1 in 91) of running tokens (in *Pericles*) and a high of 1.52% (1 in 66) in the *Two Noble Kinsmen*, with *Cymbeline*, *The Winter's Tale* and *The Tempest* all at the high end. Interestingly, *Macbeth* is also high, with 1.38%, although that play also has the highest relative frequency of *that* relatives, which might suggest a rather different syntactic profile.

The animacy parameter¹¹⁵

High relative frequency of animate antecedents was a marked feature of the late plays, but this difference disappeared in the comparison using the larger, high *which* group. Interestingly, we can see in the individual results that *Pericles* (one of the original late plays) has no animate antecedents at all, so amongst the other five post-1609 late plays the tendency

¹¹³ See Appendix 10, Tables 1 and 2.

¹¹⁴ See Appendix 10, Table 1.

¹¹⁵ See Appendix 10, Table 4.

towards animacy of *which* usage is statistically even more marked. These post-1609 plays have an average relative frequency of animate antecedents of 10.05% (1 in 10) of single word antecedents¹¹⁶, as opposed to the 7.02% (1 in 14) of the earlier thirty two. The new four show a range from 3.61% (1 in 28) to 6.25% (1 in 16), while the post-1609 plays band is from 4.88% (1 in 20) to 18.64% (1 in 5), although only one play in each band overlaps into the other range. Only *Macbeth* is above 4.48% in the new four (with 6.25%, or 1 in 16), and among the post-1609 plays only *Henry VIII* is lower than 7.69% (with 4.88%, or 1 in 20). Broadly speaking then, it would be true to say that the later group is characterised by an increase in animate antecedents, although *The Tempest* is by far the biggest contributor to that figure, with an extraordinarily high relative frequency of 18.64%. 1 in 5 *which* clauses in that play has an animate antecedent.

Closer examination of the occurrences in *The Tempest* sheds further light on the picture. A first hypothesis might be that this usage is restricted to Prospero and perhaps related to characterisation, since a number of the occurrences are in his speeches:

(151) No, not so much perdition as an hair,
 Betid to any creature in the vessel,
Which thou heard'st cry, *which* thou saw'st sink;

(1.2.37-8)

(152) Abhorred slave,
which any print of goodness will not take

(1.2.418-9)

¹¹⁶ As explained in 3.3.2.1, this set excludes sentential antecedents, to which animacy and inanimacy do not apply.

(153) No, wench; it eats and sleeps, and hath such senses
As we have, such; this gallant *which* thou see'st
Was in the wrack;

(1.2.465-7)

(154) Spirits, *which* by mine art
I have from their confines call'd to enact
My present fancies

(4.1.127-9)

The play's exploration of the nature of power manifests in one important respect in Prospero's domination of others, so it might seem possible that this animate use of *which* dehumanises the antecedent, disparagingly. As Hope points out, Shakespeare does indeed use *who* with animate antecedents elsewhere in *The Tempest* (1994, 39), suggesting that the *which* cases above could be marked usages. Certainly (152), and possibly (151), might be read this way, and in (153) the use of the impersonal pronoun "it" on several occasions prior to the *which* occurrence (as the nature of Ferdinand, this new being to appear, is discussed) has an almost identical effect, supporting the hypothesis. Prospero objectifies "spirits" in a similar way, again possibly underscoring his dominion. But he also uses *which* to refer to himself:

(155) That I am Prospero and that very duke
Which was thrust forth of Milan

(5.1.174-5)

as does Caliban, in (35), cited earlier:

For I am all the subjects that you have
Which first was mine own king

(1.2.406-7)

Ferdinand even uses it to refer to Miranda, the "goddess" (1.2.477) with whom he has fallen in love, and for whom he could hardly have higher regard:

(156) This my mean task
 Would be as heavy to me as odious; but
 The mistress *which* I serve quickens what's dead
 And makes my labours pleasures

(3.1.6-9)

We even see a reversal of the situation, as Shakespeare gives Ariel (in the form of the harpy) the relative “whom” in reference to what would now be conceived as an inanimate antecedent:

(157) the elements
 Of whom your swords are temper'd

(3.3.77-8)

The hypothesis that Shakespeare is using animate *which* to connote a character's disparagement of another does not, overall, seem to hold water. What can perhaps be said more generally is that the question of what constitutes humanity seems to be a preoccupation in *The Tempest*, so it might be appropriate for animacy to be an issue in that play. However this does not explain Shakespeare's wider tendency to use this waning variant with increased relative frequency after 1609.

Antecedent distance¹¹⁷

In the original analysis we saw a significant drop in the proportion of 2 words distant antecedents in the late plays corpus, and a significant increase in those with 5 or more words, with greater distance of course representing increased complexity. The individual results for the closer, 2 word distant antecedents reveal broadly two bands of figures, the new four plays in one, and the post-1609 plays in the other, although the ranges are not quite mutually exclusive. Three of the new four have between 13.25% and 15.05% (both 1 in 7¹¹⁸) of single

¹¹⁷ See Appendix 10, Table 5.

¹¹⁸ See section 3.0.

word antecedents that are 2 words away from the relative, and all the five post-1609 plays have between 2.56% (1 in 39) and 12.19% (1 in 8). *Macbeth* is very low in this feature (with 1 in 64), an exception to the new four band, and (interestingly, since this mirrors the distribution seen for animate antecedents) *Henry VIII* is highest in the post-1609 band (with 1 in 8).

Distant antecedents (5 or more words away from *which*) are found in a band from 2.41% (1 in 41) to 11.94% (1 in 8) in the new four, compared with 9.28% (1 in 11) to 15.25% (1 in 6) in the post-1609 plays. *The Tempest* has the highest density of these, although there is some consistency across the post-1609 group in this feature. *Antony and Cleopatra*, and particularly *Coriolanus*, have the lowest density of distant antecedents. It would be imprudent to take a single feature such as this and attempt to draw conclusions from it, but there is an interesting potential link here with McDonald's account of the language of *Coriolanus* as laconic and asyndetic (2006, 56). *All's Well That Ends Well* looks like a post-1609 play here, with 1 in 8 of its single word antecedents 5 or more words from the relative.

As far as *Pericles* is concerned, it has no occurrences of 2 word interventions, and only a density of 1 in 12 of 5 words or more (with only 2 occurrences in total), so it does not seem to belong to either the new four or the post-1609 plays.

Word groups intervening to the left¹¹⁹

The proportion of single word antecedents with phrases intervening decreased markedly in the late plays corpus, as clausal interventions rose, and there was some diminution of the difference in the high *which* analysis. The individual statistics give a clearer picture of what is going on here, however. The new four plays range from 14.06% (1 in 7) to 23.88% (1

¹¹⁹ See Appendix 10, Table 6.

in 4), although *All's Well That Ends Well* is responsible for the top figure there, and has almost double the density of the other three. The results are similarly diverse in the post-1609 plays. Three of them (*Cymbeline*, *The Winter's Tale* and *The Tempest*) have comparable proportions to three of the new four plays (at 1 in 6 or 7), but *Henry VIII* is very high – even higher than *All's Well* – with 24.39% (1 in 4), and *The Two Noble Kinsmen* is very low, at 7.69% (1 in 13). *Pericles* is the lowest of all, with 4.17% (1 in 24), (expected, since we saw a similar picture with antecedent distance, and these figures represent almost the same data). The low figures for *Pericles* would have weighted the results for the original six late plays towards the low end of the spectrum, making the drop in proportion of phrases to the left seem more pronounced than is the case in the group as a whole. Individual plays manifest this feature differently, so there does not seem to be a convincing pattern as far as the two sets of plays are concerned.

A clearer pattern exists in the case of the proportion of single word antecedents with clauses intervening to the left of *which*, however. The range of results for the new four plays is 10.84% (1 in 9) to 16.42% (1 in 6), compared with 14.56 (1 in 7) to 23.73% (1 in 4) in the post-1609 group, and the only overlap between bands is *All's Well That Ends Well*, which (as before) looks rather like a later, post-1609 play in this respect. This pattern is made even clearer in the breakdown of the data into non-finite and finite clauses, where in both cases the post-1609 plays show higher rates of occurrence. Furthermore, there is a distinct shift in the nature of the clausal interventions between the new four and the post-1609 plays. In the former clausal interventions tend to be partial finite clauses, whereas in the post-1609 group they are more likely to be non-finite or finite clauses. Again, *Coriolanus* has a very low density, so not only are there proportionally fewer distant antecedents in that play, but they also tend less

often to be clauses, supporting the characterisation of the language of *Coriolanus* discussed earlier¹²⁰.

These detailed statistics for the individual plays reveal the reasons why the second analysis, of high *which* as opposed to late plays, produced attenuated results. There is sufficient evidence here to show that it is the post-1609 plays, and not the new four, that have marked increases in both antecedent distance and the intervention of clausal elements. The exceptions are interesting, however. *Antony and Cleopatra* looks like a later play with respect to the type of clausal constructions intervening. And *Pericles* does not seem to belong to either group as far as interventions between antecedent and *which* are concerned, with only two instances of any intervention at all.

Continuative clauses¹²¹

Continuative clauses increased significantly in the late plays from 15.00% (1 in 7) to 19.95% (1 in 5) of relative *which* clauses, supporting the case for increased syntactic complexity in those plays. In the second corpus analysis on the other hand, when the new four high *which* plays were included, there was not even an attenuation of the increase, but rather a reversal of the direction of change to a decrease of 0.75%. As can now be seen from the individual statistics, it is the overall low frequency of continuatives in the new four that was the cause. *Cymbeline*, *The Winter's Tale*, *The Tempest* and *Henry VIII* have between 16.67% (1 in 6) and 27.08% (1 in 4) of relative clauses that are continuative. Only *The Two Noble Kinsmen* (10.87% or 1 in 9) has a lower density, inside the band of the new four plays, which

¹²⁰ See section 3.3.4.6.

¹²¹ See Appendix 10, Table 7.

range from 5.48% (1 in 18) to 15.66% (1 in 6). Yet again, *Pericles* is an exception in the group, with only 1 in 10 of its *which* clauses continuative.

Double pushdown constructions and recursions

Double pushdown constructions are examples of extremely complex relativisation syntax, and are very rare in the earlier part of the canon. There are only 15 occurrences in the entire earlier corpus (as used in Chapter 4), a rate of only 1 in 108 relative clauses. A highly statistically significant increase in Shakespeare's use of these constructions was found in the group of late plays, and now we can see exactly where these occur. There are no occurrences at all in either *All's Well That Ends Well* or *Macbeth*, and only one in each of *Antony and Cleopatra* and *Coriolanus* (rates of 1 in 106 and 1 in 96 respectively). However, *Cymbeline* has 5 (1 in 24), *The Winter's Tale* has 6 (1 in 20), and even the shorter texts of *Henry VIII* and *The Two Noble Kinsmen* have 2 and 3 respectively (with rates of 1 in 24 and 1 in 15), so this is clearly a distinguishing feature of the post-1609 plays. Once again, *Pericles* is an anomaly in the late group, with no instances of double pushdown constructions at all.

The individual statistics show in more detail the distribution of recursions across the group of plays under scrutiny here. Although the second corpus analysis seemed to suggest unexpectedly that they were spread evenly across the whole group of nine later plays¹²², that turns out now to be not quite accurate. They occur almost twice as densely than the earlier canon's 1 in 44 in *Cymbeline* (1 in 20), *The Winter's Tale* (1 in 29), *The Tempest* (1 in 23), and *The Two Noble Kinsmen* (1 in 23), so are arguably a feature of the post-1609 plays. However, they are found even more markedly in *Coriolanus* (5 occurrences in the play, or 1 in 19 *which* clauses), and that statistic alone may have been responsible for the increased, more

¹²² See page 216.

significant result in the analysis of the nine high *which* plays. And it would also have been helped by the fact that *Pericles* (yet again) is an anomaly, with no occurrences of recursions, so the fact that *Pericles* was removed from the high *which* corpus would have weighted the result in favour of that group.

The number of words intervening to the right

The halving of the proportion of 5 or more word interventions to the right of *which* in the six late plays¹²³ seemed to be balanced by a concomitant increase in 2 word interventions (non-finite clauses, as it later transpired), yet did not seem to be carried over into the second grouping analysis, suggesting that the new four plays did not share that combination of features. This suggestion is confirmed by the detailed data for the individual plays, where we can see a clear division between the groups in this respect. The new four have densities of 2 word interventions to the right of *which* in a band from 1.20% (1 in 83) of all *which* clauses to 3.12% (1 in 32), while the post-1609 plays range from 4.28% (1 in 23) to 8.34% (1 in 12). There is no overlap at all between the two groups, and the fact that *Pericles* only has one occurrence of 2 words to the right (3.45% or 1 in 29) confirms this to be a distinctive feature of the post-1609 plays.

Word groups intervening to the right¹²⁴

We are looking particularly here for details of three features that were pronounced in the late plays analysis, but not so in the second grouping: the drop in proportion of *which* clauses with phrases intervening to the right, the marked increase in non-finite clauses

¹²³ See section 4.5.1.

¹²⁴ See Appendix 10, Table 12.

intervening, and the increase in proportion of those with any kind of clausal (as opposed to phrasal) intervention.

The new four have a range of densities of phrases to the right from 13.70% (1 in 7) to 17.92% (1 in 6), whereas the post-1609 plays band is a lot wider, from 8.33% (1 in 12) to 17.39% (1 in 6). *The Tempest* and *The Two Noble Kinsmen* are within the band of the new four group, with similar phrase intervention densities to those found in the earlier 32 plays¹²⁵, although the other three (*Cymbeline*, *The Winter's Tale* and *Henry VIII*) have very low densities of phrasal interventions. Non-finite clauses are particularly marked in four of the post-1609 plays. The range of rates in the new four plays is between 1 in 73 and 1 in 32, compared with those in *Cymbeline*, *The Winter's Tale*, *Henry VIII* and *The Two Noble Kinsmen* of 1 in 11, 1 in 20, 1 in 12 and 1 in 11 respectively. Oddly, *The Tempest* does not seem to manifest either the low rate of phrases to the right, or the high rate of non-finite clauses. The phenomenon is, however, pretty clearly a feature of the other plays in the post-1609 group, and the very high statistical significance of the group data in Chapter 4 suggests that this is an important stylistic finding.

The proportion of *which* clauses with clausal interventions to the right of *which* is even more clear cut in its distribution. The new four range is from 1.37% (1 in 73) to 4.72% (1 in 21), with only the smallest of overlaps with the post-1609 plays band. Here the range is 4.28% (1 in 23, in *The Tempest*), to 12.50% (1 in 8, in *Henry VIII*). *Antony and Cleopatra* is the member of the new four that tends towards the late play profile as far as clausal interventions to the right are concerned, with 5 occurrences, or 1 in 21.

¹²⁵ See section 4.5.2.

Pericles and *Macbeth* are both particularly low in clausal interventions to the right, despite some deceptive appearances in the *Pericles* column, where extremely low raw frequencies (such as a total of four interventions) make for misleadingly high percentages. In *Macbeth*, all but one of its 11 interventions to the right are phrasal.

5.4 Chapter summary

We now have a much more complete picture of *which* relativisation in Shakespeare's late canon. From 1605, with *All's Well That Ends Well*, Shakespeare begins to write with a markedly increased density of *which* relatives. Although there is an inexplicable dip in this usage in the year 1607 (with *Pericles* and *Timon of Athens*) it goes on to become a plateau of sustained high density after that year. High density of *which* relativisation is a strong feature of increased syntactic complexity in itself, producing complex sentences with dependent clauses. In addition to this, however, there occurs between 1605 and 1609 a period of sporadic experimentation in the use of syntactically complex features of the *which* clause. It begins (also in 1605 with *All's Well That Ends Well*), with the use of complex antecedent relationships, a high density of sentential antecedents, and of those with (increasingly long) interventions between the antecedent head word and relative *which*. But we do not find those features prominent in another play until after 1609. Similarly in *Antony and Cleopatra* in 1606, Shakespeare experiments with high density of another feature of complex syntax: clausal interventions on either side of *which*. Again, we do not see this feature again until after 1609. In the same few years prior to 1609, Shakespeare also writes two plays that, while still having high density of *which*, are not distinguished by these features. *Macbeth* and *Coriolanus*

(perhaps as a result of their generic and narrative qualities) do not show the same experimentation in features to the left or right of *which* that were prominent in *All's Well That Ends Well* and *Antony and Cleopatra* respectively, although *Coriolanus* has an unusually high density of recursions in the *which* clause. From 1610, however, we not only see sustained high density of *which* relatives, but also from now on a strong tendency for several of the complexity features listed above to co-occur in every play.

It is not surprising that raw frequencies for these features in the individual plays are small. They are all unusual examples of intensely complex syntax, so their occurrence at all is noteworthy. But when they occur in markedly increased density in a very small group of plays, and when they also co-occur, so that four or five or even more of the complexity features appear prominently in a play together, there is undoubtedly something important happening in the syntactic style.

This shift in syntactic behaviour is not absolute – not a mechanical ‘off/on’ effect. In almost all the cases of complex relativisation behaviour seen above, there are one or two of the post-1609 plays that are exceptions and do not show high density. But as said before, all the features are examples of unusually complex syntax, and if for example in *The Tempest* Shakespeare does not use the same high density of clausal interventions to the right of *which*, we must be aware that he does use increased density of relativisation, distant antecedents, continuative clauses, and long interventions to the right in that play, so the overall picture is still of intensely complex relativisation syntax, and this simply cannot be said of the four plays written before 1608. Even more emphatically, it cannot be said in any way at all of *Pericles*. Although the six late plays were identified as a group at the beginning of this study, according to the consensus of late twentieth century criticism, they were quickly found to have this

stranger in their midst, all the more interestingly given the fact that *Pericles* is the only member of the late group written before the plague year of 1608/9. Whatever may have been discovered concerning the relationship of *which* usage to the other features of increased syntactic complexity, it is possible that *Pericles* could be an argument for there being some connection after all. Low in *which* frequency, it has either very few or none at all of almost all the features studied here. It is beyond the scope of this thesis to identify a cause for this outcome, but it surely also begs questions about the authorship of both that play and *Timon of Athens*, the other 1607 play with low *which* density.

Another interesting result from the individual analyses is the relationship between *Henry VIII* and the other play supposed to have been written in 1613, *The Two Noble Kinsmen*. Although there are some distinct similarities, they are often at opposite poles in the results tables. Generic differences may perhaps account for variations, but despite their undoubted membership of the post-1609 group in terms of features of complexity, there is a list of striking disparities in the syntactic style of these two plays. For example, *Henry VIII* is prominent for its high proportion of interventions to the left of *which*, and *The Two Noble Kinsmen* has about half that density; *The Two Noble Kinsmen* has a very high density of relativisation overall, whereas *Henry VIII* does not; *Henry VIII* has a high rate of continuative clauses, but *The Two Noble Kinsmen* has a very low rate. There are still some doubts about the dating of *Henry VIII* (Wiggins, personal communication), so perhaps these findings suggest that the doubts may be justified.

CHAPTER 6

CONCLUSIONS

6.1 Review of method

The WordSmith Tools keyword analysis was found to be a very successful method of identifying saliency in the corpus under scrutiny, and could have been taken further with an examination of other keywords, although the prominence of *which* in the list was felt to be sufficient justification for devoting this thesis to its investigation. The limitations of the keyword analysis are clear, however. Firstly, frequency software such as WordSmith Tools cannot identify absences or deletions such as the zero relative, so it has not been possible here to compare zero frequency with that of the other relatives. Without a much more sophisticated text tagger and parser than has yet been developed, this would be impossible by electronic means, and in order to produce comparable statistics for zeros the plays would have to be scrutinised manually. As far as the resulting statistical limitations are concerned, although a complete picture of Shakespeare's relativisation would require the incorporation of statistics for zeros, their absence in these figures does not undermine any of this data. The increases in his use of the relatives discussed are not affected by the presence or otherwise of zeros in the count. Furthermore, zeros are a proportionally less significant feature of Shakespeare's language than that of Fletcher, for example, as shown by Hope in his authorship attribution

study of six plays from the end of the canon, namely *Antony and Cleopatra*, *Coriolanus*, *Cymbeline*, *The Winter's Tale* and *The Tempest* (Hope 1994, 38). It was therefore felt that the advantages of corpus analysis, particularly its facilitation of analysis of all the texts in their entirety, outweighed the disadvantages.

Similarly, the computer can only identify keyness on the basis of frequency of single items, and there are numerous syntactic features with the potential for increasing syntactic complexity that could not be identified by such means. Similarly, the syntactic formations considered in this thesis can also occur with paradigmatically parallel word types (indeed *whose* and *whom* are both keywords in the late plays list, although much less key than *which*). Continuative clauses starting with the other *wh-* forms are missed here, for example, as are double pushdown constructions with those relatives, and so on. The detailed analysis of *who*, *whom*, *whose* and *that*, and even the temporal and locative *wh-* forms *when* and *where*, would therefore be an important extension of this research, and could develop understanding of the features identified in a number of ways. An additional benefit of extending the field under scrutiny to all *wh-* forms, or all relatives, would be the increase in size of the raw frequencies for each feature in each play, thus making the relative frequencies more reliable, and patterns more easily mapped.

The related question arises of what exactly is represented by the statistics in terms of style. What does it actually mean, for example, that in *All's Well That Ends Well* there are 7 partial finite clauses and only 2 finite clauses intervening to the left of *which*? The emphasis in this study has been on comparative interpretation, and that seems appropriate in this case, where we are considering the hypothesis of increased late complexity. However, it is still important to ask what these levels of frequency might mean. Is there perhaps a case for saying

that they are, on the whole, not sufficiently frequent to be significant in stylistic terms? In fact, authorship attributionists working with lexical words often use features with similar occurrence rates to the grammatical features in the present study, and find them not only significantly frequent but determining of authorship, as can be seen for example in the work of Craig and Kinney (2009, e.g.18-19). Very low raw frequencies are nonetheless a problem, the limitations of which are acknowledged throughout this study when they appear, but it is argued here that it is appropriate to accept, for example, six cases of large, multi-word phenomena such as double pushdown constructions in a full length play such as *The Winter's Tale*, or three in a short, part-play such as the Shakespeare sections of *The Two Noble Kinsmen*, as evidence of complex syntax because they are remarkable and unusual constructions. And of course when the plays are viewed as a group the changes in density of features such as this have been shown to be highly statistically significant.

Another limitation of the computer analysis is its inability to take account of verse lineation, or the difference between verse and prose. Such concerns form a crucial component of the analysis of Shakespeare's style, and can contribute importantly to the picture of the late syntax, as McDonald has shown (2006, e.g. 96-99). Although, as explained in sections 3.2 and 3.3, WordSmith Tools offers the possibility of expanding any concordance line to view any amount of co-text, it is not yet possible to conduct searches on the basis of the position of words or phrases in the line. The result is that versification and lineation can only provide supporting evidence, rather than being determining factors in the search and analysis. The present research would be very usefully extended by the consideration of the frequency and nature of occurrences of the chosen features in either verse or prose.

Positive aspects of this method are that it is initiated by an empirical test (the keywords analysis) and the findings of that analysis were statistically very reliable. Also, a major strength of this research is that the manual analysis scrutinises the entire body of data involved, all 2,207 *which* clauses in the canon, for all the categories chosen, instead of sampling, as is more often the practice elsewhere. In addition to this, Shakespeare's other relatives were counted and analysed manually in terms of their main function. This thoroughness and comprehensivity in the (localised) area of relative clauses means that although the results may be subject to occasional individual human error, as is always the case in such situations, there is a greater overall likelihood of accuracy in the statistics. The problem of this approach, however, is the long-windedness of the manual analytical and counting methods. The 2,207 clauses had to be revisited individually many, many dozens of times in order to arrive at the results listed here, something unlikely to be undertaken by busy professional scholars (see for example Smith 1990, 199). Computerisation made the study possible in the first instance with the keyword analysis of the late plays, but would not have been able to perform the subsequent analyses of individual concordance lines.

A further limitation of the present research method is that, because it is dependent on numerical frequency of predetermined categories, it does not provide as sensitive a measure of degrees of complexity as an account of individual occurrences on a case by case basis might offer. Such measurement is beyond its scope. And there are, necessarily, cases excluded from the count where there is complexity in the relative clause, for example where length or clausal interventions occur to the right of *which*, but after the next principal item (the subject, verb or object).

The analytical framework is a good gauge of the syntactic complexity of *which* usage, but by no means a perfect one. What it has shown, however, is reproducible evidence of *some* features of increased complexity.

The selection of categories in an analysis such as this is of course the product of individual and potentially subjective intervention, so requires full explanation and justification. Existing linguistic scholarship was the principal determining factor in the choices made. This deductive approach was employed in the assignment of the categories of antecedent type, antecedent distance, the information parameter, continuative clauses, subject, object and prepositional complement roles in the relative clause, the number of words intervening between *which* and the next principal grammatical item in the relative clause, and clausal length. All of these were selected because one or more commentators identified a potential for increased complexity in their usage. It was, however, necessary to design counting procedures for words intervening to the left and right of *which*, and descriptive categories for word groups intervening in the same ways. Continuative clauses, whose loose link between antecedent and relative is observed by a number of historical linguists (such as Ryden, 1966 and Rissanen 2008) and agreed to be a source of syntactic complexity¹²⁶, proved challenging to incorporate as an analytical category, since there was no existing agreed method of identifying these clauses reliably. This study responded by establishing a set of criteria whereby continuative clauses can be identified with 93.5% accuracy, and proceeded to count all the occurrences in the canon (in the same way as all the other features analysed).

An inductive approach was employed at times, in response to the occurrences of *which* themselves. On occasions when syntactic phenomena could not be accounted for by existing

¹²⁶ See 3.3.3.2.

models, categories had to be designed in order to accommodate them. This was the case to some extent with both single and double pushdown constructions, recursions, and with the word groups intervening to both the left and right of *which*. There have also been some occasions in this thesis when the necessity for measurability and numerical specificity meant overriding more subtle syntactic or formal description, such as in the case of designating the head of the antecedent noun phrase as the antecedent. Although the limitations of this kind of practice are clear, such decisions were made with the minimum compromise of linguistic accuracy. For example, in the case of antecedents, if a reduction is required for computational purposes, the reduction to the head of the noun phrase seems the most legitimate and defensible option. Decisions of a problematic kind such as this one are always acknowledged where they occur. Similarly, as pointed out in 3.3, categories employed in this study reflect modern linguistic scholarship, rather than suggesting that they reflect an analytical framework that would have been used, either consciously or unconsciously, by Shakespeare himself.

Finally, in the foregoing thesis the term “complexity” has been used without qualification, both by the present writer and a wide range of linguists and literary critics, but the notion itself is somewhat problematic. The term can depend on the familiarity of the reader or auditor with the construction in question. For example, some cases of *which* in a recursive role in the relative clause can surely be fairly easily processed by modern readers, such as the following from *Coriolanus*:

(160) There’s one thing wanting, *which I doubt not but*
Our Rome will cast upon thee

(2.2.110)

Yet cases such as this are some of the most complex, in syntactic terms, in the whole set of *which* occurrences. To clarify then, in the present thesis the term “complexity” is used in a more strictly syntactic than cognitive, psychological sense. Having made that distinction, however, the very frequency of clauses with loose relative link in the Early Modern period may well suggest that such constructions were not perceived as complex by their readers, or indeed auditors in Shakespeare’s theatre. What is in question in this study, however, is not the Early Modern but rather the modern perception of complexity. It has been made clear from the outset that observations made by literary critical writers of the twentieth and early twenty-first centuries are the basis for this investigation. It is impossible to reconstruct the perspective of an Early Modern reader, and there would be little to be gained by the attempt. Furthermore, the observation is a relative, comparative one. The perception of literary critics is that the language of the late plays is more complex than that of the earlier canon, and that has been the underpinning assumption of this thesis throughout, upon which the present analytical procedure has been founded.

6.2 Overview of findings

Firstly it must be pointed out that the present thesis only addresses one narrow aspect of the style of Shakespeare’s late plays, and it cannot therefore make definitive or comprehensive claims on the question of the constitution of the late group. What it can do, however, is direct attention to a set of features that seem to be representative of the late style, or styles, and draw distributional conclusions on that basis.

Corpus analysis of the six late plays *Pericles*, *Cymbeline*, *The Winter’s Tale*, *The Tempest*, *Henry VIII* and *The Two Noble Kinsmen*, in comparison with a reference corpus of

the previous thirty-two plays in the canon, revealed the highly statistically significant keyword *which*. Research into the background linguistic usage of the early seventeenth century revealed that this was not simply a reflection of a contemporary trend: in fact the reverse was true. The background tendency has been shown (by, for example, Dekeyser 1984) to be in the other direction, as the animacy parameter gained ground and *who* increasingly replaced *which*, suggesting that Shakespeare's increasing use of *which* was probably an individual stylistic feature.

An analytical framework was then designed in order to investigate in detail the behaviour of *which* as a relative in the late plays, again comparatively with the earlier canon. Six general aspects of *which* usage were identified as categories for analysis. These were

1. The main function of *which* (as interrogative or relative, principally)
2. The relationship of *which* with words to the left (with its antecedent)
3. The information parameter (distinguishing restrictive from non-restrictive relative clauses)
4. The role of *which* in the relative clause
5. The relationship of *which* with words to the right
6. The length of the relative clause

The density of occurrence of features of syntactic complexity was measured in terms of the proportion of all relative *which* clauses¹²⁷ where each feature occurred, and (where appropriate) in terms of its rate of use.

¹²⁷ Sometimes, where appropriate, the proportion of all single word antecedents was the measure, rather than all relative *which* clauses. See 3.3.2.3 and Table 4.2.

This analytical framework was applied in exactly the same way on three different occasions. Firstly it was used to compare the six late plays (listed above here) with the previous thirty-two. Secondly, in response to a need to clarify the composition of the late group, it was applied to an analysis of nine plays from the late part of the canon, all of which have markedly high density of the type *which*. This group comprised five of the six already listed, (with *Pericles* removed because it was found to have low density of *which*), and also the four high *which* density plays *All's Well That Ends Well*, *Antony and Cleopatra*, *Macbeth* and *Coriolanus*. Thirdly and finally the same analysis was applied to ten individual plays, namely *Cymbeline*, *The Winter's Tale*, *The Tempest*, *Henry VIII*, *The Two Noble Kinsmen*, *All's Well That Ends Well*, *Antony and Cleopatra*, *Macbeth*, *Coriolanus* and *Pericles*, in order to establish the details of the distribution of the complexity features in question.

As described in outline in 5.3, the findings of these three analyses produce an overall picture of the development in Shakespeare's relativisation syntax over the last eight years of his career. Beginning in 1605 (with *All's Well That Ends Well*) his use of relative *which* increases markedly, with some concomitant overall increase in relativisation. In 1607 there is a temporary unexplained dip in *which* density, in *Pericles* and *Timon of Athens*. Since both these are dual authored plays, textual questions are raised by these findings, which would need to be answered before further interpretation is ventured.

In addition to *which* density, a raft of features of more complex relativisation syntax was discovered in the late plays, but (as Chapter 5 established) there is no sudden shift from one stylistic phase to the other. Between *All's Well That Ends Well* in 1605 and *Coriolanus* in 1608 it seems that (although the raw frequencies involved are small) Shakespeare experiments with a range of features of complex relativisation, as they appear more markedly in one play

but not the next. It is not until after the break in output of the plague year 1608/9, not until *Cymbeline* in 1610, that we really see the plays where many of these features appear together, producing the highly complex syntactic style that has been sensed by so many critics.

In *All's Well That Ends Well* Shakespeare increases the density of his *which* usage, and experiments with more complex antecedent relationships for *which*. The play is not mentioned by McDonald, yet as we have seen it does have some late characteristics as far as its *which* syntax is concerned, namely high frequencies in a group of features all related to antecedent relationships. It has the highest density of sentential antecedents of all the plays, the second highest (after *The Tempest*) of distant antecedents, and the third highest (after *Henry VIII* and *The Two Noble Kinsmen*) of antecedents with multiple noun phrases. (This play also has a high proportion of longer interventions to the right of *which*, which could be a source of increased complexity, but was not in fact found to be typical of the late plays¹²⁸). The relatedness of these features seems to suggest a stylistic pattern, one of a consistently more complex relationship with the antecedent in that play, suggestive of early experimentation with what would become part of the late style. It may well be significant that high *which* density occurs first in a play that has many hallmarks of the romance genre, since the correlation between these two seems to characterise the late phase. *All's Well That Ends Well* is the only one of the four additional, high *which* plays to be related to the original late group in terms of narrative content. Although given a dark and ironic treatment, there are substantial romance features in the work, of the kind that are strongly prevalent in the last five plays (the folk tale elements of the miraculous healing of the king, and the wish-granting that results in the choosing of the husband, for example), and the fact that it reveals some stylistic links with

¹²⁸ See section 4.5.

those five late plays proper would seem therefore potentially significant. The connections must not, however, be overstated. In all other respects it is certainly not a member of the late group. There is no evidence of increase in clausal interventions, not even in the antecedent relationships, nor is there an increase in double pushdowns or recursions.

Although *which* density is still high in *Macbeth*, the next play he wrote, a higher proportion of those occurrences have interrogative function, and there is no prominence of the features seen in *All's Well That Ends Well*¹²⁹. Both *Macbeth* and *Coriolanus* are almost always at the low end of the frequency range in the complexity features considered here, and even though there is consensus among scholars to the effect that they were written two years apart (in 1606 and 1608 respectively) they do seem to have some similarities in their syntax. McDonald's view that *Macbeth* is characteristic of the late style in terms of repetitious sounds and sequences (2006, 50) does not hold true in the case of most of the features analysed in this thesis. Both plays are, of course, high in *which* frequency, so there is still some corroboration of McDonald's account of Shakespeare's "additive impulse" and numerous "append[ed] relative clauses" in *Macbeth* (2006, 46), as a typical feature of late style. But McDonald also says of *Macbeth* that its language is "compressed" and "reduced to essentials" (43-4), and although the play lacks the syntactic compression brought about by replacement of longer phrases by shorter clauses, it is tempting to suggest that what McDonald sees as the "laconic, separated speech" (57) of *Coriolanus* himself – blunt, unconnected and mistrustful of verbosity – which he argues is "coded historically as a masculine style" (57), in some way may prevail more generally in both these heavily masculine tragedies, resulting in a resistance

¹²⁹ *Macbeth* was also found to have a markedly high density of restrictive clauses associated with demonstrative pronouns such as *this* NP *which* or *that which*, but this was not felt to be sufficiently pertinent to the present analysis to include it in the thesis.

to the other features of complex syntax identified with the late romances. In *Coriolanus* the marked feature is the proportion of clauses that have recursive structures¹³⁰ involving *which*.

Antony and Cleopatra appears in 1606, the same year as *Macbeth*. McDonald identifies internal stylistic dynamics in *Antony and Cleopatra*, describing contrasting syntactic behavior associated (as a feature of characterisation) with Cleopatra on one side and the Romans on the other, with Antony vacillating (linguistically and otherwise) between the two poles. Obviously this analysis cannot comment in any detail on such an account, given the narrow, specialised focus here. Remarks can be made, however, about the relationship of the play to the rest of the group of ten later plays, on the basis of *which* syntax. *Antony and Cleopatra* is markedly not a member of the post-1609 group in terms of its antecedent relationships, with the lowest or almost lowest frequencies of distant antecedents, sentential clauses, multi-NP antecedents and animacy usage. It is similarly very low indeed in continuative clauses, recursions and double pushdown constructions. In this play, however, the early experimental emphasis seems to be on activity to the right of *which*, with an increase in the proportion of interventions between *which* and the next principal item in the clause, and in the replacement of phrases with clauses, both to the left and right of *which*. It is not prominent in these respects: about half way between highest and lowest frequencies, in fact, but it is in the post-1609 plays band in both cases. It is also, like *All's Well That Ends Well*, high in the number of words intervening to the right of *which*, although, as pointed out, this was not found to be a post-1609 feature.

What we see from 1610 onwards however, starting with *Cymbeline*, is a development

¹³⁰ See 3.3.4.6, where the category of recursions is exemplified by cases from *Coriolanus*.

and intensification of many complexity features in the *which* clause, along with a sustained high density of *which*.

Interventions between *which* and its antecedent in the post-1609 plays are more likely than cases in the earlier canon to be five or more words long. They also show a marked increase in clausal, as opposed to phrasal, interventions between relative and antecedent, and both of these features are strongly identified with syntactic complexity (for example see Rissanen, as above, and Berlage 2008). The increase in the use of *which* for animate antecedents is a distinctive feature of the last five plays, and although probably not related to complexity, connects significantly with other work in this field (Hope 1994, 27-53). In these plays Shakespeare is markedly less likely to use determiner *which*, and its use becomes entirely restricted to continuative clauses.

Continuative clauses are found to be significantly more likely to occur in the last five plays than in earlier works, and much more likely to be used with sentential antecedents than are restrictive or ordinary non-restrictive clauses, and they are also much more likely to have antecedents five or more words distant from the relative than non-continuatives. Thus three features of complex syntax are shown to co-occur significantly, representing an intensification of complexity in the last five plays. There is some evidence that in these last five plays Shakespeare leaves behind fixed syntactic constructions such as cataphoric *which* and partial finite clauses intervening to the left, preferring the freer but more complex *which* usage typified by other findings below here.

It is in the role of *which* in the relative clause, however, and in the related category of interventions to the right of *which*, that the most pronounced features of late style were found.

Clauses where *which* has a reduced role in the main relative clause, such as relativiser-only *which*, and single pushdown constructions, decrease in relative frequency in the last five plays. Correspondingly, those constructions where *which* has an increased syntactic role in the relative clause, namely those identified here as double pushdowns, increase their density very markedly. There are more double pushdown constructions in the last five plays than in the entire previous canon: their proportion increases by over 300%. Along with double pushdown constructions, possibly the most syntactically complex form of *which* function in the relative clause is that found in cases of recursion, because the relative is part of a rank-shifted clause inside the main relative clause, complicating the anaphoric relationship to the left, as well as the syntactic relations inside the relative clause. Recursions also show a marked increase in density in the late canon, particularly in the last five plays, (although, as already mentioned, they also feature prominently in *Coriolanus*). In *The Two Noble Kinsmen*, *The Winter's Tale* and *Cymbeline* almost one in ten *which* clauses is either a double pushdown or a recursion, a considerable indicator of complex syntax.

As far as the number of words intervening to the right of *which* is concerned, there was a slight decrease in the proportion of interventions overall, but strikingly, the number with 5 or more words intervening was halved. This might seem to imply both a simplification of syntax, with *which* closer to the next grammatical component of the clause, and also that the length of Shakespeare's late relative clauses decreased compared with the earlier canon. It has, however, been shown (in Section 4.6) that there is no appreciable change in the length of relative clause in the late plays. It must logically follow, therefore, that the section of the relative clause *after* the next principal grammatical item to the right of *which* tends to be longer in the post-1609 group, and further research could investigate this.

If syntactic relations between *which* and the next grammatical item to the right are not typically affected by increased numbers of words in the post-1609 plays, what does affect these relations however, and changes very substantially in the late group, is the nature of the intervention. Along with the increase in double pushdown constructions, the most marked feature of late style found in this study is the replacement of phrases with clauses, and particularly non-finite clauses, in interventions to the right. The increased syntactic complexity afforded by clausal as opposed to phrasal linguistic material is widely noted (for example by Rissanen 1984, 420, and Berlage 2008), and we find this very markedly here, dispatching any notion of a simpler syntax in relations to the right that might result from a decrease in distance. The marked tendency to use shorter, clausal interventions instead of longer, phrasal ones can also be seen as corroboration of some critical observations of the late style such as Orgel's "intensity" (1996, 12) and McDonald's "condensation and intensification" (2006, e.g.93).

Some aspects of relativisation associated by historical linguists with syntactic complexity were not found to increase in the late plays. Keenan and Comrie's Accessibility Hierarchy (1977) would suggest that if syntactic complexity is a feature of late style the function of *which* in the relative clause might be expected to reveal a shift towards the more complex cases. In fact there was a slight shift towards subject function, although too variable to be an identifiable late feature. Another aspect of *which* usage seen as a "measure of syntactic complexity" (Dekeyser and Ingels 1988, 33), but not found to reveal any appreciable variation in the late plays, was the sixth category listed above, clausal length.

One of the main findings of the individual plays analysis was that *Pericles* is a stranger in the late group. As was pointed out in section 1.1 earlier, *Pericles* has been seen as part of the late group since its first identification, almost certainly because of the romance narrative in

the play, which links it with *Cymbeline*, *The Winter's Tale* and *The Tempest*. McDonald declares that “Pericles, Cymbeline, The Winter's Tale and The Tempest undeniably [resemble] one another” (2006, 22). With respect to its *which* relativisation syntax, however, it would have to be argued here that *Pericles* certainly does not resemble the other five plays, as can be seen in the tables of results in Appendix 10. The play entirely lacks the pattern of a number of syntactic features to which the five plays written after 1609 conform. Only the three last acts were included in this analysis, all of which are firmly believed to be by Shakespeare (Vickers 2002, 291-332), so this anomaly raises some very interesting questions, which future research may wish to address. Further insight into the syntactic style of 1607 might be gained by an analysis of *Timon of Athens*, the other play written in that year (according to Wiggins et al), to see whether similar results appear there. We know already¹³¹ that *Timon of Athens* has an unusually low *which* frequency at 0.28% of tokens, very similar to *Pericles* at 0.29%, so it would be very interesting to see whether the correspondences extend further. As already suggested, though, the anomaly may also be the result of inaccurate authorship attribution.

Many recent writers (e.g. McDonald 2006, 24-5, Lyne 2007, and Richards and Knowles 1999) include *Henry VIII* and *The Two Noble Kinsmen* in the group only reluctantly, citing various differences of genre, thematic content or even style as their reasons. But McDonald is right that “at the level of sentence and line” (2006, 26) they are typical, and sometimes exceed even the four romances in their relative frequency of complexity features. They appear next to each other in the hierarchy of density of distant antecedents (third and fourth out of ten plays) and multi-NP antecedents (first and second) and they both have extremely high frequencies (the highest in the group) of clauses intervening to the right.

¹³¹ See section 5.0.

Where they differ in other respects, however, they are often at extremes of the spectrum, and there are many more of these cases. *Henry VIII* is prominent for its high proportion of interventions left, while the other play has only half the percentage. *The Two Noble Kinsmen* has the highest proportion of clauses intervening between *which* and antecedent, and the highest proportion of double pushdown constructions, both in the entire set of plays. *Henry VIII* has no recursions at all (although a fairly high proportion of double pushdowns), but has the highest number of continuative clauses of all. *The Two Noble Kinsmen*, on the other hand, has a very low rate of continuative clauses.

Otherwise, we see most often the three romances *Cymbeline*, *The Winter's Tale* and *The Tempest* alongside each other in their frequency patterns, particularly in the features listed above as most typical, something that seems to confirm the perceptions of over a hundred years of critical responses. Despite some individual features, such as marked use of animate antecedents in *The Tempest*, these plays have a great deal in common in their relativisation syntax, showing the most marked trend in their movement towards clausal density and complexity on either side of *which*, in addition to overall frequency of relativisation and more complex and distant antecedent relationships.

In summary, after 1609 the co-occurrence in each play of several features from the above set creates the effect of an intense aggregation of syntactic complexity. In one play after another we see multiple combinations of distant antecedents, continuative clauses, recursions, double pushdown constructions and the use of clauses rather than phrases on either side of *which*. As relative *which* clauses increase in frequency as the canon progresses, the proportion of such clauses that are highly syntactically complex also increases, accentuating the resulting effect. The late syntactic style is definitely complex.

6.3 Related studies

It seems appropriate at this point to discuss a group of other studies of function words in Shakespeare's plays which have a close relationship to the present thesis, and whose findings can now be compared usefully with those presented here. They are all authorship investigations, and they all concern *which* in the late group.

Firstly, as has been demonstrated here, *Pericles* is an interesting anomaly in the late group, whose *which* relativisation syntax is atypical of the late group in almost all respects. An interesting addition to this picture is offered by Honigmann's discovery of what he calls "striking mannerisms" (1965, 195) in the syntactic style of Wilkins (Shakespeare's collaborator in *Pericles*), found in his novels *The Painful Adventures of Pericles* and *Justine*. Honigmann describes the way that Wilkins "repeatedly begins sentences, and clauses after a colon or semi-colon, with 'which' or 'by which', 'to which', 'of which' etc" (1965, 194) and cites the following passage from the (Wilkins authored) first act of *Pericles*, where he also identifies these features:

Cleon: The which when any shall not gratifie,
Or pay you with vnthankfulnesse in thought,
Be it our Wiues, our Children, or our selues,
The Curse of heauen and men succeed their euils;
Till when, the which (I hope) shall neare be seene:
Your Grace is welcome to our Towne and vs.
Peri. Which welcome wee'le accept, feast here awhile

(1.4.101-7¹³²)

All three occurrences of "which" here are continuative clauses, the first being also a single pushdown construction, with "which" as object in the lower element "when any shall not gratifie". Jackson takes up this observation, using the feature as a test for authorship in the

¹³² Cited in Honigmann 1965, 195.

play, labelling this continuative usage of *which* “resumptive” (1990, 193) or “connective” (2003, 123). Jackson’s analysis successfully identifies three specific habits of Wilkins’s writing that differ from Shakespeare’s (2003, e.g.128), so the authorship test is a useful one, and it would certainly seem to corroborate the findings of the present research in suggesting that *which* usage is a stylistic marker of the time. Unfortunately the choice of categories is such that the fuller picture of syntactic style is missed (although a full account of syntax was not, of course, the object of Jackson’s study). His first category, where the *which* usage is paraphrasable by “and this” (123), equates with the first criterion for continuativity employed in this study, but excludes all other cases except determiner *which*. The second is the waning variant *the which*¹³³, but the third simply requires *which* to be speech-initial, and is not based on a syntactic principle. The result is that although he trawls the Shakespeare canon for the categories he has chosen, Jackson does not find there what he calls the “incidence of unusual *which*” associated with Wilkins (2003, 126-7). Ironically, he is working on the very play in the late group that does not manifest the complex late *which* syntax, the play whose final two Shakespearean acts are in so many ways anomalous in the group. In the plays written only a year later, there is “striking” and “unusual” *which* usage in plentiful supply.

Secondly, Jonathan Hope’s 1994 study *The Authorship of Shakespeare’s Plays* arrives at some of the results of this thesis, as by-products of a quest to find reliable authorship tests for collaborative plays both in the Shakespeare canon and the apocrypha. Hope sets out to establish consistent norms for linguistic behaviour in plays known to be by Shakespeare and the other authors in question, in order then to test the debated authorship plays against those norms (1994, e.g. 34). He finds “auxiliary do” to be a reliable norm, and successfully

¹³³ See 3.2.

establishes the authorship of a number of different plays on the basis of this test. He then analyses the different authors' use of relative markers with the same aim, but these prove less reliable because used less consistently, especially by Shakespeare. However, the results produced in the process have some interesting correlations with this research.

There are two sets of Shakespeare plays in his samples: one from the early part of the canon, and one from the late. The first group comprises *Two Gentlemen of Verona*, *The Comedy of Errors*, *Love's Labour's Lost*, *Richard II*, *A Midsummer Night's Dream* and *The Merchant of Venice*, and the second *Antony and Cleopatra*, *Coriolanus*, *Cymbeline*, *The Winter's Tale* and *The Tempest*. Although the present study records *which* frequencies across the whole canon in terms of percentage of total tokens, rather than Hope's method of showing the percentage of all relatives represented by each form (e.g. 35), the results are closely comparable to those in Hope's work. He finds that although Shakespeare's *which* usage in the earlier group is higher than that of Fletcher, it is "even higher" in the late group (1994, 42). It is worth noting however that he does not of course include the two plays with lower *which* frequency, *Pericles* and *Timon of Athens*, in the sample (because they are in the contested authorship group). Specifically, he also finds "strikingly anomalous" low usage in the very early play *The Comedy of Errors*, "much lower than in any other play of [that] date" (40). The data for *which* frequency presented here in Table 5.1 seems to confirm Hope's impressions of experimental relative usage in the early canon (40), since *which* frequency is certainly very variable between the plays. However, *Love's Labour's Lost* in his sample is rather an anomalously high *which* play¹³⁴, perhaps misleadingly so for Hope. Also, plays outside his sample put the picture in more context: *Titus Andronicus* has the same low *which* frequency as

¹³⁴ See 5.0.

The Comedy of Errors, with *The Taming of the Shrew* even lower, and the frequency of *which* continues to fluctuate considerably until the late period, as discussed in 5.0 above. Importantly, Hope sees Shakespeare's relativisation syntax as "clearly show[ing] a shift from 'that' and zero to 'who' and 'which' (that is, the more formal end of the scale) over the course of Shakespeare's career" (43), which corresponds precisely with the findings of this research, and he posits the idea of a generic shift being the cause of the changes (43), interestingly approaching suggestions made here, below.

Finally, an absorbing addendum to the present account of *which* in the late plays is provided by a recently published study of Shakespeare's linguistic style, again from an authorship perspective, by Craig and Kinney (2009), already referred to here on several occasions. In the present thesis, as explained in section 2.3, it proved necessary to use the Craig 1914 edition of the plays. One of the limitations of that choice was that Craig (like most other editors) offers a single text of *King Lear*, thus preventing the analysis in this research of the two different versions. Although the play is believed to have been originally written in 1605, in the version published three years later in Quarto and entitled *The History of King Lear*, a different text called *The Tragedy of King Lear* appears in the 1623 First Folio, generally thought to have been written in 1609. Considerable discussion surrounds the differences between the two texts, particularly with regard to the reasons for these (see for example Wells and Taylor 2005, 909; Craig and Kinney 2009, 181-194; and Foakes 1997, 128-132). Given its date, the Folio *King Lear* would have been a candidate for the late plays group, and so very much of interest in this study. Craig and Kinney (2009, 181-201) analysed the textual material that is in the Folio *Tragedy* but not in the earlier Quarto *History*, using software focusing on lexical and grammatical word frequency and distribution, in an attempt

to answer the question of the identity of the author of the new material. They reveal the following very interesting findings: “Repeatedly...and quite evenly through the various different sections of the play, the person(s) responsible for F uses *which* where Q has *that*” (2009, 194). And they provide details in a footnote as follows:

Fifty-five instances of *which* as a relative are common to the two versions. Q has nine not in F, all of them in sections of the text that are not in F; Q’s total is thus sixty-four. Though shorter, F has fifteen instances not in Q for a total of seventy instances. In ten cases, F has *which* where Q has *that*; three are in sections not in Q, in one case F has *which* where Q has *who*, and in another, F has *which* where there is no relative in Q. This pattern, consistent in all five acts, implies one author preferred *that*, *who* or a relative deletion to *which* as a relative, and a second author, or the first author at a different time of composition, preferred *which*.

(Craig and Kinney 2009, 194).

This shift towards preference for *which* is precisely the trend identified in this research, and by Hope (1994, 43). The work of Witmore and Hope (2006) is also corroborated by the findings of the present thesis concerning the high frequency of *which* in the late plays. Craig and Kinney’s results, however, connect even more closely with this research. As we can see from their report here, not only does the reviser (who they confirm to be Shakespeare) increasingly use *which* and prefer it over both *that* and zero relatives, going against the background linguistic trend demonstrated by Dekeyser¹³⁵, but there is also evidence here of *who* being replaced by *which*, corroborated by the results for *which* usage and the animacy parameter in this thesis. So those critics who argue for connection of the *Tragedy of King Lear* with this late group may well have a point¹³⁶. There can, I think, be no doubt that in frequency and usage of *which*, a distinctive feature of Shakespeare’s late style has been found.

¹³⁵ See 3.2.

¹³⁶ See, for example, Lyne (2007, 25) and Richards and Knowles (1999, 2), discussed in 1.1 above.

6.4 Further interpretation

6.4.1 General comments

In *Shakespeare's Late Style*, Russ McDonald is uninhibited in his extrapolation of broad and wide-ranging narrative and dramatic structures (e.g. 40, 100), even of aspects of Shakespeare's biography and psychology (e.g. 52, 66 and 69), from sentence level linguistic features. The present thesis is more guarded in such matters, particularly in light of certain debates in stylistics that have (probably productively) left their mark on the discipline. In a widely-cited and influential critique of stylistics, Stanley Fish argues that he finds in some stylistic analyses "the absence of any constraint on the way in which one moves from description to interpretation, with the result that any interpretation one puts forward is arbitrary" (1973, 96). Although his stance is rather all-or-nothing as far as interpretative 'proof' is concerned (as Toolan (1990) points out), this is undoubtedly a useful caution, and will be heeded here in making connections between Shakespeare's *which* usage and wider issues. It does however seem important to consider, albeit briefly, aspects of the wider meaning and function of the findings of this thesis.

6.4.2 The dramatic function of complexity

The discourse structure of drama is complex. As Short argues, it has "at least two levels of discourse, the author-audience/reader level and the character-character level" (1996, 169) and "[t]he overarching level of discourse is that between the playwright and the audience. Character talk is embedded in that higher discourse, allowing the audience to 'listen in' to what the characters say" (169). All the dramatic language of Shakespeare's plays comprises

speeches put into the mouths of characters, who talk to each other on stage. Even on the fairly rare occasions when a narrator addresses the audience, mediating in some way on behalf of the characters or players, (such as Gower in *Pericles*, Time in *The Winter's Tale*, the speakers of prologues in such plays as *Henry V* or *Romeo and Juliet*, or the epilogues of *As You Like It*, *A Midsummer Night's Dream* or *The Tempest*), these are still characters speaking in role. So if, as is the case here, a distinct, recurrent feature of linguistic style has been identified in the late plays, what might that mean in functional terms, in terms of the dramatic discourse? If Shakespeare's late plays use markedly more relative clauses, more complex syntax, are we seeing an expression of Short's character-character function, of the characterisation of Shakespeare's plays? Or perhaps the kind of sub-literary syntactic feature sought by authorship attributionists, as seen in 1.3 above?

McDonald addresses this point on many occasions, arguing that stylistic tendencies exist in the late plays which are not connected with character. For example:

Responding to what is probably the most obvious example in the last seven plays, the chaotic metrical scheme in Leontes' ravings about sexual infidelity, critics proudly note the indisputable parallel between the unstable verse and the unstable mind [...]

The problem with this tidy argument, of course, is that many persons in *The Winter's Tale* – and indeed all the last plays – speak similarly irregular blank verse

(McDonald 2006, 29).

In a review of McDonald's *Shakespeare's Late Style*, Brian Vickers denies the existence of any linguistic dimension in Shakespeare's plays other than that which pertains to the expression of character, saying that "[i]t follows that the language of any utterance in drama is best analysed in terms of the character who makes it, and for the purposes it is intended to fulfil" (2008, 16). This seems to contradict his earlier perspective on function words and style,

cited above¹³⁷, where he describes them as “elements that recur frequently in any piece of verse or prose and which are used by many writers with distinct preferences” (2002, 90). He argues, as McDonald anticipates, that the convolutions of Leontes’s language in *The Winter’s Tale* are entirely the expression of his emotional state (2008, e.g.22) and serve only to elucidate aspects of his character at that time. The present research is only concerned with relative clauses, but does quite clearly demonstrate that there are stylistic features in the late plays that pervade several works, and cannot therefore be conceived of as performing a characterisational function. Frequent *which* relativisation, and a double pushdown construction in this case, feature in Leontes’ convoluted ravings in the third act:

(161) Your actions are my dreams:
 You had a bastard by Polixenes,
 And I but dream’d it. As you were past all shame,—
 Those of your fact are so,— so past all truth:
 Which to deny concerns more than avails; for as
 Thy brat hath been cast out, like to itself,
 No father owning it,— *which is indeed*,
 More criminal in thee than it, – so thou

 Shalt feel our justice, *in whose easiest passage*
 Look for no less than death.

 (*The Winter’s Tale*, 3.2.80-9).

For Vickers, this marked syntax serves to express his mental state. He argues, citing Jonathan Smith¹³⁸, that

 After Mamillus’s real and Hermione’s apparent death, Leontes’ language changes in response to these disasters, taking on a grave simplicity [...] and when we see him

¹³⁷ See 1.3, p.25.

¹³⁸ ‘The Language of Leontes’. No further bibliographical details are given in the online article.

discourse from classical rhetorical features, for example. But there are alternatives to the two extremes of characterisation and sub-literary style. Very usefully for this thesis, Hope suggests that

it is easy to imagine that a writer predisposed by biography to use ‘that’ and zero forms might in fact produce a piece using entirely ‘who’ and ‘which’ forms in an attempt to write in a formal mode

(Hope 1994, 31).

And Hope offers a similar idea with regard to *Henry VIII*. Interpreting his findings concerning Shakespeare and Fletcher’s use of personal pronouns in that play, he suggests that Shakespeare may have adopted a “specific linguistic tactic” as an “attempt to write what was perceived as a new form of historical drama” (1994, 82-3). Writing “in a formal mode” as a “specific linguistic tactic” would be an expression of the author-audience/reader function in Short’s model¹³⁹. And Witmore and Hope take this further, arguing that the presence of repeated linguistic patterns, including those characterized by grammatical words, can indeed function at a rhetorical and dramatic level, although one that pervades a text at a level ‘below’ individual characterisation:

We call this assumption of pervasively coordinated rhetorical patterning the ‘principle of dramaturgical saturation’. In adopting it, we are assuming that the material constraints of the theatre and the temporal entailments of certain narrative conventions create a situation in which even the smallest particles of language are regularly pressed into the service of telling a particular type of story

(Witmore and Hope 2007, 141).

“Telling a particular kind of story” would again be an operation of the author-audience/reader function, rather than that of character-character. Perhaps in the late plays Shakespeare is engaged in telling a particularly literary, and particularly Elizabethan, kind of romance story,

¹³⁹ See p. 303.

which might then account for his widespread use of these complex relativisation strategies, also characteristic of that genre. The section which follows suggests a further step in this proposal.

6.4.3 The Elizabethan revival

If, as is argued above, the increase in *which* function is not related to characterisation, then its presence in the late plays needs to be accounted for in other ways. In her 1975 work Shakespeare's *Last Plays: A New Approach*, Frances Yates comments that "there was ... in this early Jacobean period a movement which might be called an Elizabethan revival" (19), where "the new generation is seen as reviving the values of the Elizabethan past" (26). This is manifested in "the revival of chivalry in the new generation [...] set in the form of figures of romance emerging from rocks or caves" (25). Numerous editors and commentators connect the late plays with Sidney (particularly his *Arcadia*) and Spenser, such as Kermode (1954, lx) and Butler (2005, 10-15), and both Yates (e.g. 35, 74-79) and McDonald (2006, 219-224) identify a strong connection between this wider cultural movement and the plots and themes of Shakespeare's late plays, seeing its expression in literary terms as a reawakening of interest in the Elizabethan romance writers:

The last plays are shot through with characters, episodes, language, and themes from the works of Sidney, Greene, Spenser, Lyly, Marlowe, and some of the Spanish and Italian romance writers whose tales were popular in the sixteenth century, not to mention Gower and Chaucer....Sidney seems to have exerted the most potent influence

(McDonald 2006, 222-3).

McDonald asserts here a further link, adding "language" to the list of areas influenced, although he supports this rather impressionistically, highlighting Shakespeare's "rediscovered

pleasure in artifice and verbal delights” (222), and later arguing that “Shakespeare seems to have recovered his Elizabethan enthusiasm for pattern but to have transmuted it into a poetic form so sophisticated that the verse seems both ornate and ordinary” (234). It is impossible to test these assertions, but there may be somewhat more empirical evidence available for them, from a number of perspectives.

Firstly, it seems that use of hypotaxis in general, and relativisation in particular, was linked in Early Modern linguistic usage with a set of clearly delineated values, suggesting further that Shakespeare’s late syntax is an expression of a movement against the linguistic tide of the time. Adamson describes the preoccupation in Renaissance literary language with copia, or rhetorical ornamentation (1999, 545-548), whereby the relative clause is viewed positively as a means of linguistic variation (602). As the seventeenth century progressed this was to change however, especially after the interregnum, with the onset of a desire for clarity and precision: “Neo-classical writing in general shows a marked decline in the use of parentheses and non-restrictive relative clauses, and the practice of variation, commended by renaissance critics, becomes a vice rather than a virtue” (1999, 602).

Next there is the generic connection. The omnipresence of old, usually prose, tales in the late plays is well documented: they pervade them, from their sources to their individual speeches. It is not unusual for Shakespeare to use prose narratives as sources for his plays, but the late plays are usually acknowledged as extreme in this respect. Greene’s prose romance *Pandosto* is the principal source for *The Winter’s Tale*; *The Two Noble Kinsmen* comes from Chaucer’s *The Knight’s Tale*; major elements of the plot of *Cymbeline* are acknowledged to come from Sydney’s *Arcadia* (e.g. Butler 2005, 10); the Greek prose romance *Apollonius of Tyre* is the probable earliest origin of the *Pericles* story (see e.g. Warren 2003, 13), and it also

existed as Wilkins's sixteenth century prose novel, discussed above¹⁴⁰. *The Winter's Tale* declares its genre in its title, and its main action is framed by tale telling. Hermione asks her son to "sit you by us/And tell's a tale" (2.1.31-2), and as Mamillus's "sad tale" (2.1.34-5) begins, his father enters, muttering his first jealous alarms. In the penultimate scene the three gentlemen give a diegetic account of the revelations they have seen, which are "so like an old tale, that the verity of it is in strong suspicion" (5.2.13). It is possible that there is some connection between this picture, and Shakespeare's changed linguistic style in these plays.

McDonald offers support for this view of Shakespeare reaching back to old prose tales, pointing out that "[s]ignificantly, in composing *The Winter's Tale* he used the original 1588 edition of *Pandosto*, even though the novella was reprinted in 1607" (2006, 222). The case of the Oxford editorial reconstruction of *Pericles* also seems to support the idea that frequent and complex *which* usage is associated with these prose romance works. An analysis of Wells and Taylor's (2005, 1059-1086) "reconstructed text" reveals a considerable increase in *which* frequency in the sections which the editors imported from Wilkins's novel *The Painful Adventures of Pericles, Prince of Tyre* (detailed in Appendix 9), and as discussed in 6.2.2 above, both Honigmann (1965) and Jackson (2003) also find this frequent and complex *which* relativisation in the work of Wilkins. If the Oxford text had been used in this research, instead of the Craig version which prefers the Quarto¹⁴¹, *Pericles* may well have joined the late group proper, at least in terms of its *which* frequency.

Evidence from a number of sources suggests that Shakespeare's *which* usage in the late plays is a linguistic gesture away from more typically spoken dramatic language and towards

¹⁴⁰ See 2.3 and 6.2.2.

¹⁴¹ See section 2.3.

Elizabethan prose. In his authoritative overview of Early Modern English, the historical linguist Matti Rissanen explains that “the spread of *wh-* forms is a ‘change from above’, from the formal and literary levels of the language” (1999, 295). Continuative clauses are one of the principal distinguishing features of the late plays syntax, and Nevalainen (2006, 86) and Reuter (1938, 39) both see this kind of clause as a sixteenth century phenomenon, typical of the Elizabethans. As seen earlier here in 3.2,¹⁴² Dekeyser and Ingels compare relativisation strategies across three corpora of generically distinct texts (informative prose, tragic drama and comic drama) from the second half of the sixteenth century, and find that their parameters “unmistakably rank [Prose] and Drama respectively as more and less complex” (1988, 34). They reveal that *wh-* forms were connected strongly with written (rather than written-to-be-spoken) genres, and with the more formal registers such as informative scientific or legal prose (e.g. Dekeyser and Ingels 1988, 35-6). Jackson also corroborates this perspective, noting that “while the use of a Latinate ‘connective’ *which* is common enough among seventeenth century English prose writers, it is rare in drama” (2003, 123). Although there are arguments about the Latin provenance of continuative clauses (e.g. Reuter 1938), Jackson’s observation is further confirmation of the stylistic shift in Shakespeare’s late work towards syntax more usually associated with prose.

Although Dekeyser and Ingels demonstrate the connection between *wh-* forms and Elizabethan informative prose, their sample does not unfortunately include selections from fictional prose of the sixteenth century, which is more likely to be the specific genre in question in Shakespeare’s case. A selection of passages from Sidney’s *Arcadia* therefore follows here (*italics added*), to demonstrate the prose style:

¹⁴² See pages 60-61 above.

Which, as in part it was more obscure than he could understand, so did the whole bear such manifest threatenings, that his amazement was greater than his fore curiosity—both passions proceeding out of one weakness: in vain to desire to know that of *which* in vain thou shalt be sorry after thou hast known it. But thus the duke answered though not satisfied, he returned into his country with a countenance well witnessing the dismayedness of his heart; *which* notwithstanding upon good considerations he thought not good to disclose

(The First Book or Act, 175).

While Pamela, sitting her down under one of them, and making a posy of the fair undergrowing flowers, filled Musidorus's ears with the heavenly sound of her music, *which* before he had never heard, so that it seemed unto him a new assault given to the castle of his heart, already conquered; *which* to signify, and withal reply to her sweet notes, he sang in a kind of still but ravishing tune a few verses.

(The Third Book or Act, 175)

'Noble prince,' said she, 'your words are too well couched to come out of a restless mind, and thanked be the gods your face threatens no danger of death. These be but those swelling speeches *which* give the uttermost name to every trifle, *which* all were worth nothing if they were not enameled with the goodly outside of love. Truly, love were very unlovely if it were half so deadly as you lovers (still living) term it. I think well it may have a certain childish vehemency *which* for the time to one desire will engage all the soul, so long as it lasteth.

(The Third Book or Act, 192-193)

At length, remembering the oracle, *which* now indeed was accomplished (not as before he had imagined), considering all had fallen out by the highest providence, and withal weighing in all these matters his own fault had been the greatest, the first thing he did was with all honourable pomp to send for Gynecia (who, poor lady, thought she was leading forth to her living burial), and (when she came) to recount before all the people the excellent virtue was in her, *which* she had not only maintained all her life most unspotted but now was content so miserably to die to follow her husband.

(The Last Book or Act, 361)

There are considerable limitations to exemplification as a means of supporting a case, as argued above¹⁴³, and the analysis conducted in this thesis as a whole could usefully be applied to a corpus of Elizabethan romance prose in order to compare it with Shakespeare's late

¹⁴³ See section 1.2 for example.

language more fully. But it can be seen here quite plainly, I think, that Sidney's syntax is complex, strongly hypotactic, rich in relativisation, and in *which* occurrences (italicized) in particular. There would therefore seem to be evidence for Shakespeare's increased *which* usage in the late plays (and, I would suggest, his more complex syntax there as a whole) being a manifestation of his renewed interest in the Elizabethan prose writers, both in their style and content. Previous commentators have identified the influence of writers such as Sidney and Spenser in the narrative topics of Shakespeare's late plays, but it seems highly likely from the evidence offered in this thesis that he is also alluding to them through his syntax.

APPENDIX 1

NON-SHAKESPEAREAN MATERIAL REMOVED FROM THE CORPORA

Titus Andronicus

Act 1

(George Peele) (Vickers 2002, 148-243)

Timon of Athens

Act 1 scene 2

Act 3

(Thomas Middleton) (Vickers 2002, 244-290)

Pericles

Acts 1 and 2

(George Wilkins) (Vickers 2002, 291-332)

Henry VIII

Act 1 scene 3

Act 1 scene 4

Act 2 scene 1

Act 2 scene 2

Act 3 scene 1

Act 4 scene 1

Act 4 scene 2

Act 5 scene 2

Act 5 scene 3

Act 5 scene 4

Act 5 scene 5

(John Fletcher) (Vickers 2002, 333-402)

The Two Noble Kinsmen

Act 1 scene 5

Act 2 scene 2

Act 2 scene 3

Act 2 scene 4

Act 2 scene 5

Act 2 scene 6

Act 3 scene 3

Act 3 scene 4

Act 3 scene 5

Act 3 scene 6

Act 4 scene 1

Act 4 scene 2

Act 5 scene 2

(John Fletcher) (Vickers 2002, 402-432)

APPENDIX 3

KEYWORD LIST FOR THE LATE PLAYS

1-18	19-36	37-54	55-72	73-90	91-107
Camillo	Paulina	Tarsus	Gonzalo	Iachimo	Princess
Palamon	Milford	Caliban	Tunis	Boult	She
Posthumus	Trinculo	Queen's	Briton	Beest	Its
Britain	Cromwell	Business	Tyre	Fardel	<i>Unto</i>
Pisanio	Cloten	Her	Britons	Sycorax	<i>Uncle</i>
Which	Milan	Highness	King's	Leonine	<i>Wit</i>
Bohemia	Naples	Cymbeline	Daughter	Heavens	<i>Caesar</i>
Pericles	't	Has	Rare	Being	<i>Go</i>
Marina	Sea	Island	Perdita	More	<i>Doth</i>
Imogen	Stephano	Haven	Polydore	Garments	<i>Why</i>
Monster	Helicanus	Thebes	Mitylene	Fish	<i>In</i>
Queen	Oracle	Beyond	Court	Honour	<i>And</i>
Arcite	Sicilia	Cardinal	On't	Goodness	<i>Will</i>
Gods	'em	Cleon	Tribute	Shearing	<i>France</i>
O'th	Polixenes	Lychorida	Language	Antigonus	<i>John</i>
Ariel	In't	Fidele	garment	To'th	<i>God</i>
Leonatus	Prospero	Leontes	flowers	Most	<i>Love</i>
Hermione	I'th	Thaisa	Cadwal	Temple	

Numbers 94-107 inclusive (italicized) are negative keywords, found with salient infrequency in the late plays. This list is also provided as Appendix 4 on the enclosed cd, with additional statistical data relating to raw frequency, relative frequency and keyness. (N.B. In the cd table of keywords 'R.C' denotes 'Reference Corpus', which is the corpus of thirty two earlier plays).

APPENDIX 5: PARADIGMATIC STATISTICS

Play	Who (relatives)	Whom (relatives)	Whose (relatives)	Which (relatives)	That (relatives)	That's (relatives)
<i>Two Gentlemen of Verona</i>	20 (4)	10 (7)	11 (11)	41 (39)	285(101)	22 (4)
<i>The Taming of the Shrew</i>	19 (4)	11 (8)	5 (4)	26 (23)	233 (79)	8 (0)
<i>Henry VI part 2</i>	35 (17)	15 (12)	24 (24)	27 (25)	286 (114)	8 (0)
<i>Henry VI Part 3</i>	26 (11)	21 (16)	24 (24)	39 (39)	324 (122)	9 (0)
<i>Henry VI Part 1</i>	20 (5)	21 (15)	15 (14)	42 (34)	251 (83)	3 (0)
<i>Titus Andronicus</i>	24 (8)	8 (7)	10 (9)	24 (21)	212 (72)	7 (2)
<i>Richard III</i>	55 (12)	19 (17)	15 (14)	77 (75)	434 (173)	3 (2)
<i>The Comedy of Errors</i>	19 (10)	10 (7)	7 (3)	22 (15)	217(82)	9 (1)
<i>Love's Labour's Lost</i>	24 (5)	14 (7)	9 (5)	66 (58)	284 (123)	11 (1)
<i>Richard II</i>	33 (13)	9 (8)	13 (12)	67 (65)	284 (116)	3 (0)
<i>Romeo and Juliet</i>	33 (15)	6 (6)	15 (13)	62 (58)	345 (136)	8 (2)
<i>A Midsummer Night's Dream</i>	11 (3)	5 (5)	8 (7)	38 (33)	186 (67)	1 (0)
<i>King John</i>	33 (10)	13 (10)	30 (28)	60 (58)	307 (108)	2 (1)
<i>The Merchant of Venice</i>	40 (15)	13 (10)	10 (9)	59 (53)	259 (101)	9 (0)
<i>Henry IV Part 1</i>	28 (17)	6 (4)	18 (17)	48 (46)	254(80)	12 (1)
<i>The Merry Wives of Windsor</i>	12 (5)	2 (2)	3 (3)	26 (20)	206 (50)	7 (0)
<i>Henry IV Part 2</i>	26 (9)	5 (4)	19 (17)	82 (73)	292(110)	7 (2)
<i>Much Ado About Nothing</i>	31 (15)	3 (3)	7 (7)	56 (45)	290 (87)	12 (2)

Play	Who (relatives)	Whom (relatives)	Whose (relatives)	Which (relatives)	That (relatives)	That's (relatives)
<i>Henry V</i>	34 (19)	12 (12)	22 (22)	69 (69)	310 (125)	14 (0)
<i>Julius Caesar</i>	28 (13)	1 (0)	5 (5)	47 (45)	284 (96)	4 (0)
<i>As You Like It</i>	41 (12)	6 (4)	6 (6)	58 (55)	344 (142)	5 (0)
<i>Hamlet</i>	41 (11)	6 (4)	32 (28)	63 (61)	395 (129)	16 (5)
<i>Twelfth Night</i>	22 (3)	8 (7)	7 (7)	24 (21)	274 (100)	12 (1)
<i>Troilus and Cressida</i>	37 (11)	14 (11)	25 (18)	43 (38)	305 (125)	18 (1)
<i>Measure for Measure</i>	28 (16)	13 (9)	18 (18)	58 (52)	305 (114)	17 (4)
<i>Othello</i>	32 (8)	10 (8)	14 (12)	46 (46)	376 (138)	23 (6)
<i>King Lear</i>	66 (35)	13 (11)	24 (23)	82 (75)	342 (135)	15 (6)
<i>Timon of Athens</i>	18 (9)	17 (17)	19 (19)	32 (31)	148 (60)	
<i>Macbeth</i>	41 (15)	12 (11)	14 (14)	80 (73)	229 (102)	9 (9)
<i>Antony and Cleopatra</i>	21 (13)	15 (13)	19 (18)	108 (106)	291 (103)	14 (5)
<i>All's Well That Ends Well</i>	21 (10)	17 (17)	26 (26)	87 (83)	327 (124)	17 (7)
<i>Pericles</i>	31 (19)	11 (11)	4 (4)	29 (29)	109 (52)	2 (0)
<i>Coriolanus</i>	42 (32)	16 (16)	20 (20)	102 (96)	319 (130)	19 (5)
<i>The Winter's Tale</i>	39 (27)	20 (18)	26 (26)	121 (118)	319 (128)	19 (5)
<i>Cymbeline</i>	54 (35)	25 (25)	31 (31)	124 (12)	365 (147)	16 (9)
<i>The Tempest</i>	30 (26)	20 (19)	16 (16)	72 (70)	190 (85)	15 (4)
<i>Two Noble Kinsmen</i>	21 (17)	7 (7)	15 (15)	48 (46)	150 (67)	5 (2)
<i>Henry VIII</i>	24 (15)	6 (6)	7 (7)	50 (48)	171 (55)	4 (1)

APPENDIX 9

TABLES OF RESULTS FOR THE ANALYSIS OF THE ‘NEW LATE GROUP’

Constitution of corpora:

- The new configuration of late plays = ‘High *which* group’:

All’s Well That Ends Well, Antony and Cleopatra, Macbeth, Coriolanus, Cymbeline, The Winter’s Tale, The Tempest, Henry VIII, The Two Noble Kinsmen.

- The new ‘earlier canon’ = ‘Low *which* group’:

All the other plays in the canon, (including *Pericles.*).

Table 1: Total tokens

High <i>which</i> group play	Total tokens
<i>All’s Well That Ends Well</i>	22,648
<i>Antony and Cleopatra</i>	23,909
<i>Macbeth</i>	16,568
<i>Coriolanus</i>	26,783
<i>The Two Noble Kinsmen</i>	10,240
<i>The Winter’s Tale</i>	24,701
<i>The Tempest</i>	16,164
<i>Henry VIII</i>	11,953
<i>Cymbeline</i>	26,975
Total for high <i>which</i> group plays	179,941
Total for low <i>which</i> group plays	621,623

Table 2: Raw and relative frequencies of types

	Raw (and relative) frequency in low <i>which</i> group	Rate of occurrence per 1,000 words	Raw (and relative) frequency in high <i>which</i> group	Rate of occurrence per 1,000 words	Statistical significance of change (Chi square score¹⁴⁴)
<i>which</i>	1415 (0.23%)	2.3	792 (0.44%)	4.4	204.59
<i>who</i>	867 (0.14%)	1.4	293 (0.16%)	1.6	5.11
<i>whom</i>	302 (0.05%)	0.5	138 (0.08%)	0.8	18.54
<i>whose</i>	419 (0.07%)	0.7	174 (0.10%)	1.0	15.18
<i>that</i> (inc. <i>that</i> 's)	8,416 (1.35%)	13.5	2,479 (1.38%)	13.8	0.58

Table 3: Proportion of each type that function as relatives

	Number of type in low <i>which</i> group that are relatives (proportion of type)	Number of type in high <i>which</i> group that are relatives (proportion of type)
<i>Which</i>	1,302 (92.14%)	760 (95.72%)
<i>Who</i>	334 (38.52%)	190 (64.85%)
<i>Whom</i>	242 (80.13%)	132 (95.62%)
<i>Whose</i>	383 (91.41%)	173 (99.42%)
<i>That</i>	3,062 (36.38%)	988 (39.85%)

¹⁴⁴ See p. 50.

Table 4: Raw frequency and proportion of relatives

	Relative <i>who/whom/whose</i> (% of total relatives)	Relative <i>which</i> (% of total relatives)	Relative <i>that</i> (inc. <i>that's</i>) (% of total relatives)	Totals (% of total relatives) (% of total tokens)
Low <i>which</i> group	959 (18.02%)	1302 (24.46%)	3,062 (57.52%)	5,323 (100%) (0.86%)
High <i>which</i> group	495 (22.07%) (+4.05%)	760 (33.88%) (+9.42%)	988 (44.05%) (-13.47%)	2,243 (100%) (1.25%)

Table 5: Main function of *which*

<i>Which</i> function	Low <i>which</i> group	High <i>which</i> group
Interrogative	112 (7.91%)	31 (3.91%)
Relative	1,302 (92.01%)	760 (95.96%)
Pronoun	1,244 (87.91%)	737 (93.05%)
Determiner	58 (4.10%)	23 (2.90%)
Total <i>which</i> occurrences	1,415 (incl. 1 anomaly (0.07%))	792 (incl. 1 anomaly (0.13%))

Table 6: Types of relative clause according to antecedent

	Low <i>which</i> group (% of relative <i>which</i> clauses)	High <i>which</i> group (% of relative <i>which</i> clauses)
Adnominal clauses (single NP antecedent)	942 (72.35%)	540 (71.05%)
Adnominal clauses (multiple NP antecedent)	47 (3.61%)	27 (3.55%)
Total adnominal clauses	989 (75.96%)	567 (74.60%)
Nominal clauses (beginning <i>that which</i>)	110 (8.45%)	79 (10.39%)
Total single word¹⁴⁵ antecedents	1099 (84.41%)	646 (85%)
Sentential antecedents	198 (15.21%)	113 (14.87%)
Total relative <i>which</i>	1,302	760

Table 7: The animacy parameter

	Low <i>which</i> group (% of single word antecedents)	High <i>which</i> group (% of single word antecedents)
Animate antecedent	83(7.55%)	49 (7.58%)
Inanimate antecedent	1,016 (92.45%)	597 (92.41%)
Total clauses with single word antecedent	1,099	646

¹⁴⁵ All noun phrase antecedents are reduced to single (head) words in this study for the purpose of counting, and the term ‘single word antecedent’ refers to those, as distinct from sentential antecedents.

Table 8: Distance between antecedent and relative

	Low <i>which</i> group (% of single word antecedents)	High <i>which</i> group (% of single word antecedents)
2 words distant	108 (9.83%)	59 (9.13%)
3 words	87 (7.92%)	49 (7.58%)
4 words	60 (5.46%)	24 (3.71%)
5 words	28 (2.55%)	23 (3.56%)
6 words	18 (1.64%)	5 (0.77%)
7 words	18 (1.64%)	8 (1.24%)
8 words	8 (0.73%)	9 (1.39%)
9 words	3 (0.27%)	6 (0.93%)
10 or more words	16 (1.45%)	7 (1.08%)
Total interventions with 5+ words	91 (8.28%)	58 (8.98%)
Total clauses with single word antecedents	1,099	646

Table 9: Word groups intervening to the left

Feature intervening between antecedent and relative	Low <i>which</i> group (% of all single word antecedents)	High <i>which</i> group (% of all single word antecedents)
1 phrase	192 (17.47%)	93 (14.40%)
2 or more phrases	15 (1.36%)	11 (1.70%)
Total with phrase	207 (18.83%)	104 (16.10%)
1 non-finite clause	18 (1.64%)	13 (2.01%)
1 non-finite clause and phrase(s)	10 (0.91%)	7 (1.08%)
Total with a non-finite clause	28 (2.55%)	20 (3.10%)
1 partial finite clause	57 (5.19%)	28 (4.33%)
1 partial finite clause and phrase(s)	27 (2.00%)	13 (2.01%)
Total with a partial finite clause	84 (7.64%)	41 (6.35%)
1 finite clause	28 (2.55%)	20 (3.09%)
1 finite clause and phrase(s)	11 (0.91%)	11 (1.70%)
Total with a finite clause	39 (3.55%)	31 (4.80%)
Multiple clauses	18 (1.64%)	9 (1.39%)
Total with clause	169 (15.38%)	101 (15.63%)
Total with intervention	376 (34.21%)	205 (31.73%)
Total clauses with single word antecedent	1,099	646

Table 10: The information parameter

Clause type	Low <i>which</i> group (% of relative <i>which</i> clauses)	High <i>which</i> group (% of relative <i>which</i> clauses)
Restrictive	330 (25.34%)	224 (29.47%)
Non-restrictive	972 (74.65%)	535 (70.39%)
Continuative	208 (15.97%)	118 (15.53%)
Total relative <i>which</i> clauses	1302	760 (incl. 1 anomaly)

Table 11: Analysis of continuative clauses

	Low <i>which</i> group (% of continuatives)	High <i>which</i> group (% of continuatives)
Antecedent 2 words distant	7 (3.3%)	5 (4.24%)
3 words	6 (2.8%)	8 (6.8%)
4 words	11(5.2%)	1 (0.8%)
5 words	3 (1.4%)	3 (2.5%)
6 words	5 (2.3%) (9.26%)	2 (1.7%)
7 words	9 (4.2%)	5 (4.2%)
8 words	6 (2.8%)	2 (1.7%)
9 words	0	4 (3.39%)
10 or more words	5 (2.36%)	5 (4.24%)
Total with 2 or more words distant antecedent	54 (25.47%)	33 (27.97%)
Total interventions with 5+ words	28 (13.21%)	21 (17.80%)
With determiner <i>which</i>	53 (25.00%) (n/a)	20 (16.95%)
With sentential antecedent	93 (43.9%) (n/a)	52 (44.0%)
Total continuative clauses	212	118
Total relative <i>which</i> clauses	1,302	760

Table 12: The role of *which* in the relative clause: principal functions

Role of <i>which</i>	Low <i>which</i> group (% of relative <i>which</i>)	High <i>which</i> group (% of relative <i>which</i>)
Subject (incl. single and double pushdowns)	600 (46.08%)	403 (53.03%)
Object (incl. single and double pushdowns)	476 (36.56%)	234 (30.79%)
Prepositional complement (incl. single and double pushdowns)	175 (13.44%)	91 (11.97%)
Total relative <i>which</i>	1,302	760

Table 13: The role of *which* in the relative clause: additional constructions

Role of <i>which</i>	Low <i>which</i> group (% of relative <i>which</i>)	High <i>which</i> group (% of relative <i>which</i>)
Relativiser-only	16 (1.23%) (incl. 2 prepositional complements ¹⁴⁶)	4 (0.53%)
Single pushdown	53 (4.07%) (incl. 11 not assigned a role elsewhere) ¹⁴⁷	12 (1.58%) (incl. 1 not assigned a role elsewhere)
Double pushdown	13 (1.00%)	17 (2.24%)
Recursion	27 (2.07%)	25 (3.29%)
Cataphoric <i>which</i>	13 (1.00%)	4 (0.53%)
Total relative <i>which</i>	1,302	760 (inc. 1 anomaly ¹⁴⁸)

¹⁴⁶ See 4.4 for explanation.

¹⁴⁷ See 4.4.

¹⁴⁸ Example (109) again.

Table 14: Number of words intervening to the right

	Low <i>which</i> group (% of relative <i>which</i>)	High <i>which</i> group (% of relative <i>which</i>)
1 word	111 (8.52%)	64 (8.42%)
2 words	31 (2.38%)	30 (3.95%)
3 words	66 (5.07%)	28 (3.68%)
4 words	40 (3.07%)	15 (1.97%)
5 words	14 (1.07%)	5 (0.66%)
6 words	13 (1.00%)	7 (0.92%)
7 words	19 (1.46%)	2 (0.26%)
8 words	8 (0.61%)	1 (0.13%)
9 words	6 (0.46%)	1 (0.13%)
10 or more words	6 (0.46%)	1 (0.13%)
Total with intervention	314 (24.12%)	154 (20.26%)
Total interventions with 5+ words	66 (5.07%)	17 (2.24%)
Total relative <i>which</i>	1,302	760

Table 15: Word groups intervening to the right of the relative

Feature intervening to the right	Low <i>which</i> group (% of relative <i>which</i>)	High <i>which</i> group (% of relative <i>which</i>)
1 phrase	197 (15.13%)	96 (12.63%)
2 or more phrases	36 (2.76%)	10 (1.31%)
Total with phrase(s)	233 (17.89%)	106 (13.95%)
1 non-finite clause	27 (2.07%)	27 (3.55%)
1 non-finite clause and phrase(s)	13 (1.00%)	8 (1.05%)
Total with a non-finite clause	40 (3.07%)	35 (4.60%)
1 finite clause	17 (1.30%)	9 (1.18%)
1 finite clause and phrase(s)	10 (0.77%)	0
Total with a finite clause	27 (2.07%)	9 (1.18%)
Multiple clauses (of either or both kinds)	13(1.00%)	4 (0.53%)
Total with clause(s)	80 (6.14%)	48 (6.31%)
Total with intervention	313 (24.04%)	154 (20.26%)

APPENDIX 10

TABLE OF RESULTS FOR THE ANALYSIS OF THE INDIVIDUAL PLAYS

Table 1: Raw and relative frequencies for word types

	<i>All's Well That Ends Well</i>	<i>Macbeth</i>	<i>Antony and Cleopatra</i>	<i>Coriolanus</i>
Total tokens	22,648	16,568	23,909	26,783
<i>Which</i> : total / relatives (% of all relatives)	87/83 (95.40%)	80/73 (91.25%)	108/106 (98.15%)	102/96 (94.12%)
<i>Who</i> : total / relatives (% of all relatives) ¹⁴⁹	21/10 (47.62%)	41 /15 (36.58%)	21 /13 (61.90%)	42/34 (80.95%)
<i>Whom</i> : total / relatives (% of all relatives)	17/17 (100%)	12/11 (91.67%)	15/13 (86.67%)	16/16 (100%)
<i>Whose</i> : total / relatives (% of all relatives)	26/26 (100%)	14/14 (100%)	19/18 (94.74%)	20/20 (100%)
<i>That</i> : total / relatives (% of all relatives) ¹⁵⁰	335 / 150 (44.78%)	238/116 (48.74%)	305/117 (38.36%)	338/139 (41.12%)
Total relatives (% of total tokens)	286 (1.26%)	229 (1.38%)	267 (1.12%)	305 (1.14%)

¹⁴⁹ The total for *who*, here and elsewhere, does not include *who* 's, since, as pointed out previously, there are no cases of that form with relative function in the canon.

¹⁵⁰ This total includes *that* 's, as elsewhere.

	<i>Pericles</i>	<i>Cymbeline</i>	<i>The Winter's Tale</i>	<i>The Tempest</i>
Total tokens	9,943	26,975	24,701	16,164
<i>Which</i> : total / relatives (% of all relatives)	29/29 (100%)	124/120 (96.77%)	121/118 (97.52%)	72/70 (97.22%)
<i>Who</i> : total / relatives (% of all relatives)	31/20 (64.52%)	54/36 (66.67%)	39/28 (71.79%)	30/26 (86.67%)
<i>Whom</i> : total / relatives (% of all relatives)	11/9 (81.82%)	25/23 (92.00%)	20/17 (85.00%)	20/18 (90.00%)
<i>Whose</i> : total / relatives (% of all relatives)	4/4 (100%)	31/31 (100%)	26/26 (100%)	16/16 (100%)
<i>That</i> : total / relatives (% of all relatives)	111/47 (42.34%)	381/166 (43.57%)	338/140 (41.41%)	205/93 (45.36%)
Total relatives (% of total tokens)	109 (1.10%)	376 (1.39%)	329 (1.33%)	223 (1.38%)

	<i>Henry VIII</i>	<i>The Two Noble Kinsmen</i>
Total tokens	11,953	10,240
<i>Which</i> : total / relatives (% of all relatives)	50/48 (96.00%)	48/46 (95.83%)
<i>Who</i> : total / relatives (% of all relatives)	24/17 (70.83%)	21/17 (80.95%)
<i>Whom</i> : total / relatives (% of all relatives)	6/6 (100%)	7/7 (100%)
<i>Whose</i> : total / relatives (% of all relatives)	7/7 (100%)	15/15 (100%)
<i>That</i> : total / relatives (% of all relatives)	175/59 (33.71%)	155/71 (45.81%)
Total relatives (% of total tokens)	137 (1.15%)	156 (1.52%)

Table 2: Main functions of *which*

	<i>All's Well That Ends Well</i>	<i>Macbeth</i>	<i>Antony and Cleopatra</i>	<i>Coriolanus</i>
Total <i>which</i> as relative (pronoun or determiner) (% of total <i>which</i>)	83 (95.40%)	73 (91.25%)	106 (98.15%)	96 (94.12%)
<i>Which</i> as relative pronoun (% total <i>which</i>)	80 (91.95%)	70 (87.50%)	103 (95.37%)	94 (92.16%)
<i>Which</i> as relative determiner (% total <i>which</i>)	3 (3.45%)	4 (5.00%)	3 (2.78%)	2 (1.96%)
<i>Which</i> as interrogative (% total <i>which</i>)	4 (4.6%)	7 (8.75%)	2 (1.8%)	6 (5.88%)
Anomalies	0	0	0	0

	<i>Pericles</i>	<i>Cymbeline</i>	<i>The Winter's Tale</i>	<i>The Tempest</i>
Total <i>which</i> as relative (pronoun or determiner) (% of total <i>which</i>)	29 (100.00%)	120 (96.77%)	118 (97.52%)	70 (97.22%)
<i>Which</i> as relative pronoun (% total <i>which</i>)	28 (96.55%)	119 (95.97%)	115 (95.04%)	66 (91.67%)
<i>Which</i> as relative determiner (% total <i>which</i>)	1 (3.45%)	1 (0.81%)	3 (2.48%)	4 (5.55%)
<i>Which</i> as interrogative (% total <i>which</i>)	0	4 (3.22%)	2 (1.65%)	2 (2.78%)
Anomalies	0	0	1 ¹⁵¹	0

¹⁵¹ See (102) in 4.1.

	<i>Henry VIII</i>	<i>The Two Noble Kinsmen</i>
Total <i>which</i> as relative (pronoun or determiner) (% of total <i>which</i>)	48 (96.00%)	46 (95.83%)
<i>Which</i> as relative pronoun (% total <i>which</i>)	46 (92.00%)	45 (93.75%)
<i>Which</i> as relative determiner (% total <i>which</i>)	2 (4.00%)	1 (2.08%)
<i>Which</i> as interrogative (% total <i>which</i>)	2 (4.00%)	2 (4.17%)
Anomalies	0	0

Table 3: Types of relative clause according to antecedent

Types of relative clause	<i>All's Well That Ends Well</i> (% relative <i>which</i>)	<i>Macbeth</i> (% relative <i>which</i>)	<i>Antony and Cleopatra</i> (% relative <i>which</i>)	<i>Coriolanus</i> (% relative <i>which</i>)
Adnominal	56 (67.5%)	48 (65.75%)	77 (72.64%)	74 (77.08%)
Multiple NP adnominal	4 (4.82%)	2 (2.74%)	2 (1.89%)	3 (3.12%)
Total adnominal	60 (72.3%)	50 (68.49%)	79 (74.52%)	77 (80.21%)
Nominal clauses - <i>that which</i>	7 (8.43%)	14 (19.18%)	14 (13.21%)	6 (6.25%)
Total single word antecedents	67 (80.72%)	64 (87.67%)	93 (87.73%)	83 (86.46%)
Sentential	15 (18.07%)	9 (12.33%)	13 (12.26%)	13 (13.54%)
Anomalies	1	0	0	0

Types of relative clause	<i>Pericles</i> (% relative <i>which</i>)	<i>Cymbeline</i> (% relative <i>which</i>)	<i>The Winter's Tale</i> (% relative <i>which</i>)	<i>The Tempest</i> (% relative <i>which</i>)
Adnominal	21 (72.41%)	89 (74.17%)	80 (67.80%)	51 (72.86%)
Multiple NP adnominal	1 (3.45%)	2 (1.67%)	4 (3.39%)	3 (4.28%)
Total adnominal	22 (75.86%)	91 (75.83%)	84 (71.19%)	54 (77.14%)
Nominal clauses - <i>that which</i>	2 (6.90%)	12 (10.00%)	13 (11.02%)	5 (7.14%)
Total single word antecedents	24 (82.76%)	103 (85.83%)	97 (82.20%)	59 (84.28%)
Sentential	5 (17.24%)	17 (14.17%)	21 (17.80%)	11 (15.71%)
Anomalies	0	0	0	0

Types of relative clause	<i>Henry VIII</i> (% relative <i>which</i>)	<i>The Two Noble Kinsmen</i> (% relative <i>which</i>)
Adnominal	33 (68.75%)	32 (69.56%)
Multiple NP adnominal	4 (8.33%)	3 (6.52%)
Total adnominal	36 (75.00%)	35 (76.09%)
Nominal - <i>that which</i>	4 (8.33%)	4 (8.69%)
Total single word antecedents	41 (85.42%)	39 (84.78%)
Sentential	7 (14.58%)	7 (15.22%)
Anomalies	0	0

Table 4: The animacy parameter

Antecedent type (% of single word antecedents)	<i>All's Well That Ends Well</i>	<i>Macbeth</i>	<i>Antony and Cleopatra</i>	<i>Coriolanus</i>
Animate antecedent	3 (4.48%)	4 (6.25%)	4 (4.30%)	3 (3.61%)
Inanimate antecedent	64 (95.52%)	60 (93.75%)	89 (95.70%)	80 (96.38%)

Antecedent type (% of single word antecedents)	<i>Pericles</i>	<i>Cymbeline</i>	<i>The Winter's Tale</i>	<i>The Tempest</i>
Animate antecedent	0	9 (8.74%)	10 (10.31%)	11 (18.64%)
Inanimate antecedent	24 (100%)	94 (91.26%)	87 (89.69%)	48 (81.35%)

Antecedent type (% of single word antecedents)	<i>Henry VIII</i>	<i>The Two Noble Kinsmen</i>
Animate antecedent	2 (4.88%)	3 (7.69%)
Inanimate antecedent	39 (95.12%)	36 (92.31%)

Table 5: Distance between antecedent and relative in the individual plays

	<i>All's Well That Ends Well</i> (% of s. w.a.'s ¹⁵²)	<i>Macbeth</i> (% of s. w.a.'s)	<i>Antony and Cleopatra</i> (% of s. w.a.'s)	<i>Coriolanus</i> (% of s. w.a.'s)
2 words distant	9 (13.43%)	1 (1.56%)	14 (15.05%)	11 (13.25%)
3 words	6 (8.95%)	3 (4.69%)	5 (5.38%)	5 (6.02%)
4 words	3 (4.48%)	4 (6.25%)	1 (1.07%)	5 (6.02%)
5 words	5 (7.46%)	3 (4.69%)	3 (3.22%)	0
6 words	0	0	0	0
7 words	0	0	0	2 (2.41%)
8 words	1 (1.49%)	2 (3.12%)	1 (1.07%)	0
9 words	2 (2.98%)	0	1 (1.07%)	0
10 or more words	0	1 (1.56%)	1(1.07%)	0
Total with 2+ words intervening	26 (38.80%)	14 (21.87%)	26 (27.96%)	23 (27.71%)
Total with 5+ words intervening	8 (11.94%)	6 (9.37%)	6 (6.45%)	2 (2.41%)

¹⁵² Single word antecedents.

	<i>Pericles</i> (% of s. w.a.'s)	<i>Cymbeline</i> (% of s. w.a.'s)	<i>The Winter's Tale</i> (% of s. w.a.'s)	<i>The Tempest</i> (% of s. w.a.'s)
2 words distant	0	11 (10.68%)	4 (4.12%)	3 (5.08%)
3 words	0	4 (3.88%)	13 (13.40%)	5 (8.47%)
4 words	0	5 (4.85%)	1 (1.03%)	1 (1.69%)
5 words	0	1 (0.97%)	3 (3.09%)	4 (6.78%)
6 words	1 (4.17%)	1 (0.97%)	0	3 (5.08%)
7 words	1 (4.17%)	2 (1.94%)	2 (2.06%)	2 (3.39%)
8 words	0	1 (0.97%)	2 (2.06%)	0
9 words	0	2 (1.94%)	1 (1.03%)	0
10 or more words	0	3 (2.91%)	1 (1.03%)	0
Total with 2+ words intervening	2 (8.33%)	30 (29.13%)	27 (27.83%)	18 (30.51%)
Total with 5+ words intervening	2 (8.33%)	10 (9.71%)	9 (9.28%)	9 (15.25%)

	<i>Henry VIII</i> (% of s. w.a.'s)	<i>The Two Noble Kinsmen</i> (% of s. w.a.'s)
2 words distant	5 (12.19%)	1 (2.56%)
3 words	5 (12.19%)	3 (7.69%)
4 words	3 (7.32%)	1 (2.56%)
5 words	2 (4.88%)	2 (5.13%)
6 words	0	1 (2.56%)
7 words	0	0
8 words	1 (2.44%)	1 (2.56%)
9 words	0	0
10 or more words	1 (2.44%)	0
Total with 2+ words intervening	17 (41.46%)	9 (23.08%)
Total with 5+ words intervening	4 (9.76%)	4 (10.27%)

Table 6: Word groups intervening to the left

Feature intervening between antecedent and relative	<i>All's Well That Ends Well</i> (% of s.w.a's)	<i>Macbeth</i> (% of s.w.a's)	<i>Antony and Cleopatra</i> (% of s.w.a's)	<i>Coriolanus</i> (% of s.w.a's)
1 phrase	13 (19.40%)	8 (12.50%)	13 (13.98%)	12 (14.46%)
2 or more phrases	3 (4.48%)	1 (1.56%)	1 (1.07%)	2 (2.41%)
Total with phrase	16 (23.88%)	9 (14.06%)	14 (15.05%)	14 (16.87%)
1 non-finite clause	0	1 (1.56%)	3 (3.22%)	1 (1.20%)
1 non-finite clause and phrase(s)	1 (1.49%) (3.70%)	0	0	0
Total with a non-finite clause	1 (1.49%)	1 (1.56%)	3 (3.22%)	1 (1.20%)
1 partial finite clause	6 (8.95%)	2 (3.12%)	3 (3.22%)	4 (4.82%)
1 partial finite clause and phrase(s)	1(1.49%)	3 (4.68%)	1 (1.07%)	1 (1.20%)
Total with a partial finite clause	7 (10.45%)	5 (7.81%)	4 (4.30%)	5 (6.02%)
1 finite clause	0	0	5 (5.38%) (18.52%)	2 (2.41%) (8.69%)
1 finite clause and phrase(s)	2 (2.98%)	1 (1.56%)	0	0
Total with a finite clause	2 (2.98%)	1 (1.56%)	5 (5.38%)	2 (2.41%)
Multiple verb phrases and/or clauses	1 (1.49%)	1 (1.56%)	1 (1.07%)	1 (1.20%)
Total with clause	11 (16.42%)	8 (12.50%)	13 (13.98%)	9 (10.84%)
Total with intervention	27 (40.30%)	17 (26.56%)	27 (29.03%)	23 (27.71%)

Feature intervening between antecedent and relative	<i>Pericles</i> (% of s.w.a's)	<i>Cymbeline</i> (% of s.w.a's)	<i>The Winter's Tale</i> (% of s.w.a's)	<i>The Tempest</i> (% of s.w.a's)
1 phrase	0	14 (13.59%)	14 (14.43%)	8 (13.56%)
2 or more phrases	1 (4.17%)	2 (1.94%)	0	0
Total with phrase	1 (4.17%)	16 (15.53%)	14 (14.43%)	8 (13.56%)
1 non-finite clause	0	1 (0.97%)	3 (3.09%)	0
1 non-finite clause and phrase(s)	0	3 (2.91%)	1 (1.03%)	2 (3.39%)
Total with a non-finite clause	0	4 (3.88%) (12.90%)	4 (4.12%) (13.33%)	2 (3.39%) (9.09%)
1 partial finite clause	0	2 (1.94%)	4 (4.12%)	3 (5.08%)
1 partial finite clause and phrase(s)	1 (4.17%)	1 (0.97%)	2 (2.06%)	3 (5.08%)
Total with a partial finite clause	1 (4.17%)	3 (2.91%)	6 (6.18%)	6 (10.17%)
1 finite clause	0	4 (3.88%)	3 (3.09%)	2 (3.39%)
1 finite clause and phrase(s)	0	2 (1.94%)	3 (3.09%)	2 (3.39%)
Total with a finite clause	0	6 (5.82%)	6 (6.18%)	4 (6.78%)
Multiple verb phrases and/or clauses	0	2 (1.94%)	0	2 (3.39%)
Total with clause	1 (4.17%)	15 (14.56%)	16 (16.49%)	14 (23.73%)
Total with intervention	2 (8.33%)	31(30.10%)	30 (30.93%)	22 (37.29%)

Feature intervening between antecedent and relative	<i>Henry VIII</i> (% of s.w.a.'s)	<i>The Two Noble Kinsmen</i> (% of s.w.a.'s)
1 phrase	9 (21.95%)	2 (5.13%)
2 or more phrases	1 (2.44%)	1 (2.56%)
Total with phrase	10 (24.39%)	3 (7.69%)
1 non-finite clause	3 (7.32%)	1 (2.56%)
1 non-finite clause and phrase(s)	0	0
Total with a non-finite clause	3 (7.32%)	1 (2.56%)
1 partial finite clause	3 (7.32%)	1 (2.56%)
1 partial finite clause and phrase(s)	1 (2.44%)	0
Total with a partial finite clause	4 (9.76%)	1 (2.56%)
1 finite clause	1 (2.44%)	3 (7.69%)
1 finite clause and phrase(s)	0	1 (2.56%)
Total with a finite clause	1 (2.44%)	4 (10.26%)
Multiple verb phrases and/or clauses	0	1 (2.56%)
Total with clause	8 (19.51%)	7 (17.95%)
Total with intervention	18 (43.90%)	10 (25.64%)

Table 7: The information parameter

Clause type	<i>All's Well That Ends Well</i> (% of relative <i>which</i>)	<i>Macbeth</i> (% of relative <i>which</i>)	<i>Antony and Cleopatra</i> (% of relative <i>which</i>)	<i>Coriolanus</i> (% of relative <i>which</i>)
Total relative <i>which</i> clauses	83	73	106	96
Restrictive	18 (21.69%)	26 (35.62%)	36 (33.96%)	32 (33.33%)
Non-restrictive	64 (77.11%)	47 (64.38%)	70 (66.04%)	64 (66.67%)
Continuative ¹⁵³	13 (15.66%)	4 (5.48%)	11 (10.38%)	11 (11.46%)
Anomalies ¹⁵⁴	1	0	0	0

¹⁵³ Continuatives are a subset of non-restrictive clauses.

Clause type	<i>Pericles</i>	<i>Cymbeline</i>	<i>The Winter's Tale</i>	<i>The Tempest</i>
	(% of relative <i>which</i>)			
Total relative <i>which</i> clauses	29	120	118	70
Restrictive	3 (10.34%)	39 (32.5%)	35 (29.66%)	14 (20.00%)
Non-restrictive	26 (89.65%)	81 (67.5%)	83 (70.34%)	56 (80.00%)
Continuative	3 (10.34%)	20 (16.67%)	27 (22.88%)	14 (20.00%)
Anomalies	0	0	0	0

Clause type	<i>Henry VIII</i>	<i>The Two Noble Kinsmen</i>
	(% of relative <i>which</i>)	(% of relative <i>which</i>)
Total relative <i>which</i> clauses	48	46
Restrictive	10 (20.83%)	14 (30.43%)
Non-restrictive	38 (79.17%)	32 (69.56%)
Continuative	13 (27.08%)	5 (10.87%)
Anomalies	0	0

¹⁵⁴ These are occurrences of *which* that are relatives, but neither restrictive nor non-restrictive (only the aposiopesis case discussed in 3.3.6.5).

Table 8: Analysis of continuative clauses

	<i>All's Well That Ends Well</i> (% of contins.)	<i>Macbeth</i> (% of contins.)	<i>Antony and Cleopatra</i> (% of contins.)	<i>Coriolanus</i> (% of contins.)
Total continuative clauses	13	4	11	11
Antecedent 2 words distant	1 (7.69%)	0	0	2 (18.18%)
3 words	1 (7.69%)	0	1 (9.09%)	1 (9.09%)
4 words	0	0	0	1 (9.09%) (20.00%)
5 words	2 (15.38%)	0	1 (9.09%)	0
6 words	0	0	0	0
7 words	0	0	0	1 (9.09%)
8 words	0	1 (25.00%)	0	0
9 words	1 (7.69%)	0	0	0
10 or more words	0	0	0	0
Total with intervention	5 (38.46%)	1 (25.00%)	2 (18.18%)	5 (45.45%)
Total interventions with 5+ words	3 (23.08%)	1 (25.00%)	1 (9.09%)	1 (9.09%)
With determiner <i>which</i>	2 (15.38%)	3 (75.00%)	2 (18.18%)	2 (18.18%)
With sentential antecedent	4 (30.77%)	2 (50.00%)	5 (45.45%)	4 (36.36%)

	<i>Pericles</i>	<i>Cymbeline</i>	<i>The Winter's Tale</i>	<i>The Tempest</i>
	(% of contins.)	(% of contins.)	(% of contins.)	(% of contins.)
Total continuative clauses	7	20	27	14
Antecedent 2 words distant	0	1 (5.00%)	0	0
3 words	0	1 (5.00%)	1 (3.70%)	1 (7.14%)
4 words	0	0	0	0
5 words	0	0	0	0
6 words	1 (14.28%)	1 (5.00%)	0	1 (7.14%)
7 words	1 (14.28%)	1 (5.00%)	2 (7.41%)	1 (7.14%)
8 words	0	0	1 (3.70%)	0
9 words	0	2	1 (3.70%)	0
10 or more words	0	1 (5.00%)	1 (3.70%)	0
Total with intervention	2 (28.57%)	7 (35.00%)	6 (22.22%)	3 (21.43%)
Total interventions with 5+ words	2 (28.57%)	5 (25.00%)	5 (18.52%)	2 (14.28%)
With determiner <i>which</i>	1 (14.28%)	1 (5.00%)	3 (11.11%)	4 (28.57%)
With sentential antecedent	3 (42.86%)	7 (35.00%)	14 (51.85%)	5 (35.71%)

	<i>Henry VIII</i> (% of contins.)	<i>The Two Noble Kinsmen</i> (% of contins.)
Total continuative clauses	13	5
Antecedent 2 words distant	1 (7.69%)	0
3 words	0	1 (20.00%)
4 words	1 (7.69%)	0
5 words	1 (7.69%)	0
6 words	0	0
7 words	0	0
8 words	0	0
9 words	0	0
10 or more words	1 (7.69%)	0
Total with intervention	4 (30.77%)	1 (20.00%)
Total interventions with 5+ words	2 (15.38%)	0
With determiner <i>which</i>	2 (15.38%)	1 (20.00%)
With sentential antecedent	7 (53.85%)	4 (80.00%)

Table 9: The role of *which* in the relative clause: principal functions

Role of <i>which</i>	<i>All's Well That Ends Well</i> (% of relative <i>which</i>) ¹⁵⁵	<i>Macbeth</i> (% of relative <i>which</i>)	<i>Antony and Cleopatra</i> (% of relative <i>which</i>)	<i>Coriolanus</i> (% of relative <i>which</i>)
Subject	40 (48.19%)	42 (57.53%)	63 (59.43%)	49 (51.04%)
Object	33 (39.76%)	21 (28.77%)	27 (25.47%)	34 (35.42%)
Prepositional complement	9 (10.84%)	10 (13.70%)	15 (14.15%)	10 (10.42%)
Total relative <i>which</i>	83	73	106	96

¹⁵⁵ Note that these categories are not all mutually exclusive. See 4.4 for explanation and details.

Role of <i>which</i>	<i>Pericles</i> (% of relative <i>which</i>)	<i>Cymbeline</i> (% of relative <i>which</i>)	<i>The Winter's Tale</i> (% of relative <i>which</i>)	<i>The Tempest</i> (% of relative <i>which</i>)
Subject	Subject	15 (51.74%)	59 (49.17%)	72 (61.02%)
Object	Object	11 (37.93%)	52 (43.33%)	34 (28.81%)
Prepositional complement	Prepositional complement	3 (10.34%)	9 (7.5%)	9 (7.63%)
Total relative <i>which</i>	29	120	118	70

Role of <i>which</i>	<i>Henry VIII</i> (% of relative <i>which</i>)	<i>The Two Noble Kinsmen</i> (% of relative <i>which</i>)
Subject	Subject	23 (47.92%)
Object	Object	13 (27.08%)
Prepositional complement	Prepositional complement	12 (25.00%)
Total relative <i>which</i>	48	46

Table 10: The role of *which* in the relative clause: additional constructions

Role of <i>which</i>	<i>All's Well That Ends Well</i> (% of relative <i>which</i>)¹⁵⁶	<i>Macbeth</i> (% of relative <i>which</i>)	<i>Antony and Cleopatra</i> (% of relative <i>which</i>)	<i>Coriolanus</i> (% of relative <i>which</i>)
Relativiser-only	0	0	0	1 (1.04%)
Single pushdown	1 (1.20%)	0	2 (1.89%) (incl.1 not assigned a role elsewhere)	1 (1.04%)
Double pushdown	0	0	1 (0.94%)	1 (1.04%)
Recursion	2 (2.40%)	2 (2.74%)	1 (0.94%)	5 (5.20%)
Cataphoric <i>which</i>	0	1 (1.37%)	1 (0.94%)	0
Anomalies	1 (as above)	0	0	0
Total relative <i>which</i>	83	73	106	96

Role of <i>which</i>	<i>Pericles</i> (% of relative <i>which</i>)	<i>Cymbeline</i> (% of relative <i>which</i>)	<i>The Winter's Tale</i> (% of relative <i>which</i>)	<i>The Tempest</i> (% of relative <i>which</i>)
Relativiser-only	0	0	3 (2.54%)	0
Single pushdown	1 (3.45%)	1 (0.83%)	3 (2.54%)	0
Double pushdown	0	5 (4.17%)	6 (5.08%)	1 (1.43%)
Recursion	0	6 (5.00%)	4 (3.39%)	3 (4.28%)
Cataphoric <i>which</i>	0	1 (0.83%)	1 (0.85%)	0
Anomalies	0	0	0	0
Total relative <i>which</i>	29	120	118	70

¹⁵⁶ Note that these categories are not all mutually exclusive. See 4.4 for explanation and details.

Role of <i>which</i>	<i>Henry VIII</i> (% of relative <i>which</i>)	<i>The Two Noble Kinsmen</i> (% of relative <i>which</i>)
Relativiser-only	0	0
Single pushdown	2 (4.17%)	0
Double pushdown	2 (4.17%)	3 (6.52%)
Recursion	0	2 (4.35%)
Cataphoric <i>which</i>	0	0
Anomalies	0	0
Total relative <i>which</i>	48	46

Table 11: Number of words intervening to the right

	<i>All's Well That Ends Well</i> (% of relative <i>which</i>)	<i>Macbeth</i> (% of relative <i>which</i>)	<i>Antony and Cleopatra</i> (% of relative <i>which</i>)	<i>Coriolanus</i> (% of relative <i>which</i>)
1 word	7 (8.43%)	7 (9.59%)	10 (9.43%)	9 (9.37%)
2 words	1 (1.20%)	2 (2.74%)	2 (1.89%)	3 (3.12%)
3 words	2 (2.40%)	1 (1.37%)	5 (4.72%)	3 (3.12%)
4 words	2 (2.40%)	0	4 (3.77%)	2 (2.08%)
5 words	0	0	1 (0.94%) (4.17%)	0
6 words	1 (1.20%)	1 (1.37%)	2 (1.89%)	1 (1.04%)
7 words	0	0	0	0
8 words	0	0	0	1(1.04%)
9 words	0	0	0	0
10 or more words	1 (1.20%)	0	0	0
Total with intervention	14 (16.87%)	11 (15.07%)	24 (22.64%)	19 (19.79%)
Total interventions with 5+ words	2 (2.40%) (14.28%)	1 (1.37%) (9.09%)	3 (2.83%) (12.50%)	2 (2.08%) (10.53%)

	<i>Pericles</i> (% of relative <i>which</i>)	<i>Cymbeline</i> (% of relative <i>which</i>)	<i>The Winter's Tale</i> (% of relative <i>which</i>)	<i>The Tempest</i> (% of relative <i>which</i>)
1 word	4 (13.79%)	9 (7.50%)	8 (6.78%)	5 (7.14%)
2 words	1 (3.45%)	6 (5.00%)	6 (5.08%)	3 (4.28%)
3 words	0	10 (8.33%)	2 (1.69%)	2 (2.86%)
4 words	0	2 (1.67%)	0	1 (1.43%)
5 words	0	1 (0.83%)	1 (0.85%)	2 (2.86%)
6 words	0	1 (0.83%)	1 (0.85%)	0
7 words	0	1 (0.83%)	0	1 (1.43%)
8 words	0	0	0	0
9 words	0	0	0	0
10 or more words	0	0	0	0
Total with intervention	5 (17.24%)	30 (29.13%)	18 (15.25%)	14 (20.00%)
Total interventions with 5+ words	0	3 (2.50%)	2 (1.69%)	3 (4.28%)

	<i>Henry VIII</i>	<i>The Two Noble Kinsmen</i>
	(% of relative <i>which</i>)	(% of relative <i>which</i>)
1 word	2 (4.17%)	7 (15.21%)
2 words	4 (8.34%)	3 (6.52%)
3 words	2 (4.17%)	1 (2.17%)
4 words	2 (4.17%)	2 (4.35%)
5 words	0	0
6 words	0	0
7 words	0	0
8 words	0	0
9 words	1	0
10 or more words	0	0
Total with intervention	10 (20.83%)	13 (28.26%)
Total interventions with 5+ words	1 (2.08%)	0

Table 12: Word groups intervening to the right

Feature intervening to the right	<i>All's Well That Ends Well</i> (% of relative which)	<i>Macbeth</i> (% of relative which)	<i>Antony and Cleopatra</i> (% of relative which)	<i>Coriolanus</i> (% of relative which)
Phrase	11 (13.25%)	9 (12.33%)	18 (16.98%)	15 (15.62%)
More than one phrase	1 (1.20%)	1 (1.37%)	1 (0.94%)	1 (1.04%)
Total with phrase(s)	12 (14.46%)	10 (13.70%)	19 (17.92%)	16 (16.67%)
Non-finite clause	1 (1.20%)	1 (1.37%)	1 (0.94%)	2 (2.08%)
Non-finite clause and phrase(s)	1 (1.20%)	0	1 (0.94%)	1 (1.04%)
Total with non-finite clause	2 (2.41%)	1(1.37%)	2 (1.89%)	3 (3.22%)
Finite clause	0	0	3 (2.83%)	0
Finite clause and phrase(s)	0	0	0	0
Total with finite clause	0	0	3 (2.83%)	0
Multiple clauses	0	0	0	0
Total with clauses	2 (2.41%)	1 (1.37%)	5 (4.72%)	3 (3.22%)
Total interventions	14 (16.87%)	11 (15.07%)	24 (22.64%)	19 (19.79%)

	<i>Pericles</i>	<i>Cymbeline</i>	<i>The Winter's Tale</i>	<i>The Tempest</i>
	(% of relative <i>which</i>)			
Phrase	4 (13.79%)	12 (10.00%)	11 (9.32%)	9 (12.86%)
More than one phrase	0	3 (2.50%)	0	2 (2.86%)
Total with only phrase(s)	4 (13.79%)	15 (12.5%)	11 (9.32%)	11 (15.71%)
Non-finite clause	1 (3.45%)	9 (7.50%)	5 (4.24%)	1 (1.43%)
Non-finite clause and phrase(s)	0	2 (1.67%)	1 (0.85%)	1 (1.43%)
Total with non-finite clause	1 (3.45%)	11 (9.17%)	6 (5.08%)	2 (2.86%)
Finite clause	0	2 (1.67%)	0	1 (1.43%)
Finite clause and phrase(s)	0	0	0	0
Total with finite clause	0	2 (1.67%)	0	1 (1.43%)
Multiple clauses (of either kind, or a combination).	0	2 (1.67%)	1 (0.85%)	0
Total with clauses	1 (3.45%)	13 (10.83%)	7 (5.93%)	3 (4.28%)
Total interventions	5 (17.24%)	30 (25.00%)	18 (15.25%)	14 (20.00%)

	<i>Henry VIII</i> (% of relative <i>which</i>)	<i>The Two Noble Kinsmen</i> (% of relative <i>which</i>)
Phrase	3 (6.25%)	8 (17.39%)
More than one phrase	1 (2.08%)	0
Total with only phrase(s)	4 (8.33%)	8 (17.39%)
Non-finite clause	4 (8.33%)	3 (6.52%)
Non-finite clause and phrase(s)	0	1 (2.17%)
Total with non-finite clause	4 (8.33%)	4 (8.69%)
Finite clause	2 (4.17%)	1 (2.17%)
Finite clause and phrase(s)	0	0
Total with finite clause	2 (4.17%)	1 (2.17%)
Multiple clauses (of either kind, or a combination)	1 (2.08%)	0
Total with clauses	6 (12.50%)	5 (10.87%)
Total interventions	11 (22.92%)	13 (28.26%)

APPENDIX 11

THE RECONSTRUCTED *PERICLES* TEXT: PRINCIPAL DIFFERENCES FROM CRAIG (1914)

1. Marina's epitaph

Act IV scene IV (Craig (1914))

The fairest, sweet'st and best lies here,
Who, wither'd in her spring of year:
She was of Tyrus the king's daughter, 40
On whom foul death hath made this slaughter.
Marina was she call'd; and at her birth,
Thetis, being proud, swallow'd some part o' the earth:
Therefore the earth, fearing to be o'erflow'd,
Hath Thetis' birth child on the heavens bestow'd:
Wherefore she does, and swears she'll never stint,
Make raging battery upon shores of flint.

Scene 18 (Wells and Taylor (2005))

The fairest, sweetest, best lies here, 34
Who withered in her spring of year.
In nature's garden, though by growth a bud,
She was the chiefest flower: she was good.

2. Act 4 scene 6 (Craig (1914))

Lysimachus: Why, your herb-woman. She that sets roots of shame
and iniquity. O! You have heard something of my
power, and so stand aloof for more serious wooing. 50
But I protest to thee, pretty one, my authority shall
not see thee, or else look friendly upon thee. Come,
bring me to some private place; come, come.

Marina: If you were born to honour, show it now;
If put upon you, make the judgement good
That thought you worthy of it.

Lysimachus: How's this? How's this? Some more; be sage.

Marina: For me,

That am a maid, though most ungentle fortune
Hath plac'd me in this sty, where, since I came,
Diseases have been sold dearer than physic.
O! that the gods
Would set me free from this unhallow'd place,
Though they did change me for the meanest bird
That flies i' the purer air!

60

Lysimachus: I did not think
Thou couldst have spoke so well; ne'er dreamed
thou couldst.

Scene 19 (Wells and Taylor (2005))

Lysimachus: Why, your herb woman;
She that sets seeds of shame, roots of iniquity.
O, you've heard something of my pow'r, and so
Stand off aloof, for a more serious wooing.
But I protest to thee,
Pretty one, my authority can wink
At blemishes, or can on faults look friendly;
Or my displeasure punish at my pleasure,
From *which* displeasure, not thy beauty shall
Privilege thee, nor my affection, *which*
Hath drawn me here, abate with further ling'ring.
Come bring me to some private place. Come, come.

90

Marina: Let not authority, *which* teaches you
To govern others, be the means to make you
Misgovern much yourself.
If you were born to honour show it now;
If put upon you, make the judgement good
That thought you worthy of it. What reason's in
Your justice, who hath power over all,
To undo any? If you take from me

100

Mine honour, you're like him that makes a gap
Into forbidden ground, *whom* after
Too many enter, and of all their evils
Yourself are guilty. My life is yet unspotted;
My chastity unstained ev'n in thought. 110
Then if your violence deface this building,
The workmanship of heaven, you do kill your honour,
Abuse your justice, and impoverish me.
My yet good lord, if there be fire before me,
Must I straight fly and burn myself?
Suppose this house –
Which too many feel such houses are –
Should be the doctor's patrimony and
The surgeon's feeding; follows it that I
Must needs infect myself to give them maint'nance?

Lysimachus: How's this? How's this? Some more. Be sage. 120

Marina: For me
That am a maid, though most ungentle fortune
Have franked me in this sty, where since I came
Diseases have been sold dearer than physic –
That the gods would set me free from this unhallowed place,
Though they did change me to the meanest bird
That flies i'the purer air!

Lysimachus: I did not think
Thou could'st have spoke so well, ne'er dreamt thou could'st.

N.B. Relative pronouns in the text imported from Wilkins have been italicized for clarity.

BIBLIOGRAPHY

Primary sources

Shakespeare, W. *William Shakespeare: The Complete Works*, ed. Craig, W.J., 1914. London: Oxford University Press. [Online]. New York: Bartleby.com. 2000. Available at: <http://www.bartleby.com/70/> [Accessed 10.11.2007]

Sidney, Sir P. *The Countess of Pembroke's Arcadia (The Old Arcadia)*, ed. Duncan-Jones, K., 1985. Oxford World's Classics. Oxford: Oxford University Press.

Secondary sources

Adams, R., 1989. *Shakespeare: the Four Romances*. New York: Norton.

Adamson, S., 1999. Literary language. In Lass, R. ed., *The Cambridge History of the English Language, Volume III: 1476-1776*. Cambridge: Cambridge University Press. Ch. 7.

Adamson, S., 2007. Prescribed reading: pronouns and gender in the eighteenth century. In *Historical Sociolinguistics and Sociohistorical Linguistics*. [Online]. Available at: http://www.let.leidenuniv.nl/hsl_shl/Adamson.htm. (Accessed 01.03.2009).

Archer, D., 2007. Computer-assisted literary stylistics: the state of the field. In Lambrou and Stockwell, eds. *Contemporary Stylistics*. London: Continuum. Ch. 20.

Archer, D., 2009. Does frequency really matter? In Archer, D. ed., *What's In a Word-list? Investigating word frequency and keyword extraction*. Farnham: Ashgate. Ch.1.

Attridge, D., 1987. Closing statement: Linguistics and poetics in retrospect. In Weber, J.J. ed., *The Stylistics Reader: From Roman Jakobson to the Present*. London: Arnold, 1996, Ch. 2.

Baillie, W.M., 1974. *Computers in the Humanities*. International Conference on Computers in the Humanities (University of Minnesota, 1973). Edinburgh: Edinburgh University Press, pp.73-81.

Barber, C., 1976. *Early Modern English*. London: Andre Deutsch.

- Barnet, S., 1972. *The Complete Signet Classic Shakespeare*. New York: Harcourt Brace Jovanovich, Inc.
- Barton, A., ed. and introd. 1968. *The Tempest*, by William Shakespeare. London: Penguin.
- 1971. Shakespeare and the limits of language. In *Essays, Mainly Shakespearean*. Cambridge: Cambridge University Press, 1994, pp. 51-69.
 - 1980. Leontes and the Spider: language and speaker in Shakespeare's last plays. In *Essays, Mainly Shakespearean*. Cambridge: Cambridge University Press, 1994, pp. 161-181.
 - 1986. Enter mariners wet: realism in Shakespeare's last plays. In *Essays, Mainly Shakespearean*. Cambridge: Cambridge University Press, 1994, pp.182-203.
- Bate, J., ed. and introd. 1995. *Titus Andronicus*, by William Shakespeare. London: The Arden Shakespeare.
- Berlage, Eva., 2008. Noun Phrase complexity: a syntactic approach. In ISLE 1 (First Triennial Conference of the International Society for the Linguistics of English). Freiburg, Germany 8-11 October 2008. Albert-Ludwigs-Universitat Freiburg.
- Bevington, D., 1997. *The Complete Works of Shakespeare*. 4th edn. New York: Longman.
- Blake, N., 2002. *A Grammar of Shakespeare's Language*. Basingstoke: Palgrave.
- Brown, J.R., ed. and introd. 1955. *The Merchant of Venice*, by William Shakespeare. London: The Arden Shakespeare, 2003.
- Burrows, J.F., 1987. *Computation into Criticism: A Study of Jane Austen's Novels and an Experiment in Method*. Oxford: Clarendon.
- Burton, D., 1980. *Dialogue and Discourse: A sociolinguistic approach to modern drama dialogue and naturally occurring conversation*. London: Routledge and Kegan Paul.
- Carter, R., and Simpson, P., eds. 1989. *Language, Discourse and Literature: An Introductory Reader in Discourse Stylistics*. London: Routledge.
- Craig, W.J. ed. and introd., 1894. *William Shakespeare: The Complete Works*, (1905). London: Oxford University Press.
- Craig, H., and Kinney, A., 2009. *Shakespeare, Computers and the Mystery of Authorship*. Cambridge: Cambridge University Press.
- Culler, J., 1975. *Structuralist Poetics: Structuralism, Linguistics and the Study of*

- Literature*. London: Routledge and Kegan Paul.
- Culpeper J., 2001. *Language and Characterisation: People in Plays and Other Texts*. Harlow: Pearson.
- 2009. Keyness: Words, parts of speech and semantic categories in the character-talk of Shakespeare's 'Romeo and Juliet'. In *International Journal of Corpus Linguistics*, 14 (1), pp.29-59.
- Cusack, B., 1970. Shakespeare and the Tune of the Time. In *Shakespeare Survey*, 23, pp. 1-12.
- Dahl, M. X., 2004. *The Authorship of 'The First Part of Henry the Sixth': An Examination of Some Traditional and Non-Traditional Stylometric Attributions, in Relation to the First Folio History Play 'The First Part of Henry the Sixth'*. PhD. Bristol: University of Bristol.
- Dekeyser, X., 1984. Relativisers in early Modern English: A dynamic quantitative study. In Fisiak, J. ed., *Historical Syntax. Trends in Linguistics. Studies and Monographs 23*. Berlin: Mouton, pp. 61-87.
- Dekeyser, X., and Ingels, M., 1988. Socio-historical aspects of relativisation in late 16th century English: ca. 1550-1600. In *Studia Anglica Posnaniensia*. 21, pp. 25-39.
- Dillon, G., 1978. *Language processing and the reading of literature: Toward a Model of Comprehension*. Bloomington (IN): Indiana University Press.
- Dobson, M.W., Wells, S. eds., 2001. *The Oxford Companion to Shakespeare*. Oxford: Oxford University Press.
- Dowden, E., 1901. *Shakespeare: A Critical Study of His Mind and Art*. (12th edn.). (1875). London: Kegan Paul, Trench, Trubner.
- Edwards, P., 1958. Shakespeare's Romances: 1900-1957. *Shakespeare Survey*, 11, pp. 1-18.
- Evans, G. Blakemore, 1997. *The Riverside Shakespeare*. 2nd edn. Boston: Houghton Mifflin.
- Fischer, O. 1992. Syntax. In *The Cambridge History of the English Language, Volume II: 1066-1476*. Cambridge: Cambridge University Press. Ch. 4.
- Fish, S., 1973. What is stylistics and why are they saying such terrible things about it? In J.J. Weber, ed. *The Stylistics Reader* (1996). London: Arnold, pp. 94-116.
- Foakes, R.A., ed. and introd. 1997. *King Lear*, by William Shakespeare. London: The Arden Shakespeare (2003).

- Gossett, S., ed. and introd. 2004. *Pericles*, by William Shakespeare. London: The Arden Shakespeare.
- Harbage, A., 1969. *William Shakespeare: The Complete Works*. London: Allen Lane. The Penguin Press.
- Hart, A., 1934. Shakespeare and the vocabulary of *The Two Noble Kinsmen*. In *Shakespeare and the Homilies: And Other Pieces of Research into the Elizabethan Drama*. Melbourne: Melbourne University Press, pp. 242-257.
- Honigmann, E.A.J., 1965. *The Stability of Shakespeare's Text*. London: Edward Arnold.
- Hope, J., 1994. *The Authorship of Shakespeare's Plays: A Sociolinguistic Study*. Cambridge: Cambridge University Press.
- 2003. *Shakespeare's Grammar*. London: The Arden Shakespeare.
 - 2004. Shakespeare and language: an introduction. In Alexander, C.M.S. ed., *Shakespeare and Language*. Cambridge: Cambridge University Press, pp.1-17.
- Houston, J. P., 1988. *Shakespearean Sentences*. Baton Rouge (La): Louisiana State University Press.
- Howard-Hill, T. H., 1972. *Ralph Crane and Some First Folio Shakespeare Comedies*. Charlottesville: University Press of Virginia.
- Huddleston, R., and Pullum, G., 2002. *The Cambridge Grammar of the English Language*. Cambridge: Cambridge University Press.
- Hunston, S., 2002. *Corpora in Applied Linguistics*. Cambridge: Cambridge University Press.
- Hunt, M., 1990. *Shakespeare's Romance of the Word*. London: Associated University Presses.
- Jackson, M.P., 1990. 'Pericles', Acts I and II: new evidence for George Wilkins. In *Notes and Queries*, 235, pp. 192-196.
- 1997. Phrase lengths in 'Henry VIII': Shakespeare and Fletcher. In *Notes and Queries*, 242, pp.75-80.
 - 2002. Pause patterns in Shakespeare's verse: Canon and Chronology. *Literary and Linguistic Computing*, 16, pp. 37-46.
 - 2003. *Defining Shakespeare: Pericles as Test Case*. Oxford: Oxford University Press.
- Jakobson, R., 1960. Linguistics and poetics. In T. Sebeok, ed. *Style in Language*. Cambridge, (MA.): Massachusetts Institute of Technology Press, pp. 350-77

- Jespersen, O., 1949. *A Modern English Grammar On Historical Principles. Part III: Syntax. Second Volume*. Copenhagen: Einar Munksgaard.
- Jowett, J., 1983. New Created Creatures: Ralph Crane and the Stage Directions in ‘The Tempest’. *Shakespeare Survey*, 36, pp. 107-120.
- 2007. *Shakespeare and Text*. Oxford Shakespeare Topics. Oxford: Oxford University Press.
- Keenan, E.L., 1975. Variation in universal grammar. In Fasold, R. and Shuy, R., eds. *Analysing Variation in Language*. Washington DC: Georgetown University Press, pp. 136-49.
- Keenan, E.L., and Comrie, B., 1977. Noun phrase accessibility and universal grammar. In *Linguistic Inquiry*, 8, pp. 63-99.
- Kemp, W., 1979. On that that that that became that which which became what. In *Papers From The Fifteenth Regional Meeting, Chicago Linguistic Society*. Chicago: Chicago Linguistic Society, pp. 185-196.
- Kermode, F., ed. and introd. 1954. *The Tempest*, by William Shakespeare. The Arden Shakespeare. London: Methuen.
- 1971. *Shakespeare, Spenser, Donne*. New York: Viking.
 - 2000. *Shakespeare's Language*. London: Penguin.
- Knight, G.W., 1947. *The Crown of Life: Essays in Interpretation of Shakespeare's Final Plays*, Oxford: Oxford University Press.
- Leech, G., and Short, M., 1981. *Style in Fiction*. Harlow: Longman.
- Louw, B., 1997. “The role of corpora in critical literary appreciation”. In A. Wichman, S. Fligelstone, T. McEnery, G. Knowles, eds., *Teaching and Language Corpora*. Addison Wesley Longman: Harlow, pp. 240-251.
- Lyne, R., 2007. *Shakespeare's Late Work*. Oxford Shakespeare Topics. Oxford: Oxford University Press.
- Mahlberg, M., 2007(a). ‘Corpus stylistics: bridging the gap between linguistic and literary studies. In Hoey, M., Mahlberg, M., Stubbs., & Teubert, W. eds., *Text, Discourse and Corpora*. London: Continuum.

- 2007 (b). A corpus stylistic perspective on Dickens' 'Great Expectations.' In Lambrou and Stockwell, eds. *Contemporary Stylistics*. London: Continuum. Ch. 2.
- Mandala, S., 2007. *Twentieth-Century Drama Dialogue as Ordinary Talk: Speaking Between the Lines*. Aldershot: Ashgate.
- McDonald, R., 2006. *Shakespeare's Late Style*. Cambridge: Cambridge University Press.
- McMullan, G., 2007. *Shakespeare and the Idea of Late Writing: Authorship in the Proximity of Death*. Cambridge: Cambridge University Press.
- Moessner, L., 1992. Relative constructions and functional amalgamation in Early Modern English. In Rissanen, M., Ihalainen, O., Nevalainen, T., Taavitsainen, I., *History of Englishes: New Methods and Interpretations in Historical Linguistics*. Berlin: Mouton de Gruyter.
- Mulholland, J., 1967. 'Thou' and 'you' in Shakespeare: a study in the second person pronoun. *English Studies*, 48, pp. 1-9.
- Mustanoja, T.F., 1960. *A Middle English Syntax. Part 1: Parts of Speech*. Memoires de la Societe Neophilologique de Helsinki 23. Helsinki: Societe Neophilologique.
- Neilson, W.A., & Hill, C.J., 1942. *The Complete Plays and Poems of William Shakespeare*. Cambridge (Mass.): Riverside
- Nevalainen, T., 2006. *An Introduction to Early Modern English*. Edinburgh: Edinburgh University Press.
- Orgel, S., ed. and introd. 1996. *The Winter's Tale*, by William Shakespeare. Oxford: Oxford University Press.
- Ed. and introd. 2000. *King Lear: The 1608 Quarto and 1623 Folio Texts*. London: Penguin.
- Orgel, S., and Braunmuller, A.R., 2002. *William Shakespeare: The Complete Works*. The New Pelican Text. Middlesex: Penguin.
- Pafford, J.H.P., ed. and introd. 1963. *The Winter's Tale*, by William Shakespeare. London: The Arden Shakespeare, 2006.
- Palfrey, S., 1997. *Late Shakespeare: A New World of Words*. Oxford: Clarendon.
- Quirk, R., 1985. *A Comprehensive Grammar of the English Language*. Harlow: Longman.

- Rayson, P., 2008. From key words to key semantic domains. In *International Journal of Corpus Linguistics*, 13, (4), pp. 519-549.
- Reuter, O., 1938. On continuative relative clauses in English. A feature of English syntax and style ascribed to Latin influence. In *Commentationes Humanarum Litterarum* (Societas Scientiarum Fennica), Tomus IX (3), pp. 1-61.
- Richards, J. and Knowles, J. eds., 1999. *Shakespeare's Late Plays: New Readings*. Edinburgh: Edinburgh University Press.
- Rissanen, M. (1984) in Fisiak *Historical Syntax Trends in Linguistics. Studies and Monographs* 23 Berlin: Mouton.
- Rissanen, M., 1999. Syntax. In Lass, R. ed., *The Cambridge History of the English Language, Volume III: 1476-1776*. Cambridge: Cambridge University Press. Ch.4.
- Romaine, S., 1980. The relative clause marker in Scots English: diffusion, complexity and style as dimensions of syntactic change. In *Language and Society*, 9, pp. 221-49.
- 1982. *Socio-historical Linguistics, its status and methodology*. Cambridge: Cambridge University Press.
- Ryan, K., 1999. *Shakespeare: The Last Plays*. Harlow: Longman.
- Ryden, M., 1966. *Relative Constructions in Early Sixteenth Century English, With Special Reference to Sir Thomas Elyot*. Uppsala: Almqvist and Wiksells.
- 1983. English relatives revisited. In *Moderna Språk*, 77 (1983), pp. 209-18.
- Salmon, V., and Burness, E. eds., 1987. *A Reader in the Language of Shakespearean Drama*. Amsterdam: John Benjamins.
- Scott, M., 2008. *WordSmith Tools Version 5.0*. Liverpool: Lexical Analysis Software.
- 2009. In search of a bad reference corpus. In Archer, D. ed. *What's In a Wordlist?* Farnham: Ashgate. Ch. 6.
- Shakespeare, W., *The Norton Facsimile: The First Folio of Shakespeare*, ed. Hinman, C., 1996. (2nd edn.). New York: W.W. Norton & Company.
- Short, M., 1996. *Exploring the Language of Poems, Plays and Prose*. Harlow: Pearson.
- Smith, M.W.A., 1990. Shakespearean chronology: a new approach to the method of word-links. In *Notes and Queries*, 235, pp.198-204.

- 1991. Statistical inference in 'A Textual Companion to the Oxford Shakespeare' *Notes and Queries*, March 1991 (38), pp.73-78.
- Strachey, L., 1904. Shakespeare's final period. In *Independent Review*, 3. (August 1904), pp. 414-429.
- Stubbs, M., 2005. Conrad in the computer. In *Language and Literature*, 14, (1), pp. 5-24.
- Sutherland, J., 1959. The Language of the Last Plays. In J. Garrett, ed. *More Talking of Shakespeare*. London: Longmans, pp.144-158.
- Sykes, H.D., 1919. *Sidelights on Shakespeare*. Stratford-upon-Avon: The Shakespeare Head Press.
- Tarlinskaja, M., 1987. *Shakespeare's Verse: Iambic Pentameter and the Poet's Idiosyncrasies*. American University Studies, 41. New York: Peter Lang.
- Taylor, E.W., 1976. Shakespeare's use of *eth* and *es* endings of verbs in the First Folio. *CLA Journal*, XIX (4), pp. 437-457.
- Taylor, G. & Wells, S., 1987. *William Shakespeare: A Textual Companion*. Oxford: Clarendon.
- Thompson, A., and Taylor, N., eds. and introd. 2006. *Hamlet* by William Shakespeare. London: The Arden Shakespeare.
- Tillyard, E.M.W., 1938. *Shakespeare's Last Plays*. London: Chatto and Windus.
- Toolan, M., 1990. *The Stylistics of Fiction*. London: Routledge.
- 2009. *Narrative Progression in the Short Story: A Corpus Stylistic Approach*. Linguistic Approaches to Literature. Amsterdam: John Benjamins.
- Traversi, D., 1954. *Shakespeare: the Last Phase*. London: Hollis and Carter.
- Tucker Brooke, C.F., ed. 1908. *The Shakespeare Apocrypha*. Oxford: Clarendon. [Online] Project Gutenberg etext #1542 (November 1998). Available at: <http://www.gutenberg.org/etext/1542> [Accessed 20.11.2007].
- UCREL log likelihood wizard, at <http://ucrel.lancs.ac.uk/llwizard.html>. (Last consulted 18.10.10).
- Vaughan, V.M. & Vaughan, A.T., eds. and introd. 1999. *The Tempest*, by William Shakespeare. London: The Arden Shakespeare.
- Vickers, B., 2002. *Shakespeare Co-Author*. Oxford: Oxford University Press.

- 2008. Approaching Shakespeare's Late Style. In *Early Modern Literary Studies*, [Online].13.3 (January, 2008) 6.1-26. Available at: <http://purl.oclc.org/emls/13-3/revmcd.htm> [Accessed 3 March 2008].

Wales, K., 1978. An aspect of Shakespeare's dynamic language: a note on the interpretation of King Lear III. VII. 113: 'He childed as I fathered!' *English Studies*, 59, pp. 395-404.

Warren, R., 1990. *Staging Shakespeare's Late Plays*. Oxford: Clarendon.

Warren, R., ed. and introd. 1998. *Cymbeline*, by William Shakespeare. Oxford World's Classics. Oxford: Oxford University Press.

 - ed. and introd. 2003. *Pericles*, by William Shakespeare. Oxford World's Classics. Oxford: Oxford University Press.

Weber, J.J., ed. 1996. Introduction. In *The Stylistics Reader: From Roman Jakobson to the Present*. London: Arnold.

Wells, S., & Taylor, G., 1987. *William Shakespeare: a Textual Companion*. Oxford: Oxford University Press

Wells, S., and Taylor, G., 2005. *William Shakespeare: The Complete Works*, 2nd. edn. Oxford: Oxford University Press.

Wiggins, M., Richardson, C., and Merry, M. (forthcoming). *British Drama 1533-1642: a Catalogue*. Oxford: Oxford University Press.

Witmore, M., and Hope, J., 2007. Shakespeare by the numbers: on the linguistic texture of the late plays. In R. Lyne and S. Mukherji, eds. *Early Modern Tragicomedy*. Cambridge: Brewer. Ch.9.

Woudhuysen, H.R., ed. and introd. 1998. *Love's Labour's Lost*, by William Shakespeare. London: The Arden Shakespeare.

Yates, F., 1975. *Shakespeare's Last Plays: A New Approach*. London: Routledge and Kegan Paul.