Charmed and chattering tongues: Investigating the functions and effects of key word clusters in the dialogue of Shakespeare's female characters

by Jane Elizabeth Judson Demmen

Lancaster University

Dissertation submitted in total fulfilment of the requirements

for the MA in English Language by Research

Submitted August 2009

ABSTRACT

This study is a corpus-based investigation of "key" word clusters (recurrent word combinations of statistical significance) in the dialogue of male and female characters in Shakespeare's plays, with particular focus on the female characters. The results show that women and men are represented as using language differently in some respects, through variation in the kinds of formulaic language (identified through the key word clusters) which they tend to use statistically frequently.

Unlike many previous studies of Shakespeare's plays, this one is empirically based, and the quantitative key cluster results are derived from **all** the dialogue in the plays (not from selected extracts or selected characters). Using these as a starting point, I categorise their functions in the data using a classification system adapted from Culpeper and Kytö (forthcoming), based on Halliday's (e.g. 1994) interpersonal, textual and ideational metafunctions of language. I then conduct a detailed qualitative analysis of key word clusters which show the most interesting concentration patterns in the female dialogue. I discuss these in the context of the male dialogue, noting some contrasts between the language of women and men in the plays.

The main findings of the study concern the stylistic characterisation of women and men (i) in general (which I term "global effects") and (ii) as individuals in the collection of plays. However, I also suggest that some trends in the results may be linked to historical sociolinguistic variation, reflecting Shakespeare's contribution to debates about the role and behaviour of women at the time the plays were written. In addition, the study indicates that some dramatic devices which operationalise the communication of the play to the audience are more prevalent in the female dialogue.

ACKNOWLEDGEMENTS

I am very grateful to the Arts and Humanities Research Council for funding my MA in

English Language by Research, the outcome of which is this dissertation.

My study was supervised by Dr. Jonathan Culpeper (Lancaster University), whose

advice, help, encouragement and great good humour have once again been invaluable.

Dr. Mike Scott (Liverpool University) kindly gave me access to his corpus of

Shakespeare's plays for this study. This considerably accelerated the start of the

project, and allowed me to spend far more time on the detailed analysis of my results

than would otherwise have been possible.

I could not have undertaken the MA without the support and encouragement of my

husband, Phillip Demmen. For this, and much more, I am immensely grateful.

Finally, I would also like to thank the following people (in alphabetical order), for so

willingly giving me their time and help during the study:

    Mr. Shaun Austin
    Mr. Alistair Baron
    Dr. Andrew Hardie
    Mrs. Carole Judson
    Mr. Robert Judson
    Dr. Ursula Lutzky
    Dr. Paul Rayson
    Mr. Brian Walker
    Mrs. Marjorie Wood

Jane Judson Demmen, August 2009

TABLE OF CONTENTS

TABLE OF CONTENTS continued

TABLE OF CONTENTS continued

1. OPENING LINES: INTRODUCING THIS STUDY

1.1 Introduction

Shakespeare's plays have a longstanding literary and cultural presence in the English language, and they also continue to generate academic interest and discussion. In particular, much has been written about the ways women are represented, and there is a wealth of feminist literary criticism on this aspect of the plays (e.g. Kahn, 1977 and 1981; French, 1981; Jardine, 1983; Dusinberre, 1996; Findlay, 1999). Few linguists, in comparison, have thus far contributed to the debate on language and gender in Shakespeare's plays, though Crystal (2008:221) claims that "Shakespeare gives us a remarkable picture of the range of social situations in Elizabethan England", and Culpeper (2000:18) argues that the plays make a "contribution to the vibrant debate at the time about the role of women in society".

Crystal (2003:62) states that although scholars have been studying Shakespeare's works for many years, "linguistically-inspired approaches such as stylistics, pragmatics, sociolinguistics, and computational linguistics" are more recent. Stubbs (2005) argues that corpus methodology reveals aspects of interest in literary texts which have not been noticed by literary critics using manual analysis, a view which is echoed by others including Mahlberg (2007a, 2007b) and Starcke (2005, 2006).

A number of important corpus studies of Shakespeare's plays have already been made (e.g. U. Busse, 2002; B. Busse, 2006; D. Crystal and B. Crystal, 2002; Crystal, 2008; Culpeper, 2002; Scott and Tribble, 2006). However, to date there have been no detailed corpus-based comparisons focusing solely on the language used by female and male characters across the collection of plays. My study aims to help fill this gap,

in order to add some empirically-based discussion to the wealth of non-corpus-based commentary on language and gender in the plays. Mine is a corpus-based stylistics study, i.e. one which utilises quantitative data derived by electronic means to address linguistic description in literary texts. Touching on all the areas mentioned by Crystal (2003:62), above, I will investigate how male and female characters in the plays are characterised through what they tend to say relatively frequently in their dialogue. To achieve this, I will identify the statistically significant "key word clusters" which occur in the dialogue of male and female characters from all the plays.

A "cluster" is a type of electronically-derived recurrent word combination (hereafter "RWC"); others include n-grams and lexical bundles. Culpeper and Kytö (forthcoming) argue that although there are compatibility issues, they are all essentially based on statistical calculations of how often words occur and co-occur in texts. In this study, clusters are the RWCs identified by Scott's (1999) corpus analysis software programme *WordSmith Tools*[1] (hereafter "*WordSmith*"), as defined in the programme's Help menu (see also section 1.5 of this study). "Key" results (single "keywords" or combinations of words) are those which occur as statistically significant in one text when compared to a reference corpus of other texts (again see section 1.5). (Hereafter, references to numbered sections are abbreviated, e.g. "S.1.5".)

Taking the quantitative key cluster results as a starting point, I will analyse them qualitatively by classifying their functions in the dialogue, using a framework adapted from Culpeper and Kytö (forthcoming). This is based on Halliday's (e.g. 1994) "interpersonal", "textual" and "ideational" language functions (see S.3.4). I will then

---

[1] See http://www.lexically.net/wordsmith/index.html (accessed 12.08.2009)

conduct detailed stylistic analyses of the patterns of results with different functions, using concordance data from *WordSmith*, to compare the ways women and men in the plays use language relatively frequently. In this way, my corpus-based study of key word clusters will enable me to see what is distinctive in the language used by characters of each sex, based on empirical evidence from the dialogue of **all** male and female characters in Shakespeare's plays (that is, all in the particular edition which I use, see S.3.1; Appendix I lists the plays which are included). Following Talbot (1998:7), amongst others, in this study "sex" refers to a biological characteristic accorded to people in plays, and "gender" means the socially-based set of characteristics which are constructed through their language and other behaviour.

I often allude to "female (and male)" data to make clear that my main focus is on female characters, though in the context of the male characters with whom they interact in all the plays. A discussion of female dialogue would be incomplete without discussing male dialogue, to the extent that space and the structure of the corpus allow (see S.3.2), and some interesting contrasts are analysed in Chapter 4.

In the rest of this introductory chapter, I present the case for carrying out the study in S.1.2, followed by an outline of how it will proceed and what it will include (S.1.3). In S.1.4 I set out my formal research questions, and in S.1.5 I clarify some definitions of terms and concepts used. In Chapter 2 I explain the background to this study, discussing relevant theories and existing research. In Chapter 3 I give full details of my quantitative and qualitative methodology, including the framework of functional classification for my results. In Chapter 4 I present the key cluster results, categorised by function, and I analyse and discuss them using numerous examples from the plays.

Finally, in Chapter 5, I summarise my findings and offer some conclusions about the outcomes of the study.

1.2   The need for a corpus-based study of language and gender in Shakespeare's plays

As noted in the previous section, the representation of women in Shakespeare's plays is a topic of longstanding debate and literary criticism. At first, it is difficult to see how a corpus linguistics study can be linked to what has been said in non-corpus-based disciplines such as literary criticism. However, as I state below and argue further in Ss.2.5 and 2.6, one approach can inform the other.

As Crystal (2008:230-231) argues, Shakespearean dialogue is closely tied to natural early modern English ("EME"), just as present-day English ("PDE") is to contemporary drama (for definitions of historical periods, see S.1.5). However, as Short (1996:173-179) points out, dramatic dialogue differs in some crucial ways from natural speech, notably because it is artificially constructed by one or more writers. The dialogue in Shakespeare's plays has a primary dramatic purpose as part of an instrument for communicating a story to an audience in a way which informs, entertains, shocks and/or takes a position (e.g. on some political, cultural or social issue). It is therefore necessary to question constantly the stylistic purpose of any apparent sex-based differences (or similarities) in the male and female dialogue. As my analyses in Chapter 4 show, this is often a complex and multi-layered process.

It is far beyond the scope of this study to review all the existing non-corpus-based research concerning language, gender and the portrayal of women in Shakespeare's plays, though I provide some relevant background to this in S.2.5. Instead, my aim is

to show the male and female dialogue in the plays in a new light by taking a fresh, empirically-based approach which has not been used by other scholars. In Chapter 4 I mention the findings and claims of other non-corpus-based studies where these can usefully be compared with my own results, but I make no attempt at direct comparisons between my results and those from critical literary studies, following Busse's (2006:51) similar principle in her corpus study of vocatives in Shakespeare's plays. I concur with linguists such as Semino and Short (2004:7-9) and literary critics such as Jardine (1983:ix), who argue that there is a place for both approaches. A quantitatively-based study shows what the language in the plays is like from a statistical perspective – hugely useful, as I demonstrate in Chapter 4 – but inevitably many interesting aspects of the dialogue will simply never occur as statistically significant, and will therefore go undiscussed with this approach. This kind of limitation is also acknowledged by other corpus linguists working with literary texts, including Busse (2006:10) and Mahlberg (2007b).

In contrast, the absence of an empirical basis to most critical literary studies avoids the limitation of what can be included, but runs the risk of accusations of selectivity and subjectivity in choosing which parts of the dialogue to base claims upon. Yet the unceasing flow of literary criticism of Shakespeare's plays is a reason the plays have remained popular: over several hundred years people have continued to find new ways of using them to interpret social, political and cultural life. For me, that in itself makes the language in the plays worth studying from an alternative, empirical perspective, on top of the benefits of analysing literary texts using quantitative data (argued by corpus stylisticians such as Wynne, 2006; Stubbs, 2005; and Semino and Short, 2004). However, the quantitatively-based approach is vastly under-represented at the present

time, particularly with regard to the analysis of language and gender in Shakespeare's

plays. As stated in S.1.1, in this study I seek to help redress the balance.

The route to successful interpretation of the representation of male and female

characters in plays from statistically significant quantitative data in dramatic dialogue,

though worthwhile (as I argue above), is neither brief nor particularly easy. Linguists

specialising in stylistic analysis of drama such as Short (1998:12) and Culpeper

(2001:39-42) argue that plays are designed to be performed, and as they point out,

dialogue is only one of several aspects of performance. Audiences interpret drama

through a combination of verbal, visual and environmental influences, whereas

scholars and readers have only the written text. Nevertheless, both Short (1998) and

Culpeper (2001) take the view that, with careful analysis, dramatic texts can be

successfully interpreted without watching them performed. This is demonstrated in the

outcomes of corpus studies of literary texts by Culpeper (2002), and others mentioned

in the next section and in Chapter 2. However, I shall keep in mind Culpeper's

(2001:42) point that readers of dramatic texts may not interpret them in identical ways

to audiences who see a performance.

Since my study essentially asks how men and women are characterised in the plays by

the language they use relatively frequently, it is essential to mention the

characterisation process itself. I must consider what the relationship is between the

things characters tend to say frequently, and the ways this creates impressions of them

as members of collective social groups and as individuals. Culpeper and McIntyre

(2006:779) argue that "dialogue is social interaction, and it is in social interaction that

character is displayed and shaped." Clearly a qualitative analysis of quantitative results

6

needs to encompass the social aspects of language, such as linguistic politeness and deference, hence my decision to categorise the key cluster results into a framework which takes into account pragmatic and discoursal functions (discussed in S.3.4). As I explain further in S.2.1, my analysis of the patterns of results in the male and female dialogue is particularly informed by Culpeper's (2001) work on characterisation, since he focuses extensively (though not exclusively) on Shakespeare's plays. In particular, Culpeper (2001) argues that character impressions are influenced by the audience or reader's existing knowledge and schematic assumptions:

    (i)   about the ways real people behave;

    (ii)   about the ways familiar and (stereo)typical dramatic characters behave; and

    (iii)   about the ways members of particular social groups behave.

For example, Culpeper states that "gender is a fundamental social category which people use in making sense of others" (ibid.:12). This is highly relevant to the study of characters in drama. It is reasonable to suppose that the kinds of language men and women in the plays use statistically frequently will have some relationship with an early modern audience's understanding of what is possible and appropriate social behaviour in real life, and in dramatic worlds. This must be so, since it constructs the frame of reference essential to grasping the story and being able to tell when characters are displaying unusual behaviour with an intended dramatic effect (i.e. creating an individual character trait, or showing a character's response to a particular situation).

Since Culpeper (2000:18) argues that Shakespearean dialogue is part of early modern commentary on women in social life at that time (noted in S.1.1 above), I anticipate that my results will include some traces of stylistically unusual language used by female and male characters. This would have been potentially thought-provoking to the audience, perhaps challenging their existing assumptions about women and men.

1.3 The approach taken in this study

Linguists such as Coates (2007) and Talbot (1998) argue that variationist sociolinguistics studies in PDE show that men and women use language differently. However, many studies of naturally-occurring language have now moved on from what Cameron (2005:487) terms a "binary gender *difference*" approach (her italics), i.e. assuming biological sex to be the underlying cause of variation, to one which allows for the construction of male/female gender based on social factors, and for variety within groups of speakers of the same biological sex (see also Culpeper, 2001:16-17). One of the ways dramatic dialogue is not like natural speech (argued in S.1.2) is that social factors are mapped on to biological sex by the dramatist(s). Therefore, there is a strong argument for investigating dramatic dialogue in terms of binary difference, to see how the language of male and female characters is actually portrayed by the writer(s) – in Shakespeare's case, a male writer. It is important to avoid predicting corpus-based results, since this can lead to a biased analysis, but I believe that there will be at least some identifiable differences in the linguistic content of the male and female dialogue that will surface in key word clusters, based on:

(i)   the extent of the language and gender debate on the representation of women in the plays;

(ii)   the fact that they originate from a male writer's perspective.

The nature of the variation cannot be predicted in advance, but a smaller, previous pilot study (Demmen, 2007) did confirm that the methodological approach of investigating Shakespearean dialogue through a lexically-focused electronic analysis was suitable for application to a larger study. I can therefore anticipate producing a useful set of results. The pilot study examined key "lexical bundles" in a sample of Shakespeare's comedy, tragedy and history plays. (A lexical bundle is another type of RWC, defined by Biber et al., 1999; Scott and Tribble, 2006:12, 32 argue that it is effectively the same as a word cluster.) This successfully showed some differences in the ways characters tend to use language, contributing to the construction of different kinds of dramatic effects in each genre. The pilot study did not extend to the breakdown of any dialogue by sex of characters – my main focus in the present study. The present study involves a much larger corpus (all the plays, as stated in S.1.1), requiring annotation that has not previously been carried out, to split it into sub-corpora of male and female dialogue (discussed in S.3.2).

I had initially intended to focus the present study on female characters who are wives in the plays, but it soon became clear that taking a social role as a starting point would be problematic. It would have been difficult to define clearly which characters should be included, for two reasons. Firstly, characters have multiple social roles which influence their dialogue (e.g. Goneril and Regan are both wives, but their social roles as daughters of King Lear seem more prominent). Secondly, few female characters are wives for the duration of a play; often, as a result of the plot, wifehood begins part way through the play or ceases through widowhood. I concluded that examining

9

variation based on a social role was not feasible in this corpus study, since it could not

be addressed in a sufficiently systematic way.

Furthermore, limiting the study to wives only would not have made the best use of the

corpus of all Shakespeare's plays to which I had access (see S.3.1). Busse (2002:9)

argues that a corpus of Shakespeare's plays constitutes "a special case of a closed and

complete set of data", so it seemed preferable to make the most of all the available

dialogue in the corpus, not just that of a few characters. This, and the absence of any

detailed corpus-based studies comparing the language of male and female characters

in the plays (argued in S.1.1 above), led to my decision to focus on characters of

different sex rather than social role. Social role is, however, an important factor in my

analyses (see e.g. S.4.3.2).

Having explained the reasons behind my focus upon female and male dialogue in the

plays, I will now briefly justify the choice of key word clusters as my quantitative

data, and explain how my study differs from others in related research areas.

As I discuss further in S.2.2, the investigation of RWCs in texts is underpinned by two

important ideas, developed by Sinclair (1966, 1991, 2004) and later Hoey (1991,

2005) and Wray (2002); see also Scott and Thompson (2001):

(i)   language is managed in ready-made "chunks" or sequences of words; and

(ii)   these chunks form patterns of co-occurrence and repetition in texts.

Such patterns are now routinely traced and investigated through the analysis of collocations, colligations and concordance data in corpora. Collocational relationships underlie RWCs such as key word clusters and lexical bundles, as explained in S.2.2.

Stubbs (2005:13) recommends the investigation of key **combinations** of words in literary texts in order to learn more about pragmatic and discoursal aspects of dialogue and narrative, since these are usually created by words working together, rather than alone. Stubbs (2005), Mahlberg (2007a, 2007b) and Starcke (2005, 2006) have shown that studies of RWCs reveal new and interesting insights into the language of well-known prose fiction novels, revealing what is distinctive about them (as discussed further in S.2.3). Some of the language features Mahlberg (2007a:9, 19) identifies in her corpus of novels by Charles Dickens coincide with those which have already been mentioned by literary critics, and as Mahlberg says, "the computer can help to trace and analyse them systematically".

The works of dramatists and playwrights seem neglected by comparison (though Culpeper and Kytö's forthcoming study includes lexical bundles and EME drama). Yet drama seems particularly well suited to investigation via the formulaic language which surfaces in RWCs. Although it differs from natural speech (mentioned in S.1.1), Blake (2002:283-290, citing Gilbert, 1997) claims that it is more formulaic than natural speech. Blake identifies a number of recognisable "conversational strategies" in Shakespeare's plays (summonses, responses, greetings, pre-request formulas, announcements, pre-parting, post-parting and parting formulae), and argues that these help audiences engage with the plot (2002:283-290). It seems reasonable to anticipate that some fragments of these conversational strategies will occur in key word cluster

11

results from the male and female characters' dialogue, so I should be able to see whether any are particularly associated with female (or male) characters.

Crystal (2008:173-175) has recently called for further research into collocations in Shakespeare's plays, strengthening my view (in S.1.1) that to date the plays have been studied little, if at all, from this perspective. The studies of keywords in Shakespeare's plays by Culpeper (2002) and Scott and Tribble (2006) have highlighted interesting stylistic features that add to the understanding of characterisation in the plays, using examples of individual male and female characters. However, they do not make a systematic comparison of differences in the dialogue of all men and women in the plays, and their work does not extend to key **combinations** of words.

Culpeper and Kytö (forthcoming) include a range of EME genres in their research on RWCs in historical texts, though they do not focus in detail on any single author as I do here. They usefully discuss how to address potential hazards such as EME spelling variation (see S.3.2.5 of this study), and as indicated in S.1.1 their system of functional classification of results proves adaptable for my study (see S.3.4). Hota et al. (2006) and Argamon et al. (2007) do include male/female variation and some data on RWCs in their corpus studies of Shakespearean dialogue, but as I argue in S.2.4, their methodology differs considerably from mine. Their studies are oriented towards the development of text mining software, whereas I seek to provide detailed qualitative stylistic analyses of dialogue in the plays using well-established corpus analysis software tools. I follow other linguists including Scott and Tribble (2006), Culpeper and Kytö (forthcoming), Mahlberg (2007a) and Culpeper (2002) in using *WordSmith* (discussed in S.3.3).

In this section I have pointed out the current lack of empirically-based studies of language and gender in Shakespeare's plays and justified the rationale behind my approach. Having introduced the study, the research gap which it aims to address, and the way I intend to carry it out, I am now in a position to set out my formal research questions.

1.4   <u>Research questions</u>

In Ss. 1.1 to 1.3 I touched briefly on existing research into:

- the language of women in Shakespeare's plays
- the characterisation process
- the study of literary texts using corpus methodology and
- RWCs.

From this, I anticipate finding empirical variation between the male and female dialogue in the plays, with explanatory factors likely to be both social and stylistic. This is because:

(i)   longstanding literary critical arguments, especially feminist ones, exist about the distinctive ways in which women are represented in the plays;

(ii)   dialogue in the plays would have been constructed in relation to schemas about male and female behaviour which is possible (a) in real life and (b) in drama;

13

(iii)   Shakespeare would have included some deliberately unexpected ways of using language to create impressions of individual characters' personalities and the relationships between characters, based on his own schemas and in anticipation of those held by his audiences.

To investigate these ideas, I will need to answer the following research questions:

1.4.1   What key word clusters are present in the dialogue of female (and male) characters in Shakespeare's plays?

1.4.2   What are the functions of the key word clusters in the female (and male) dialogue?

1.4.3   What patterns of functions are evident in clusters which occur in female (and male) dialogue, e.g. concentrations or contrasts?

1.4.4   What character impressions are created in the dialogue through the formulaic language female (and male) characters tend to use frequently, as evidenced by the data in the key clusters?

1.4.5   Does statistically significant formulaic language in the female (and male) dialogue contain traces of language and gender issues of the historical period in which Shakespeare's plays were written, as distinct from stylistic effects with a dramatic purpose?

Before proceeding further, I will clarify some definitions and potentially ambiguous terms which I use.

1.5   <u>Definitions and conventions used in this study</u>

In addition to those given in the preceding sections, the following definitions apply in

this study unless otherwise stated:

- Fennell's (2001:1) definition of early modern English, from 1500-1800;
  Fennell defines modern or present-day English as being from 1800 onwards
  (ibid.), but in this study present-day English refers to usage from the 20th
  century onwards;

- the early modern social rank structure set out by the historical sociolinguists
  Nevalainen and Raumolin-Brunberg (2003:28-43);

- Blake's (2002:13) term "Shakespeare's English (ShE)", distinguishing usage
  in the plays from general usage in EME;

- Sinclair and Coulthard's (1975:27-28) concept of "discourse", i.e. conveying
  the "functional" aspects of language compared to the "formal" aspects
  conveyed by grammar; also their concept of "discourse acts", i.e. utterances
  which elicit, direct or inform (ibid.:28);

- "text" means any spoken or written "language event" (based on Scott and
  Thompson, 2001:4, who also cite Hoey, 1991);

- "play-text" refers to the data under consideration in this study, following
  Culpeper and McIntyre (2006:775), to emphasise that I am discussing a written
  form of the plays, not a performed form; and

- Wales's (1989) definitions of stylistic and general linguistic terms, e.g. "co-
  text" means words surrounding other words in a text, and "context" refers to
  wider situational circumstances.

There are many specific terms associated with corpus linguistics, and it is useful to

define a few which are particularly important in this study:

- I use Baker et al.'s (2006) definitions of general corpus linguistics terms (e.g.

  "tags", "annotation", "reference corpus", "collocation", "precision" and

  "recall").

- I use the term "word" to mean an orthographic word-form as identified by

  Scott (1999), i.e. a string of letters bounded by space or punctuation (see

  definitions in the *WordSmith* Help menu). *WordSmith* expresses some

  calculations in numbers of "tokens", but rather than make a distinction I refer

  consistently to "words". See also Sinclair (2004:131-148) and Hoey

  (2005:156-158) for further discussion of words and lexical items.

- Linguists use different definitions for types of "recurrent word combination", a

  term which itself is defined by Altenberg (1998) as "any continuous string of

  words occurring more than once in identical form". I use this (abbreviated to

  "RWC" as noted in S.1.1) as a general reference to this type of linguistic form;

  other corpus studies use the general term "multi-word units". I use other

  theorists' chosen terms when referring specifically to their work on different

  types of RWC. Space does not permit a detailed discussion of all these, but see

  Culpeper and Kytö (forthcoming) for a helpful summary. As noted in S.1.1,

  there are issues of compatibility between different types of RWC, hence the

  need for clear definitions.

- It is important to emphasise that my results are all **key** word clusters, though

  for economy I often refer to them as "word clusters" or simply "clusters". In

  the *WordSmith* Help menu, Scott (1999) states that "[c]lusters are words which

  are found repeatedly in each others' company" and he argues that they are

products of relationships such as collocation, colligation or semantic prosody. I discuss the theoretical background to word clusters more in Chapter 2, but those which form the results in this study are recurrent string of words identified by the *WordSmith* corpus software as occurring with statistically significant frequency in one dataset when compared with another (as a reference corpus).

- I use Scott's (1999) definitions of "key" and "keyness" (see also Baker et al., 2006:97-98; Baker, 2004; Culpeper, 2002:13-14). Key results (words or word clusters) are those occurring with unusually high (positively key) statistical frequency or low (negatively key) statistical frequency in one text when compared to a reference corpus of other texts. Culpeper (ibid.:13) notes that these are distinct from Williams's (1976) culturally-derived keywords (see also Stubbs, 2001:145-169; 1996:166-195).

The contents of clusters in my results are given in single quotation marks, e.g. 'I know not'. Boldface text indicates my own emphasis, and the convention […] shows where I have omitted less relevant lines in quotations, in order to conserve space.

My functional category labels (shown in Table 3, S.3.4.2) are indicated by initial capitals, e.g. Directive, Topical, Question, to distinguish them from similar terms in general use in linguistic description. They are also shown in the order of metafunction: category: sub-category (e.g. Ideational: Topical: States clusters).

In this chapter of "opening lines", I have introduced the reasons for undertaking the study, its purpose and objectives, and the way it will be carried out. This is a corpus-

based analysis of the entire canon of one outstandingly well-known EME playwright,

which seeks to reveal how women (and to some extent men) are characterised through

the things they tend to say relatively frequently, based on the functions of statistically

significant language formulae which occur in their dialogue (in the form of key

clusters). The outcomes will add some much-needed empirically-based discussions to

the body of commentary on language and gender in Shakespeare's plays, and will add

to what is known about the ways Shakespeare created some of the memorable

individual personalities in his works. In Chapter 2, I now discuss in greater detail the

background to this study and other research which informs my own.

2. SETTING THE SCENE: BACKGROUND TO THIS STUDY

This chapter begins with a brief discussion of corpus-based stylistics, and some further

background to the process of characterisation in drama (S.2.1). In S.2.2 I introduce

theories suggesting that language is processed in "chunks" at the lexical level, and I

explain the ways these ideas have been applied by corpus linguists. I then present the

case for investigating RWCs by introducing a selection of relevant studies from a

range of language genres, including literary texts (S.2.3). In S.2.4 I focus specifically

on other studies of Shakespearean dialogue, and I explain what my own research will

contribute. S.2.5 contains a brief evaluation of other approaches to language and

gender in Shakespeare's plays. Finally, in S.2.6, I explain what existing criticism from

non-empirically-based disciplines can usefully offer my research.

2.1 Exploring literary texts via corpus analysis

As stated briefly in S.1.2, scholars such as Semino and Short (2004), Stubbs (2005),

Wynne (2006) and Mahlberg (e.g. 2007a) have applied corpus methodology in

stylistics analysis. Their arguments of the benefits of empirically-based approaches to

literary texts derive from claims that corpus methodology can locate language patterns

in texts which would be difficult or impossible to find by manual analysis. Sinclair's

(1966, 1991, 2004) work has been particularly influential here, as discussed further in

the next section. Sinclair argues that literature should be investigated in the same ways

as other language genres (2004:51). He goes on to add that "[p]atterns of language

that are not remarked upon in non-literary text are invested with meaning in stylistics"

(ibid.:70). Corpus linguists who have studied literary texts, e.g. Stubbs (2005) and

Mahlberg (2007a, 2007b), have found this to be true as I stated in S.1.3, adding a new

perspective to some of the claims made by literary critics. Both Stubbs and Mahlberg do, however, acknowledge opposing views to this approach, including:

- the apparent reduction of literary texts to a decontextualised list of features;
- claims that corpus analysis rather subjectively searches out features already known to be of interest.

Space does not permit a detailed discussion of the criticisms of corpus stylistics, but see e.g. Archer (2007). Semino and Short (2004:7-9) and Mahlberg (2007a) take the view that corpus methodology complements other, more qualitative methods of stylistics analysis, and they argue that the researcher can have the best of both worlds by combining the two approaches. I propose to do this in my study, since it is clear from existing corpus studies of drama including Shakespeare's plays (e.g. Culpeper, 2002; U. Busse, 2002; B. Busse, 2006; Scott and Tribble, 2006) that patterns of quantitative data can usefully aid research into characterisation and other dramatic effects. However, Stubbs (2005:6) emphasises that such patterns require careful qualitative analysis in order to learn something useful about the text(s) under consideration.

Corpus methodology offers the advantage of basing an interpretation of the male and female dialogue in Shakespeare's plays on results which are statistically important. Busse (2006) and Mahlberg (2007b) have shown that it is possible to make some links between the outcomes of corpus-based research and the claims of literary critics, although it is simply not feasible to compare these directly, nor to include all the critical literary theory on the texts under consideration (even if some findings seem to

20

coincide). As I argued in S.1.2, the corpus-based approach offers a fresh, alternative

perspective on Shakespeare's plays, based on all the dialogue in the canon –

something which would be well beyond my manual resources.

In seeking to say something about the way women are characterised through what they

tend to say relatively frequently in their dialogue (in comparison to men), it is

important that I touch first on what scholars have said about the actual process of

characterisation. How do impressions of people featuring in a drama form between the

production or instantiation of their dialogue (onstage or in the text) and the mind of the

audience or reader? The work of Culpeper (2001) and Short (1996) particularly

informs my approach to analysing the evidence for characterisation in the key cluster

results in Chapter 4. Both include discussions of Shakespeare's plays, Culpeper's

research having a particular (though not exclusive) focus upon them, as noted in S.1.2.

I also mention studies by other stylisticians who have researched characterisation in

the plays using both corpus-based and non-corpus-based methods, e.g. Bousfield

(2007) and Cooper (1998).

Space allows for just a brief introduction to some ideas which are especially important

to my study. These are:

(i)   the cognitive aspects of characterisation argued by Culpeper (2001);

(ii)   Short's (1996) concept of "discourse levels" in drama; and

(iii)   the relevance of theories about real-world social interaction, such as those

        surrounding linguistic politeness, in the analysis of dramatic dialogue.

Culpeper (2001) argues that audiences and readers construe characters through a set of cognitive mechanisms which are activated by triggers in the language of the dialogue spoken by characters (onstage or in the play-text). These triggers, which Culpeper terms "textual cues", set off the characterisation process – the registration of impressions about personality traits and behaviour of the characters in the minds of the audience/reader. Culpeper (2001) distinguishes three types of textual cues which help activate characterisation in the mind of the audience/reader:

- "explicit cues" occur in what characters say about themselves and about each other through "self-presentation" and "other-presentation" (2001:167-172);

- "implicit cues" reside in the linguistic features of characters' dialogue, i.e. their speaking styles in terms of speech structure, syntax, sound, the words and expressions they tend to use and the kinds of conversational implicature they habitually employ (2001:172-229);

- "authorial cues" are those which are imposed on the characters by the writer of the dialogue, such as their names and the actions which are written into the stage directions (2001:229-232).

Bearing in mind Blake's (2002:283-290) argument of the formulaic nature of conversational strategies in drama (mentioned in S.1.3), I anticipate that my key cluster results will show some evidence of contributing to implicit and explicit cues. I do not expect authorial cues to feature, since I use only the dialogic text in the plays (see S.3.1). It does seem reasonable to expect some evidence of communication between author and audience in the results, however, because of the different "discourse levels" which operate in drama, bringing me to my second main point in

this section. The model of discourse structure (also called "discourse architecture") put

forward by Short (1996:169-172) is now cited regularly by other stylisticians,

including Culpeper (2001:38-39), Culpeper and McIntyre (2006:775-776) and

Bousfield (2007:217). Short (1996:169) argues that drama typically features "at least

two levels of discourse, the author-audience/reader level and the character-character

level", and he represents the structure of dramatic discourse diagrammatically, as

shown in Figure 1:

Addresser 1	Message	Addressee 1
(Playwright)		(Audience/Reader)
Addresser 2	Message	Addressee 2
(Character A)		(Character B)
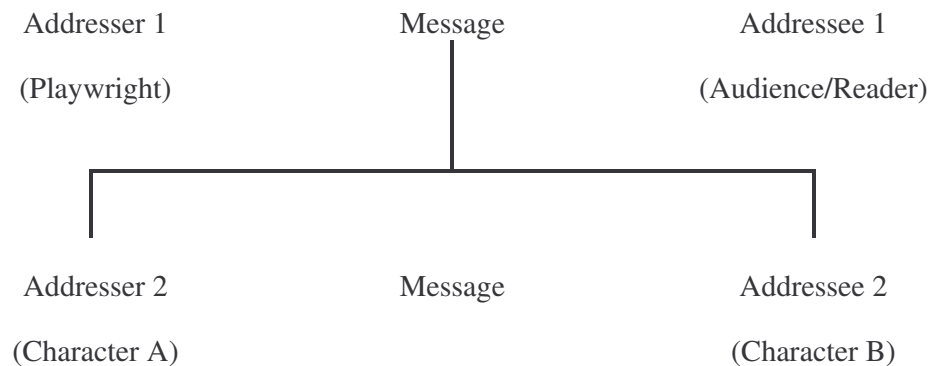
Figure 1. Discourse structure of drama (reproduced from Short, 1996:169)

Short's model shows that the characters (A and B) communicate onstage but, as Short

points out, other offstage characters do not hear the message in their dialogue

(1996:169). The audience, however, is at a higher, more privileged level, able to hear

all the messages exchanged on stage, and indeed those voiced by characters alone,

through soliloquies. Short (1996:169-170) argues that dramatic irony arises through

this dual discourse structure, since the audience often has greater knowledge than the

onstage characters, and he includes a Shakespeare play (*Othello*) amongst his

examples.

Short's concept of discourse structure is important for characterisation. The audience (at the higher level) hears all the explicit cues such as self- and other-presentation present in the dialogue of characters at the onstage level, whilst the onstage characters themselves have access to much less information about one another. Over and above characterisation, the dramatic irony Short demonstrates as being created between the two discourse levels is hugely important to the success of the play as a piece of entertainment. It will therefore be interesting to see whether the key cluster results in the male and female dialogue of the plays show any evidence for contrasts or trends which arise from the dual discourse levels in the plays.

The third important set of ideas to keep in mind in analysing my results encompasses linguistic theories of social interaction and other factors arising at the pragmatic level of dialogic language. Again, space permits only the briefest mention of these, beginning with the importance of existing knowledge, beliefs, assumptions, memories, schemas and ways of inferring meaning which audiences and readers already hold. These factors influence the audience/reader's construal of dramatic characters. This is argued by Short (1996:222-253; 1998) and Culpeper (2001:57-70), plus other scholars including Cooper (1998:56) and Busse (2006:92). Of particular relevance to my study are Culpeper's (2001:47-112) claims that characters are interpreted through the audience/reader's existing knowledge and schemas of:

- the ways real people behave;
- the ways (stereo)typical dramatic character types behave (e.g. villains and heroes); and

- the ways members of particular social groups behave (e.g. people of different

  sex).

Short (1996:227-231) and Culpeper (2001:63-70) argue that much of this kind of knowledge is organised schematically, and they discuss "schema theory" in some detail. As Short (1996:231) neatly puts it, "Schemas are organised representations of background knowledge which readers bring along to texts." The important point is that essentially audiences and readers are predisposed to interpret characters and their behaviour in certain ways, through the influences of their own experiences. Busse (2002:186) argues that Shakespeare would have consciously manipulated the audience's likely assumptions about social behaviour.

In analysing the language that female and male characters tend to use relatively frequently (based on my results), it will therefore be important to remain aware of what possible schematic assumptions and prior knowledge these might be tapping into on the part of the audience or reader. Of course, as Short (1996:231) and Culpeper (2001:68) emphasise, not all audience members and readers of texts have identical schemas since people's experiences vary. Cooper (1998:64) concludes that either of two contrasting interpretations of the character Kate (also known as Katherine/Katharine/Katharina, in *The Taming of the Shrew*) would be valid depending on the assumptions held about conversational implicature.

Scope for different interpretation is one reason Shakespeare's plays remain an intriguing source of comment and criticism, as argued in S.1.1, but it also means that there will always be the possibility of multiple evaluations of characters and their

behaviour, both as individuals and social groups. Furthermore, the existing knowledge and beliefs of an early modern audience/reader would be likely to differ from those of today, since as Cooper (1998:56) observes, cultural and social views have altered considerably since the plays were written. The risk of (mis)interpreting historical pragmatics behaviour through a modern perspective has been pointed out, e.g. by Klein (2002:871); in my study it is further complicated by the stylistic possibilities in dramatic dialogue. Critical literary and socio-cultural/historical commentary, discussed further in S.2.6, can provide useful background information on everyday life at the time Shakespeare's plays were written, but views vary among writers from different critical standpoints and fields of study. Whilst a variety of differing opinions may not lead to any definitive conclusions about early modern social life, it usefully widens the scope of possible schematic assumptions which Shakespeare's early audiences would have held. This helps minimise the risk of misinterpreting language behaviour as stylistically remarkable when in fact it may simply reflect social conventions of the day (and as noted above, there would have been a range of opinions about what was conventional).

Real language behaviour brings me finally to what Culpeper (2001:235) terms "the social dynamics of interaction". Both Culpeper (2001) and Short (1996:195-221) comment on the role of speech act theory (Austin, 1962; Searle, 1969), linguistic politeness and face (Brown and Levinson, 1987) and other aspects of linguistic pragmatics in the interpretation of character behaviour in dramatic interaction. Culpeper (2001, 1998) further argues that the effects of impoliteness in dramatic dialogue are especially important, and Bousfield (2007) demonstrates this in a detailed analysis from Shakespeare's *Henry IV Part I*. Since these aspects of interactional

26

language take place at the pragmatic level, it will be essential that my framework for analysing the functions of the key cluster results (in S.3.4) allows them to emerge. However, I will not expand on linguistic politeness, deference or speech act theory in this study, to allow discussion space for other theories which underpin my research more closely.

The above introduction to the influence of real language use upon character interpretation leads me to a further important issue which needs clarification: exactly what will count as a "stylistic effect" in my data. Culpeper (2001:113-155) discusses theoretical bases put forward for the ways audiences and readers distinguish between character-forming behaviour and behaviour which is simply a response to circumstances. He notes Kelley's (1972) principle of "discounting" character behaviour that clearly arises from the character's circumstances, rather than as a personality trait (see also Culpeper and McIntyre, 2006:779). As stated in Chapter 1, my study takes in evidence for the ways women and men in Shakespeare's plays are characterised as social groups, based on the frequently-occurring language forms they use in the canon, as well as the way individual male and female characters are constructed. Individual character traits are likely to be created through comparatively unusual uses of language, i.e. those which are stylistically remarkable. In my data, however, whilst a single character's response to a particular set of circumstances may not be stylistically interesting on its own, multiple characters doing the same thing will be notable since these may represent a wider stylistic trend associated with one sex or the other. In principle, therefore, I will not discount any result as potentially stylistically interesting, on a collective level if not an individual level. I make the distinction between these types of results clear in my analyses in Chapter 4.

Having outlined the general arguments for using corpus methodology to investigate Shakespeare's plays, and some important ideas which will be relevant to my stylistic analysis, I now explain the theories and research which led me to focus specifically on key word clusters.

2.2   Theories of formulaic language and lexical patterns in texts

The analysis of clusters and other types of RWC has grown out of arguments that language is processed and managed in ready-assembled units or "chunks". Wray (2002) and Sinclair (1991) both advance ideas of language being retrieved and processed in a two-stage process: firstly by drawing on ready-prepared constructions already stored in the speaker's mind, and secondly by constructing new units of language if no suitable pre-existing units are available. In Sinclair's (1991:109-115) terms these are the "idiom principle" and the "open-choice principle", respectively. Hoey (2005) conceptualises a single mental storage process, "lexical priming", whereby words are stored with "primed" information about the ways the speaker knows them to be typically used and encountered (e.g. the usual collocations, semantic connotations, pragmatic uses, styles, domains and genres in which they are found). Crucial aspects of the claims of Hoey, Sinclair and Wray are:

(i)   that these ready-made chunks of language are stored at the lexical level, not in complete grammatical structures; and

(ii)   that the semantic and pragmatic meaning of the chunks depends on the co-occurrence of the words in particular sequences.

28

These ideas underlie a relatively new perspective on linguistic analysis which encompasses the study of RWCs but has much wider research implications. Scott and Thompson describe this as:

> a shift from the clause to the text: from a focus on language as a set of syntactic structures in isolation to a focus on language as a set of functional resources in use. (Scott and Thompson, 2001:1)

Stubbs sums up the situation thus:

> There is no clear consensus, but the current trend, in independent traditions, seems to be towards a model of language in which lexis plays a central role. (Stubbs, 2001:218)

Space permits only the briefest summary of this area, since I must allow space to discuss the nature of word clusters and other studies of RWCs in the following sections. Sinclair's (e.g. 1966, 1991, 2004) work, mentioned in S.1.3, has been a major driving force in this area, which originates with the ideas of J.R. Firth (e.g. 1957). There is now an established research discipline focusing analysis upon electronically-derived statistical relationships between words, such as collocations and colligations. It is these relationships which lie behind the lexical patterns that have been widely investigated in texts by linguists such as Hoey (1991), Biber et al. (1999:993-1024), Stubbs (1996, 2001), Scott and Thompson (2001) and Scott and Tribble (2006). Essentially, these theorists take the view that lexical, grammatical and semantic qualities of words are interdependent, and that language meaning and language function are produced by words **in their surrounding context**, not standing alone. Form and meaning are conceived as working in tandem, not as separate components of language. For example, Biber et al. (1999:995) argue that "[i]n both conversation and academic prose, an important proportion of discourse is made up of recurrent lexical bundles". They state that lexical bundles are any word combinations

which recur in the same sequence with statistical frequency, i.e. not through idiomatic

or grammatical usage, but as a result of "extended collocational association"

(ibid.:991-992). In other words, although lexical bundles form identifiable patterns in

the text which the researcher can analyse, they are actually a product of the mental

language process rather than a structural or topical phenomenon in the text.

This approach to language has resulted in a need for methods of analysis which can

take in lexical, grammatical, semantic and pragmatic levels, since the "functional

resources" mentioned by Scott and Thompson (2001:1), above, are contained in more

than one language level. Many linguists (though not all) have used corpus

methodology to address this need, and again space permits discussion of only those

which particularly inform my study. As stated briefly in S.1.1, I will use word clusters

as the specific linguistic form from which to derive lexical patterns in my corpus,

which I shall then examine for the functions they have in the female (and male)

dialogue. The interpretative process – from form to function to effects in the texts – is

crucial, since it must produce useful information, i.e. information which can explain

what is going on in the plays and therefore successfully answer my research questions

in S.1.4. The practical steps in this process, which I shall follow in my study

(discussed fully in the next chapter), are summarised by Biber, Conrad and Cortes

(2004) with regard to lexical bundles:

> [W]e group together bundles that serve similar functions, based on the
> typical meanings and uses of each bundle. We used concordance
> listings to examine the use of each bundle in its discourse contexts.
> Once the bundles were assigned to groups, we attempted to determine
> the discourse functions associated with each of the groups. Biber,
> Conrad and Cortes (2004:383)

However, it is far from easy to implement such a process successfully. A framework for analysis must be devised, but this cannot be done accurately until the results are known. Culpeper and Kytö (forthcoming) argue that the starting point for a classification system should be the results themselves (what kind of linguistic features they are and what functions they serve). However, Stubbs (2005:6) also points out that to some extent any research begins with known features of language. Some idea of the expected results must be borne in mind when building and annotating the corpus, so as to ensure the right information will be accessible. I discuss these methodological matters, including the classification of cluster functions, in Chapter 3.

Before moving on to discuss relevant corpus studies in detail, it is important to note that formulaicity has influenced other areas of linguistics, e.g. second language acquisition (see e.g. Weinert, 1995 and De Cock, 1999) and conversation (see e.g. Lenk, 1999 and Aijmer, 1996), which I shall not discuss further here. I must also point out that Wray (2002) and Hoey (2005), mentioned above, acknowledge that views such as those of Chomsky do not coincide with their arguments, nor indeed with Sinclair's ideas. However, language processing itself is not the focus of my study; as emphasised in Chapter 1, I seek to analyse Shakespearean dialogue in a new, more empirical way. The weight of research into lexical patterns in texts leads me to believe that this is a robust methodology and, as I continue to show in this chapter, one which is suitable for my study.

As language processing involves larger units than single words, argued above, it therefore makes sense to look at the presence and effects of word **combinations** in texts, not just at single words. Accordingly, in the next section I give further details of

corpus studies involving RWCs, and I expand on the advantages of focusing on "key" clusters in my corpus (see definitions in S.1.5).

2.3   Keywords, key word clusters and other types of recurrent word combinations

Baker (2004:346-347) states that corpus studies of keywords have become a popular way of investigating what is distinctive about language in different genres. Keywords analysis has successfully shown new insights into literary genres (amongst others), for example prose fiction (Stubbs, 2005, mentioned in S.1.2) and indeed Shakespearean dramatic dialogue (Culpeper, 2002; Scott and Tribble, 2006 and Murphy, 2007, discussed in the next section). Stubbs (2005:11) argues that "[t]extual frequency is not the same as salience", claiming that keywords provide a much more accurate picture of what is worthy of analysis than mere word frequency counts. This is clearly demonstrated by Culpeper (2002:16-18), who shows that frequently-occurring words in Shakespeare's *Romeo and Juliet* are often simply words that are in common usage, whereas keywords indicate what is statistically unusual in one character's speech when compared to that of other characters. Key results (words and, by extension, word clusters) are therefore more likely to contribute to effects which are particular to individual characters, or groups of characters belonging to a particular social category (e.g. sex, rank or role). Stubbs includes keywords and key combinations of words in his analysis, arguing that "studying only individual words in the text is inadequate" (2005:13). He gives several reasons for looking at longer units than single words:

- "any text makes references to other texts" (ibid.), i.e. a word might in fact be part of a quotation from another text, but this is masked when it is examined on its own;

- "collocations create connotations" (ibid.:14), i.e. meanings are partly
  determined by the context (coinciding with the arguments of theorists such as
  Sinclair, discussed in S.2.2);

- "words occur in recurrent two-, three-, four- and five-word lexico-grammatical
  patterns" (ibid.:17-18), i.e. word meanings are partly determined by their
  grammatical function, and this depends on their recurrence as part of a
  particular sequence of words (cf. S.2.2).

This suggests that although the keywords studies of Shakespearean dialogue
mentioned above are valuable empirical studies, more can be learned from
investigating key RWCs such as clusters.

Having justified the case for investigating key RWCs, I now mention a selection of
RWC studies in different genres, to demonstrate that it is a well-established approach.
I then discuss studies of RWCs in Shakespeare's plays in more detail, since these are
closely relevant to my research.

Biber et al.'s (1999:993-1024) concept of lexical bundles and Scott's (1999) similar
concept of word clusters, introduced in S.1.3 (and see S.1.5), have had a major
influence on the investigation of the different functions of RWCs. The combined
strength of Biber et al.'s and Scott's research, and others who have followed them in
investigating bundles or clusters, is substantial. For example, Partington and Morley
(2002) examine lexical bundles in the political debate reported in newspapers, and
Culpeper and Kytö (forthcoming) investigate lexical bundles in historical EME texts
(discussed in the next section). Biber, Conrad and Cortes (2004), mentioned in S.2.2,

compare lexical bundles in university textbooks and classroom teaching. Other studies

using similar types of RWC include Altenberg (1998), whose research into the

phraseology of RWCs in the London-Lund Corpus lends further weight to the

argument that language is formulaic (in S.2.2). Moon (1998) finds fixed expressions

and idioms with different functions in a corpus of PDE texts. I discuss the functional

classification systems used by these theorists in S.3.4 in deriving a suitable analysis

framework for my own results.

Turning specifically to literary genres, as mentioned in S.1.3 there are several relevant

studies of prose fiction. Stubbs (2005) includes "phraseology" as well as single

keywords in his investigation of Conrad's *Heart of Darkness*, and Stubbs and Barth

(2003) claim that "chains" of language contribute to different styles of language in

their comparison of fiction, learned writing and "belles" (personal

letters/memoirs/biographies). Starcke (2005, 2006) and Mahlberg (2007a, 2007b)

examine word clusters in the novels of Jane Austen and Charles Dickens, respectively,

adding a new dimension to what has been said in the many critical literary studies of

these popular works. Mahlberg (2007a) shows that combinations of words which

literary critics have argued as being distinctive in Dickens's work can be analysed

more formally and in more detail with the aid of corpus software. In Mahlberg's

(2007a) study, only her comparisons of Dickens's prose with that of other fiction

writers are based on key results. The clusters supporting her main functional

classification and analysis in her 2007a study, and her 2007b study, are simply the

RWCs which occur most frequently in Dickens's prose, so essentially they show what

it is like. In my study I focus only key results, since I want to make an internal

comparison between two datasets from the same body of works, i.e. the male and

34

female dialogue in the Shakespeare canon. I make brief mention of Mahlberg's

(2007a) functional classification system of results in S.3.4, though as I explain, I adopt

another approach which better suits my dramatic dialogue (because it lacks the

substantial narrative content typical of prose fiction).

This brief summary of research shows that RWCs are a useful way into finding out

more about the construction of distinctive text-types, styles and genres. The studies

mentioned span a variety of non-literary and literary genres, supporting Sinclair's

(2004:51, 70) argument that the same methodology can be successfully applied to both

(mentioned in the previous section). The prose fiction studies outlined, plus the studies

of dramatic dialogue in the next section, show that distinctive lexical patterns lie at the

heart of functions contributing to effects such as characterisation. This is the kind of

effect I wish to investigate in my study of Shakespearean dialogue.

2.4   Other studies of keywords and key recurrent word combinations in Shakespeare's plays

Culpeper and Kytö's (forthcoming) study, following Biber et al. (1999) and introduced

briefly in S.2.3, shows that functions of lexical bundles in historical dramatic dialogue

can be classified and analysed successfully. Importantly, Culpeper and Kytö address

the problem of EME spelling variation when using corpus methodology which relies

on orthographic matching to find lexical patterns. This is a major hurdle in the corpus

analysis of texts written before the standardisation of English spelling, which I discuss

fully in S.2.5. Moreover, as mentioned in S.1.3, I had conducted a previous pilot study

which compared the functions of lexical bundles in a sample of plays from each genre

(comedies, tragedies and history plays). The results had been encouraging, showing

some identifiable differences which could be linked to varying stylistic effects such as

dramatic tension and characterisation. The combined evidence from my pilot study and Culpeper and Kytö (forthcoming) indicated that identifying the functions of key word clusters in the male and female dialogue in the plays would be a viable approach.

Before proceeding, however, I evaluated other studies which have taken a similar route into Shakespearean dialogue, to see what I could usefully contribute to this area. The investigation of RWCs in Shakespeare's plays is a largely untouched area. Crystal's (2008:173-175) claim that collocations in Shakespeare's plays are under-researched (see S.1.3) seems to confirm this, since collocational relationships underlie RWCs (argued in S.2.2). I mentioned in S.1.3 that Hota et al. (2006) and Argamon et al. (2007) do include RWCs in their corpus studies of male and female dialogue in Shakespeare's plays but, as noted there, our aims and methodology differ. They do not include any substantial qualitative analysis with examples from the plays to support their claims. For example, Hota et al. (2006) present a small selection of nouns and verbs from their results, e.g. men talk about 'swords' and use the verbs 'avoid', 'fight' and 'wrought' (ibid.:3). From this they claim that "[m]ale characters seem to be aggressive" (ibid.), but they provide no contextual information about the situations in which male characters use these words, or how, or to whom. This is unfortunate, since it can lead to the kind of criticisms of corpus stylistics mentioned by Wynne (2006) and Stubbs (2005:6) (see S.2.1).

Although I located no other studies of RWCs in Shakespearean dialogue, as indicated in S.2.3 Culpeper (2002), Scott and Tribble (2006:59-70) and Murphy (2007) have investigated keywords in the plays. These show the value of key results in pointing to distinguishing style features in different characters' dialogue, including differences

36

between male and female characters. For example, Scott and Tribble (2006:62-63)

discuss the presence of keywords which are exclamations (e.g. "O" and "Ah"), and

argue that in *Romeo and Juliet* "it seems that the female characters are nearly twice as

likely as males to exclaim". This has important implications for the way female

characters are portrayed (possibly as more reactive or excitable, though Scott and

Tribble do not themselves make this suggestion). Culpeper (2002) shows that the

keyword "if" in Juliet's dialogue helps convey her sense of anxiety. His empirically-

based findings provide a useful and interesting comparison to the many claims about

Juliet made by literary critics (e.g. Findlay, 1999:135-138).

This demonstrates the complementary nature of corpus-based and non-corpus-based

approaches to stylistics (argued in Ss.1.2 and 2.1). I anticipate that key cluster results

in my data will show some similar patterns which are likely to be stylistically

significant (for their contribution to character creation in the plays). I also expect

results to be sociolinguistically interesting, in terms of the reasons women might be

represented in drama as behaving in certain ways during this historical period, bearing

in mind Culpeper's (2000:312) argument that Shakespearean dialogue provides some

insight into language and gender issues in early modern England (in S.1.1). Scott and

Tribble (2006:63) caution that many keywords arise simply because of what the play

is concerned with, and that not all are stylistically important. I shall keep this in mind

in considering my key results (in Chapter 4); it is likely that some results which are not

stylistically important will nevertheless be historically interesting.

As pointed out in S.1.1, much has been said about language, sex and gender in Shakespeare's plays and the representation of women. In the next section I put this into some sort of context in relation to my own approach.

2.5   **Approaches to language, sex and gender in literary texts**

I begin with some preliminary references to other works, since language, gender and literary criticism are huge areas which I cannot possibly cover in detail here. As clarified in S.1.1, I follow Talbot's (1998:7) definitions of "sex" as a biological category and "gender" as a social construction. See Sunderland (2006) and Cameron (2005) for wider discussions of language and gender, and Nevalainen (1996) for more on language and gender in EME. For a general background to literary theory and feminist criticism, see e.g. Wolfreys (1999). Culpeper (2000) provides a useful introduction to the variety of literary and linguistic criticism about female characters in Shakespeare's plays in his analysis of Katherina in *The Taming of the Shrew*.

Livia (2003:142) argues that the study of sex and gender in literary texts is typically approached in one of two ways:

    (i)   some scholars attempt to determine male and female **authorial styles**, based on their sex;

    (ii)   others examine the **representation of characters** of different sex, through their dialogue and/or the ways they are described.

As my quantitative data is from one author only, with no other for comparison, my study seems to fall squarely into the second type. Nevertheless, since the social

context of the plays would have contributed to the audience's prior knowledge and

schematic assumptions about the plays, argued in S.2.1 as being relevant to

characterisation of men and women, I must acknowledge that Shakespeare was a male

author writing dialogue for characters of both sexes, all of whom were played by male

actors at that time (argued by e.g. Jardine, 1983:9 and Astington, 2001:109). It

therefore seems sensible and practical to follow Kahn in regarding male authorship as

an inherently interesting aspect of the plays (Kahn, 1981:9, quoted by Jardine,

1983:5), and to bear it in mind as a possible explanation for some of my results.

Interestingly, whilst Livia's research has been into modern French literature, not

historical English drama, she finds no basis for a distinctive "male" and "female" style

of writing. She states that:

> Instead, we have found that there are conventions of masculine and
> feminine style which any sophisticated writer, whether male or female,
> can follow. (Livia, 2003:156)

Livia (ibid.) also finds that authors in her research do not limit themselves to

constructing "conventional" sexual and gender identities, but also create "alternative"

or "oppositional" identities, including characters of no identifiable gender.

Shakespeare does appear to make some creative choices in assigning sex and gender

roles to characters in the plays, as I discuss further in S.3.2.2. For example, the chorus

of Time (*Twelfth Night*) self-identifies himself as male in the dialogue, although as

Boyce (1990:641) points out, Time is "a virtually abstract figure […] distinctly not

human". Shakespeare also re-assigns sex roles in several plays (e.g. *The Merchant of

Venice*) where characters spend part of the play disguised as members of the opposite

sex, and I discuss the methodological implications of this in S.3.2.2.

39

Over the last 25 years there has been considerable commentary on sex and/or gender in Shakespeare's plays from feminist writers, as I said in S.1.1, so I now spend some time evaluating this particular approach. As indicated in S.1.2, space permits the mention of only a few whose work particularly informs my study.

Mills (1995:3-21) provides a comprehensive summary of feminist approaches to language study, usefully discussing these in light of stylistics, sociolinguistics and critical linguistics. Mills takes a feminist approach to stylistics, but warns of the dangers of evaluating texts stylistically without putting them into social and ideological contexts. She claims that too often feminist criticism of texts assumes a general, undefined but agreed idea of the way women should be portrayed, which can then be used to justify existing views rather than explain new evidence (ibid.:14-15). Loomba (1989) highlights the presence of tacit assumptions behind commentary from a single cultural standpoint. She points out that most feminist criticism of EME drama is from a Western cultural perspective, and that other non-Western interpretations of sex and ethnicity in Shakespeare (and other plays) are possible. The arguments of Mills and Loomba show how easily subjectivity can creep into an analysis of the plays, so in my qualitative analysis in Chapter 4 it will be important to question continually the schematic assumptions not only of the early modern audiences for whom Shakespeare's work was intended (mentioned in S.2.1), but also my own.

As discussed further in S.2.6 below, Jardine (1983:ix) initially voiced objections to the traditional feminist approaches to Shakespeare, though she takes a more equivocal view a few years later (in her second edition). She describes her research position as that of an analyst with "special interest" in the female characters in the plays (ibid.:6-

7), claiming that this offers advantages over what she sees as two typical schools of

feminist criticism directed at the plays. Jardine (1983:1-8) describes these as:

- the "aggressive" approach, decrying Shakespeare as sexist; and

- the "non-aggressive" approach, lauding Shakespeare for representing the entire

    spectrum of EME womanhood.

However, Jardine (1983:6) claims that both these feminist views are flawed, since:

(i)   they are based on the premise that the construction of Shakespearean

        characters approximates real people; and

(ii)   because oft-neglected historical evidence suggests neither view to be

        accurate.

Jardine proposes approaching sex and gender in Shakespeare's plays by looking for

traces of "specific cultural issues", i.e. those based on historical evidence, and at the

ways these are reflected in the plays (ibid.).

Findlay's (1999) feminist critical study is one which draws on historical evidence,

from the writings of 16th and 17th century women who formed part of the theatre-going

audience. Findlay (1999:114) echoes many scholars of EME historical texts in

describing the idealised, model female as "chaste, silent and obedient" in the

patriarchal society of the day. Citing the social historian Wrightson (1982), amongst

others, Findlay (1999:127-163) argues that women had much less social power than

men. Women also had fewer rights, for example in owning property and in choosing

41

whom to marry (see also Greenblatt et al., 1997:7-12 on EME class structure and on the lower socio-economic power of women compared to men). However, Findlay and others, e.g. Bryson (1998) in her study of instructional or "didactic" works idealising proper manners and social behaviour in the 16th and 17th centuries, agree that this was a time of changing social codes of conduct. Women were making attempts to break free from the constraints of the past. Additionally, Bach (2007) argues that dominant ideologies of sexuality and sexual relationships have shifted considerably since Shakespeare's plays were written, with platonic male friendships being privileged above male-female romantic/sexual relationships at that time.

At first glance Culpeper's (2000) claim that the interpretation of dramatic characters is inevitably rooted in a schematic knowledge of the way real people behave (in S.2.1) seems to contrast with Jardine's (1983) argument, above, about the extent to which realism in drama is over-interpreted. However, this does not have to be the case. Jardine is commenting on the **evaluative** aspect of interpreting the plays. She is pointing out that Shakespeare's female characters exist as part of a dramatic plot, and there is no evidence that Shakespeare believed their behaviour or their treatment by other characters to be (a) typical or (b) desirable of that among real people. Culpeper, on the other hand, is looking more at the **cognitive** side, as discussed in S.2.1, and explaining that we understand what is going on in the play by filtering it through our knowledge of what real people do. This is a useful distinction to keep in mind when interpreting results, since both arguments are valid.

Jardine's (1983:6-7) position of "special interest" in the female characters, mentioned above, seems a practical approach for a linguistics researcher to take, since it is not

possible to separate the plays from their male author (cf. Kahn, 1981), nor to

completely remove my own Western female cultural viewpoint (cf. Loomba, 1989).

Moreover, this position allows for finding possible stylised characterisations of sex

and gender which are **not** like those of real people (based on Livia, 2003:156 above),

as well as those which **are** (bearing in mind Culpeper's 2001 arguments about

schematic knowledge, discussed in S.2.1).

Adopting a position of special interest also allows me to consider useful background

information on early modern social life from as wide a range of sources as possible

when analysing my results, including selected studies by scholars working in other

non-corpus-based disciplines (e.g. social and cultural historians), as well as literary

critics and feminist writers. I discuss this briefly in the last section in this chapter.

**2.6   Finding a relationship between corpus studies and non-corpus-based
research**

As I argued in S.1.2, non-empirically-based studies run the risk of the criticism of

subjectivity, but the stances taken and interpretations made have to be considered as

part of what people continue to do with Shakespeare's plays. Ultimately, drama is not

owned by linguists or literary critics, but is simply what its consumers make of it, on

whatever basis they choose. As Rose (1988:5) says in her critical literary study of love

and sexual behaviour in Renaissance drama, "Dramatic forms are viewed as ways of

categorizing experience and making it meaningful to people". Arguably,

meaningfulness includes, but is not limited to, empiricism. As indicated briefly in

S.2.5 above, Jardine, writing in the 1989 preface to the second edition of her 1983

publication, qualifies her earlier criticisms of much critical feminist writing on the

plays, allowing that there is a place for "individual subjectivity" (Jardine, 1983:ix).

This does not dilute the value of my empirical approach in this study – as a scholar of corpus stylistics that is my way of making the plays meaningful – but it puts it into a wider context.

My qualitative analysis in Chapter 4 will therefore be mainly informed by the work of other corpus linguists specialising in historical and/or literary texts (e.g. Culpeper and Kytö, forthcoming; Mahlberg, 2007a, 2007b; Busse, 2002). However, I will follow them in mentioning the work of selected literary critics where this seems useful, concentrating on those who incline towards historical evidence as a basis for their claims (e.g. Findlay, 1999; Jardine, 1983), in line with my own empirical approach. Occasionally, I will refer to relevant information from other non-empirical studies of EME drama to help contextualise my results, for instance where this heops determine whether or not results show a stylistic effect. For example, I include Habermann (2003), whose discussions of the way slander was represented in drama make clear the differences in expectations of social behaviour for women and men, and the consequences of breaching these, especially in terms of loss of reputation.

In this chapter I have introduced a great deal of theory spanning several major themes, supported by many existing studies. This has been necessary to explain my research area and my chosen approach adequately, and I shall finish this chapter by drawing all the background threads together in a short summary:

- Despite the existence of contrary views, corpus stylistics is now a respected methodology (argued in S.2.1), and the study of lexical patterns and formulaic aspects of language are both established fields (discussed in S.2.2).

- The weight of theory in favour of analysing key clusters is substantial (presented in S.2.3), since clusters add significantly to what can be learned from single keywords studies. Furthermore, there is sufficient evidence to show that this methodology can be usefully applied to literary texts (argued in S.2.3), including EME dramatic dialogue (S.2.4).

- Despite several keywords analyses which successfully trace stylistic effects such as characterisation in the plays (discussed in S.2.4), there is still a need for study which looks beyond keywords to key word **combinations** (clusters), and which uses empirical results from female (and male) dialogue as the starting point from which to provide a thorough qualitative stylistic analysis. My methodology differs from that in the few existing studies of RWCs in Shakespeare's male and female dialogue, and offers the advantage of much more substantial qualitative analysis.

- Little of the considerable body of commentary on the language and representation of women in Shakespeare's plays will be directly comparable to my results, but non-corpus-based research from selected critical literary studies (discussed in S.2.5) and research by social and cultural historians (S.2.6) will provide useful background information to help explain my results.

In the next chapter, I go on to discuss the methodology of my study.

3.   **BUILDING THE SET: METHODOLOGY AND A FRAMEWORK FOR ANALYSING RESULTS**

My methodology has four successive stages, discussed in Ss. 3.1 to 3.4. I begin with the choice of a corpus from which to obtain my quantitative results (in S.3.1), then I explain the process of adapting it to meet my research needs (in S.3.2). Next, I detail the method of obtaining the quantitative results from my corpus using *WordSmith* (S.3.3), and lastly (in S.3.4) I explain the choice and adaptation of a framework of functional categories through which to classify and analyse the results.

3.1   Choosing a corpus of Shakespeare plays

A number of corpora of Shakespeare's plays exist which have been used in studies by other scholars. For example, B. Busse (2006) and U. Busse (2002) use Spevack's (1968-80) concordances, and Murphy (2007) and Hota et al. (2006) used *The Nameless Shakespeare* (Mueller, 2005). The choice of a corpus depends upon accessibility and suitability for the study, and in this section I explain why I used the corpus constructed by Dr. Mike Scott (Liverpool University)[2]. Having claimed that I use a corpus offering "total representation" of the plays (in S.1.3), it is important to qualify this by acknowledging some debate and variety over the plays which are considered to constitute the "Shakespeare canon" (see e.g. Busse, 2002:193). Scott's corpus is based on the 1916 edition of *The Oxford Shakespeare*, edited by W.J. Craig, which Scott sourced from the *Online Library of Liberty*[3], a source of literary works which may be used for academic purposes. This edition includes the 36 plays in the First Folio (about which see e.g. Greenblatt et al., 1997:65-74), plus *Pericles*.

---

[2] For details of access to Mike Scott's Shakespeare corpus, see http://www.lexically.net/wordsmith (accessed 12.08.2009)
[3] See http://oll.libertyfund.org, hosted by Liberty Fund, Inc. (accessed 08.11.2007)

Scott's corpus contains a directory for each genre of plays (comedies, tragedies and histories). Each directory contains a text file with the contents of each whole play, plus a sub-directory for each play. The sub-directories contain individual text files of the dialogue of each character in the play (or group of characters who speak simultaneously), plus a list of the dramatis personae. The files have already been substantially edited to render them suitable for corpus analysis. Non-dialogic text is annotated so that it is excluded from computations made by the corpus software (discussed further in S.3.2.3). This is imperative in studies like mine which focus on just the spoken dialogue. The annotation also advantageously retains things like stage directions which can later provide useful contextual information in the analysis of results (e.g. the location of characters, their activities whilst speaking, and the co-presence of other characters). Though the structure and annotated contents of Scott's corpus offered clear advantages for my study, annotations which separate the dialogue by sex of speaking characters needed to be added (requiring considerable work), so before embarking upon this I investigated other options.

At the time I did not have access to an electronic copy of Spevack's (1968-80) concordances. Culpeper (2002:14) used an electronic version of Craig's 1914 edition of *The Oxford Shakespeare*, available online[4]. However, this would have required downloading, editing and annotating to bring it to the corpus-ready point of Scott's corpus, so it was clearly not a better option. *The Nameless Shakespeare* (Mueller, 2005) can be accessed via *WordHoard* (Mueller, 2006). *WordHoard* is a computer interface which enables morphological, syntactical, semantic and narrative information

---

[4] See http://www.bartleby.com (accessed 08.11.2007)

to be mined using CLAWS[5]. The EME spelling has also been regularised to overcome

the electronic difficulties of matching variant forms, an important aspect discussed in

S.3.2.5. However, Murphy (2007) and Hota et al. (2006) used selections of text in

their studies, not whole play-texts, and their data manipulation requirements differed

from mine. Murphy (2007) focused his analysis mainly on the soliloquies, i.e. non-

interactional dialogue only, whereas my study includes interactional and non-

interactional dialogue. Therefore, I needed an easy way of identifying onstage

addressees present during the interactional dialogue, plus any relevant stage directions

about their behaviour. Although *WordHoard* is an extremely powerful tool, providing

scope for rapid, broad-brush analysis of the plays by sex of character and genre, it

does not easily provide this kind of information. Hota et al.'s (2006) study includes

interactional dialogue, but is oriented to the "machine learning" process of the

programme *Annotated Text Manager* (*ATMan*), and they do not include the kind of

socio-pragmatic information which will be crucial to my qualitative analysis of the

corpus results (see also Ss. 1.3 and 2.4).

The data in Scott's corpus, although requiring further adaptation, was far easier to

manipulate than *The Nameless Shakespeare* in order to obtain the results I needed, and

offered easier access than *WordHoard* to contextual information necessary for my

subsequent qualitative analysis. Accordingly, I opted to adapt Scott's corpus for my

study, as I explain in the next section.

---

[5] Constituent Likelihood Automated Word-tagging system. See
http://www.comp.lancs.ac.uk/ucrel/claws (accessed 12.08.2009)

48

3.2   Adapting Scott's corpus to suit the needs of this study

Despite Scott's corpus being the most suitable choice as a starting point, the adaptation

process took several months. First, the text files had to be converted for use with the

software available to me, as explained in S.3.2.1. Then, the files were split into male

and female dialogue, presenting some problems which I outline in S.3.2.2. In S.3.2.3 I

compare the various possible ways of annotating the files and justify the route I took.

During the annotation process I identified a few necessary changes to the content of

the plays as it is organised in Scott's text files, detailed in S.3.2.4. In S.3.2.5 I discuss

the application of software which regularised EME spelling variation in the texts.

3.2.1   Converting the files

In order to access the texts in Scott's corpus, I had to convert all the files from

Unicode to ASCII text format, since the software available to me for the project

(*WordSmith* version 3.0) required ASCII. Simply re-saving Scott's plain text files

using the *Microsoft Word*[TM] "Save As" option did not successfully convert them. The

only way to achieve this was to copy the contents of one file at a time, paste them into

new blank documents and save them as different files. (I used the on-screen Windows

Default option and checked the boxes for "insert line breaks" and "allow character

substitution.) Scott's corpus contained well over a thousand files, but fortunately I

was able to use a file conversion programme, *uchack*[6], to carry out the conversion

procedure on multiple files at once. However, not all the files converted successfully,

for reasons which remained unclear despite several attempts, so these had to be done

manually in the above manner.

---

[6] *uchack* was provided to me by Dr. Andrew Hardie (Lancaster University)

3.2.2   Issues arising in separating the male and female dialogue

Once the files were converted, the process of splitting them according to the sex of each character could begin. I argued in S.1.3 that this was more straightforward than identifying characters with a particular social role, but it nevertheless presented some difficulties. As stated in S.2.5, Shakespeare makes some creative choices in assigning sex, probably for dramatic effect (e.g. by assigning sex to a non-biological concept such as time). Wherever possible I used the context of the play to determine the sex of characters. This is often made clear through the pronouns used in the dialogue, or through the social roles and relationships identified in the text (e.g. wife, mother, son, uncle). Lists of dramatis personae were also helpful, as some indicate sex via the character's role (e.g. "brother of [character]"). Where the sex of a character was not determinable through these means, I sought clarification from Crystal and Crystal (2002:514-591) and Boyce (1990), rather than relying on my own (subjective) interpretation.   For example, the sex of the many unnamed servant characters who are identified only by role in the plays (e.g. Servant 1, Servant 2) is not automatically clear unless some clue is given in the way they are addressed by other characters.   Boyce (1990:580-583) indicates that all the unnamed servant characters are male, and I tagged them as such in the absence of other information. This may well have been a matter of convenience if the sex of the servant(s) in the play was immaterial to the plot though, as all the actors would originally have been male (argued in S.2.5).

In some cases the sex of groups of speakers (e.g. "All" in *A Midsummer Night's Dream*, I:ii) could be determined simply by looking at the play-text to see who is actually on stage. In the example, it is a group of male characters (Quince, Snug, Bottom, Flute, Snout and Starveling), so the "All" dialogue was classified as male.

However, if group-speech appeared to be made by characters of both sexes, there were two choices: either include it in both the male and the female dialogue, or have a separate category for mixed-sex dialogue. Although it is not unreasonable to include it as both male and female, because any key clusters which occur belong equally to male and female characters, this is problematic because it also artificially duplicates the frequency of those clusters, skewing the results. Consequently, I created the separate category "both" for mixed-sex dialogue.

More complex was the question of how to categorise dialogue spoken by characters who are disguised as members of the opposite sex (mentioned in S.2.5). The disguise of women as men is a plot device in *As You Like It*, *Twelfth Night*, *Cymbeline* and *The Merchant of Venice*. I considered whether or not to assign a separate category to this dialogue, since it might show interesting results. Findlay (1999:106) argues that whilst masquerading as men, female Shakespearean characters have the chance to express views and emotions that they cannot whilst appearing to be women (due to social constraints). However, the size of the female dialogue data file was not overly large, and separating out the dialogue spoken by cross-dressed characters would have reduced it even further. I therefore decided not to separate it from the main body of female dialogue, but to examine the context of cluster results manually to identify any instances in which females are disguised as males (or vice versa). This was possible since I had annotated the individual character files so that the dialogue could be traced to the speaking character (explained in the next section), justifying the decision to build this into my approach.

As shown above, identifying the sex of characters in the plays was an interpretation of Shakespeare's intentions, rather than an exact science. My data files of "male dialogue" and "female dialogue" can therefore best be described as containing the dialogue spoken by characters who are **represented as belonging to** one biological sex or the other in the plays. The sex of a few characters (with very little dialogue) was unclear, so I classified it separately as "unknown".

Clarifying the sex of speaking characters was the main issue in splitting the male and female dialogue, but there was also a minor issue of deciding how to allocate the content of letters and other documents originated by one character but actually spoken by another in the plays. Example 1 shows a character reading a letter, the text of which is shown in italics.

(1)   Countess: *I have sent you a daughter-in-law: she hath recovered*
                  the king, and undone me. I have wedded her, not bedded
                  her; and sworn to make the 'not' eternal. You shall hear
                  I am ran away: know it before the report come. If there
                  be breadth enough in the world, I will hold a long
                  distance. My duty to you.
                       Your unfortunate son,
                       Bertram.

                  This is not well: rash and unbridled boy,
                  To fly the favours of so good a king!
                  […]

                                    All's Well that Ends Well, III:ii

In example 1, Bertram's letter is read aloud onstage by his mother the Countess, and so forms part of her dialogue. However, it could also be considered part of Bertram's dialogue since he is the originator. However, I decided that Bertram's letter, and other similar cases of characters quoting words originated by others, should form part of the speaker's dialogue, not the originator's dialogue. The reasons for this were twofold:

(i)   Shakespeare purposely presents the reading of the letter as part of the

        Countess's actions and speech on stage, not as being spoken directly by

        Bertram to his mother: it is clearly important to the play that she quotes

        him from the letter;

(ii)   leaving quoted written (or indeed spoken) text as part of the speaker's

        dialogue meant that the question of whether to include the originator as a

        character in the play, despite his or her non-appearance onstage, did not

        then arise.

Having decided how to distinguish the dialogue of different-sex characters, I then annotated the texts in the corpus files, as I now explain.

3.2.3   Annotating the texts to create male and female data files

As indicated briefly in S.3.1, Scott's corpus files were already usefully edited so that any non-dialogic text in the plays is automatically excluded from computations of results made by *WordSmith*. Stage directions, speakers' names, act and scene numbers are all placed between pairs of angled brackets (e.g. <Exeunt>, <Juliet>), a method used in a range of corpus annotation coding languages (e.g. SGML and XML, see McEnery et al., 2006:23-27). I followed this method in annotating the play-texts further.

In order for *WordSmith* to produce the required key cluster results, I needed two separate data files for male and female dialogue so that each could be compared separately with a reference corpus. I also wanted the option to break down the results from these by genre; this later proved to be crucial to making any direct comparisons

53

between male and female results (as discussed in S.3.3.1). I had two clear choices of route in annotating the texts, both of which were substantial undertakings with advantages and drawbacks:

(i)   Option 1 was to construct single-sex data files by joining together individual character text files, using the software programme *SimpleFileJoiner*[7].

(ii)   Option 2 involved extracting the speech of male and female characters from the whole-play files, using software such as *Multi-Lingual Corpus Tools* (Piao et al., 2002; see also McEnery et al., 2006:74-75).

McEnery et al. (2006:75) caution that the quantity and type of metatextual information added to the corpus texts must be designed to produce results which will actually **answer** the researcher's questions, and not be merely interesting, and they also point out that the researcher should consider the cost-benefit of annotation. With this in mind, I considered both options carefully, and I now explain the rationale for eventually choosing the first option.

Joining the files (option 1) seemed quite straightforward: all the male characters' files in one play would be joined into a single file, as would all the female characters' files. Then all the single-sex files for each genre of plays would be joined, resulting in files of all male comedy dialogue and all female comedy dialogue (and the same for histories and tragedies). Finally the single-sex genre files would be joined, resulting in

---

[7] *SimpleFileJoiner*, Version 2.0 (2005), Peretek, Inc. See http://www.peretek.com/sfj.php (accessed 24.06.2008)

two files of aggregated male and female dialogue from all the plays. This process

would certainly produce data files from which I could obtain results that would answer

my research questions in S.1.4. However, there was one drawback: the dialogue in the

individual character files of Scott's corpus was not annotated with the identity of the

speaker, or the title of the play. This would have to be done before joining the

individual character files into larger files of multiple characters, so that the key cluster

results could be traced back to the speaking characters using the *WordSmith* Concord

function (see S.3.3), to obtain the contextual information mentioned in S.3.1 above.

With over a thousand individual text files to annotate this would be no small task, and

a test suggested it would take about 45 minutes per play. However, it could not be

avoided, since the information in the tags would be crucial to explaining patterns of

results. I needed to allow for the possibility that some patterns of key clusters might

prove to be typical of characters in particular social roles (suggested in S.1.3) or

specific situations (e.g. females disguised as males, mentioned in S.3.2.2). However,

this could not be determined without tags showing the speaker's name and the play's

title (to allow for different characters bearing the same name featuring in some plays).

I then considered option 2, extracting male and female dialogue from the whole-play

files. Unlike that in the individual character files, the dialogue in the whole play files

was already annotated with the names of the speakers, although not the sex of the

speakers or the titles of the plays. Annotation in the form of simple tags denoting the

sex of the speaker would of course be required for the extraction software to identify

the relevant data. As with option 1, I felt the play title should be included to ensure

that key clusters were not attributed to the wrong characters in data files containing

dialogue from more than one play. Both tags could have been inserted at the same

time, e.g. <Othello><M> before a speech made by Iago (a male character in *Othello*).

In principle, a tag denoting the genre of the play could also be added at the same time,

as could other tags such as the sex of the addressee. This might have been useful

when analysing the patterns of results, and was not possible with option 1.

Ultimately, either option required the tagging of every speech of more than a thousand

characters in all 37 plays, either on a character-by-character basis in option 1 (the file

joining method), or on a play-by-play basis in option 2 (the dialogue extraction

method). Testing out each way, I found that option 1 was quicker and easier, since

every speech in one individual character file required the insertion of the same two

tags (character name and play title). This could be done in one swift move using the

Microsoft Word™ Find and Replace function, and I estimated it would take about half

an hour per play on average. Example 2 below shows part of the text file of a

character after it was tagged with his name <WILLIAM> and a short form of the play

title <AYLI>. The other XML-type tags showing speech number, act and scene

numbers and the proportion of the play that has elapsed in percentage terms were

already present in Scott's corpus.

```
(2)    <WILLIAM><AYLI><SPEECH 1><ACT 5><SCENE 1><83%>
            Good even, Audrey.
       <WILLIAM><AYLI><SPEECH 2><ACT 5><SCENE 1><83%>
            And good even to you, sir.
       <WILLIAM><AYLI><SPEECH 3><ACT 5><SCENE 1><83%>
            Five-and-twenty, sir.
```

As You Like It, V:i

Option 2 would have required going through each play-text annotating the speeches

one by one, rather than via global replacement, since the dialogue is in interactional

format. Although the play title tag would be the same throughout the play, one of four

speaker sex tags would be required (male <M>, female <F>, unknown <X>, or a

group comprising speakers of both sex <G>). A test indicated that this would take

about two hours per play, more than twice the time estimated for annotation using

option 1.

Option 2 offered the advantage of greater future potential, for example if I later wished

to extend the annotation to include sex of addressees and extract all the dialogue from

characters of one sex to another. However, as the scope of including this in my study

was likely to be limited, given the amount of discussion space available in Chapter 4,

the benefit seemed outweighed the time it would take to implement. It would also be

prudent to investigate other software with text extraction and concordance functions

before embarking on such a large undertaking (e.g. *XAIRA*, see McEnery et al.,

2006:75-76). Accordingly, I went ahead with the first option: annotating the

individual character files then joining them.

The contents of the files were annotated with tags showing play title and speaker name

for each speech turn of dialogue. The total number of files annotated was 1,313, and

Table 1 below shows these broken down by sex and genre.

Table 1. Number of character dialogue files in my corpus, by sex and genre

Sex of characters	Comedies	Tragedies	Histories	All files
Male	358	367	421	1146
Female	79	39	38	156
Both	4	1	2	7
Unknown	1	3	0	4
All	442	410	461	1,313

Table 1 confirms that there are many more male characters than female (1146 to 156), as mentioned in S.3.1, and that the proportion of male characters to female characters is very much higher in all three genres, although less so in comedies. In comedies, the male-to-female character ratio is about 5:1, in histories it is 11:1 and in tragedies 9:1. About twice as many female characters appear in comedies than in histories or tragedies. Although Table 1 shows that comedy, tragedy and history plays contain similar numbers of character files overall, for a corpus-based comparison this is less important than the relative **amounts** of dialogue (numbers of words) in the male and female data files, as I discuss later in this section.

I joined the files of individual characters' dialogue to make two files containing the aggregated male and aggregated female dialogue from all the plays, and six files containing the male and female dialogue from each genre. These are shown in Table 2 further on in this section. I excluded the files of dialogue belonging to groups of both sexes and characters of unknown sex, since these were too small to provide useful results when analysed separately. They are, however, included in the reference corpus of all the dialogue in the plays, which I used to obtain my key results. There was some argument for including the mixed-sex group dialogue in both data files, since it forms part of the dialogue of both male and female characters, but I felt that the dynamics of mixed-sex group dialogue should really be analysed separately from individually-spoken dialogue. As there was very little, it was preferable to concentrate on the main data files of female and male dialogue.

I decided upon a reference corpus of all the dialogue in all the plays after testing different ways of obtaining my quantitative results, discussed further on in S.3.3.1. To

construct this, I simply copied and joined files of individual plays containing all the

dialogue, building three single text files containing all comedy plays, all history plays

and all tragedy plays. I then copied these and joined them into one file containing the

text of all 37 plays. Although I could have joined the single-sex data files to create the

reference corpus, doing so would have resulted in files containing blocks of text

(males, females, groups of both, unknown), and I felt it would be more useful to have

the reference corpus still in the original interactional format of the plays. This holds

potentially useful contextual information, as discussed in S.3.1 above. Table 2 below

shows all the data files and file sizes in number of words.

Table 2. Data files in my corpus of Shakespeare's plays

	Male dialogue	Female dialogue	All dialogue
All plays	669,650	144,451	816,649
Comedy plays	265,810	82,153	348,686
Tragedy plays	196,511	33,878	229,467
History plays	207,328	28,416	238,487
Word counts from *WordSmith Tools* (Scott, 1999)			

Table 2 shows that the data files of male dialogue are all larger than those of female

dialogue, though the amount of female dialogue in comedies is more than twice that in

tragedies or histories (generally following the proportions of male to female characters

shown earlier in Table 1). Since key results are based on statistical comparisons

between two text files, substantially different ratios of male to female dialogue and

numbers of male to female characters are a possible influence on my results. I point

this out in several cases in Chapter 4. The slightly higher proportion of female

dialogue in comedies compared to histories and tragedies may increase the likelihood of some results occurring as key in comedies. This is because the more female dialogue there is, the more likely it is that some forms will eventually recur, and at some point this will reach a statistically significant level when compared to the male dialogue. The disproportionate ratio of male to female dialogue was also a factor in determining the reference corpus (see S.3.3.1).

As indicated at the start of this chapter, during the process of constructing the data files to suit my study I made some minor modifications to the content as it appears in Scott's corpus. I note these in the next section.

3.2.4   Modifications made to the content of files in Scott's corpus

Annotating the dialogue one play at a time, explained in S.3.2.3 above, afforded an opportunity to cross-reference Scott's text files of dramatis personae with those in Crystal and Crystal (2002), highlighting just a few small queries which were relevant to a comparison of male and female characters. There is no listing for Lady Mortimer's character in Scott's files for *Henry IV Part I*, although she appears on Crystal and Crystal's (2002:530) list of characters in the play. Boyce (1990:357-358) states that Lady Mortimer's dialogue is in Welsh not English, and is omitted from some editions of the plays. In Scott's corpus Lady Mortimer's speeches are mentioned only in the stage directions; no dialogue is there. Non-English dialogue would probably occur as key clusters with very low keyness value and frequency, below the thresholds applied in my study (see S.3.3.2) so the exclusion of Lady Mortimer's dialogue was not important.

Scott's text file for the character Luce in *A Comedy of Errors* contained the dialogue of both Luce and Luciana, so I separated this into two distinct character files and amended the speaker tags in my whole play-text file. I looked at the online version of the plays which Scott's corpus is based on, detailed in S.3.1, and found that Luciana's name is abbreviated to "Luc.", a similarity which probably accounts for her amalgamation with Luce. Two similar queries arose. The courtesan's dialogue in *A Comedy of Errors* is located in a character file labelled "courtier" in Scott's corpus, so I amended the file name – a necessary distinction to make in my study, since a courtier would be male and a courtesan female. Lady Macduff's dialogue (*Macbeth*) was all in Lady Macbeth's data file in Scott's corpus, so again I separated this and altered the speech tags in the whole play-text file, so as to be able to identify the speaker of any potentially interesting key cluster results. Similarly, Scott's corpus files combine the characters of the Duchesses of Gloucester and York in *Richard II*, so I separated them.

Scott's *Macbeth* files also contained a single character file for "All", which includes the dialogue of two separate groups in different locations in the play. One group comprises the three (female) witches, and the other is an all-male group, so I separated the dialogue in order to include it with my female and male data, respectively.

Scott's corpus files contained individual character text files for both a doctor and a physician in *King Lear*. Crystal and Crystal (2002:548) list only one character fulfilling this function, as does Boyce (1990:157), and since only one character seems to be listed in Scott's online source, I combined them into one.

Having made these minor changes in the construction of my data files, there was one further important stage in preparing my corpus: addressing EME spelling variation.

3.2.5 Regularising early modern English spelling variation

Crystal (2008:39) claims that although the spelling in Shakespeare's plays had been modernised in EME terms (in that many Middle English spellings had been changed to the spellings which still exist today), multiple variation in the spelling of single words was still common at the time and is found in the plays.

As noted in S.2.4, spelling variations present a problem in a corpus study where the results will be based on frequency counts of words (and word clusters) which are identified by software through the matching of orthographic word-forms (see definitions in S.1.5). Without intervention, the recall rate will be reduced since *WordSmith* will count the same word spelled differently as two separate word-forms, not as multiple occurrences of the same word or cluster. These will be lost, yielding results of lower accuracy, although it is impossible to say whether this would have a substantial effect on the eventual outcomes.

The editors of some collections of Shakespeare's plays have standardised spellings at their own discretion, and Wells and Taylor et al. (1987:155-157) discuss the modernisation of *The Oxford Shakespeare*. The 1916 Craig edition of *The Oxford Shakespeare* (the basis for Scott's corpus, see S.3.1) does appear to be the modernised version, and spelling variation is not mentioned as a problem by Scott and Tribble (2006) in their keywords analysis of Shakespeare's plays.

However, editors of the plays have their own reasons for fixing some spellings whilst

leaving others as variants. Rayson et al. (2007:2) claim that even modernised EME

texts still contain variant forms which could lower the recall rate in corpus analysis.

See Rayson et al. (2005, 2007) for detailed discussions of difficulties presented by

structural and functional differences between EME forms and the PDE English

language forms which corpus software is typically programmed to interpret. The

problem of spelling variation was also noted by Culpeper and Kytö (2005), who

addressed it by applying the variant detecting software *VARD* to their EME texts prior

to the corpus analysis process. *VARD*, now in its second version, was developed as

part of UCREL's[8] semantic tagging programme for EME. Archer et al. (2003:26)

explain its origins; see Rayson et al. (2005, 2007), Baron and Rayson (2008) and

Baron et al. (2009) for discussions of subsequent development and testing. Some of

this has been carried out since the work on spelling variation in my corpus was

completed in 2007, for which I used version 2.1.5.

VARD version 2.1.5 regularises early modern spellings by matching the word-form in

the text to a lexicon of manually-input known early modern word-variants[9]. It applies

rules to help disambiguate words (e.g. "then" could be spelled "than", but "than" was

also a word-form in its own right, so a disambiguating rule distinguishes these through

the context of the word-form). The disambiguating rules increase the precision of the

changes made by *VARD*. As Rayson et al. (2007:4-6) explain, *VARD* deploys rules

derived by phonetic matching and manually-input letter replacement rules, as well as

the list of known variant spellings. Additionally, *VARD* now allows the user to set the

---

[8] the University Centre for Computer Corpus Research on Language, Lancaster. See
http://www.comp.lancs.ac.uk/ucrel (accessed 12.08.2008)
[9] in version 2.2 this has now been replaced by an extended dictionary derived from SCOWL (Spell
Checking Oriented Word Lists); see Baron et al., 2009

"threshold confidence measure" (ibid.:6), i.e. the point between precision and recall at which *VARD* will actually carry out the potential regularisation of spelling it has identified. This offers the corpus linguist the advantage of more control over what is changed by *VARD*. As Rayson et al. (2007:9-10) demonstrate, a lower threshold will yield a larger number of potential variants (higher recall) but with a higher risk of inaccurate changes (lower precision). Essentially, *VARD* replaces fewer variants at higher thresholds because it needs more evidence to do so.

Although I used the first version of *VARD* in my pilot study (see Ss.1.3 and 2.4), I did so mainly based on Culpeper and Kytö's (forthcoming) view that it was likely to improve accuracy, and I was unable to quantify the likely benefits. The second version produces a set of statistics showing the number of variants found and the number actually changed, so experiments can be conducted using different thresholds. Rayson et al.'s (2007:6-11) test of *VARD* with a selection of Shakespeare's comedy plays indicated that it increased the accuracy of results by 3% (from 82% to 85%), using a threshold of 70%. For my present study I conducted tests on several plays, and consistently found the biggest difference occurred between thresholds of 40% and 50%. For example, *VARD* found 662 variants in *Macbeth*, and replaced 352 of these using a 40% threshold, 199 at 50%, 197 at 60% and 183 at 70%. Since few tests using *VARD* had been published at the time, I conferred with the developers about my test results (Baron, personal communication 18.02.2008), and consequently decided to apply a threshold of 50% to my Shakespeare corpus. While this is slightly less cautious than the 70% threshold used by Rayson et al. (2007), it still relies on evidence from more than one source to make the change (the known variants list, together with either phonetic matching or letter replacement). As the *Macbeth* example above

64

reveals, my tests showed very little difference between 50-70% thresholds. The 40%

threshold relies on the known variants list alone, or on a combination of phonetic

matching and letter replacement (without the known variants list).

Applying *VARD* to my Shakespeare corpus involved only about two hours of human

effort, although the larger files of multiple plays required overnight computer

processing time (fortunately, without any user intervention). In total (counting each

play only once) *VARD* located 11,416 variant forms in my corpus and replaced 2,220

of them at the 50% threshold. Rayson et al. (2005) do point out that scholars might

hold different opinions about the evidence upon which spelling modifications are

based. However, this seems inevitable, and applies equally to earlier editorial

processes of spelling modernisation. It is therefore important that the changes can be

viewed by the user, and that the case for making them is documented. The changes to

variant spellings made by *VARD* can be viewed in detail if XML is selected as the

output option when processing the texts. Variant forms found but not altered (due to

insufficient evidence at the selected threshold) are simply annotated with a tag

between angled brackets showing the start and finish, as in example 3a:

```
(3a)  Queen      And Edward, my poor son, at
      Margaret:  <variant>Tewksbury</variant>.
```

<div align="right">

Richard III, I:iii
</div>

Variant forms which **are** altered are similarly annotated with tags at the start and

finish, but the preceding tag includes information indicating the evidence on which

VARD has based the alteration, as shown in example 3b:

(3b) Queen
      Margaret:
```
Why, so I did; but <replaced
variant="look&apos;d" foundBy="pls"
replacementType="pr"
ed="1">looked</replaced> for no
reply.
```

Richard III, I:iii

As noted above, Scott's corpus is derived from an edition of the plays in which

spellings had already undergone modernisation by the editors. It is therefore possible

that even without the increase in recall rate afforded by *VARD*, sufficient numbers of

matching orthographic forms would have been recalled in order to identify the most

significant key clusters in the text. However, Scott and Tribble's (2006) keywords

results required single orthographic forms to be matched. My key cluster results

require the matching of three (or more) consecutive orthographic forms for each result.

As some of my single-sex datasets are likely to be quite small anyway, it was

especially important to maximise the recall rate in order to obtain as many results as

possible. I also wanted to take up the opportunity of using a fairly new tool which

offers a substantial advantage to corpus linguists studying EME texts.

In this study, quotations from the plays contain spellings regularised by *VARD*.

Regularisation of the spelling completed the preparation of my corpus, enabling me to

start the quantitative analysis.

3.3   **Obtaining the key word cluster results using *WordSmith Tools***

I conducted some tests in preparation for the quantitative analysis using *WordSmith* in

order to obtain the best possible quantitative results, in numerical frequency and

statistical strength. In this section, first I discuss the choice of reference corpus (in

66

S.3.3.1). Next I explain tests determining the settings of minimum frequency and p

value (S.3.3.2), then cluster length (S.3.3.3).

In testing and obtaining the key cluster results, I used *WordSmith's* WordList and

KeyWords functions. I also used the Concord function to examine my results, and to

classify them functionally (further on in S.3.4). Space limitations preclude an

explanation of *WordSmith* functions; for this see Scott's (1999) *WordSmith* Help

Menu, and see also Scott and Tribble's (2006:30-31) examples of wordlist results.

3.3.1   The reference corpus

The choice of reference corpus influences the type of results obtained. Scott and

Tribble (2006:59) used all the dialogue in all the plays as a reference corpus in their

keywords study, whilst pointing out (ibid.:64) that Culpeper (2002) took a different

approach in his analysis of six characters from a single Shakespeare play. Culpeper

(2002) compared each character's dialogue with a reference corpus of all the **other**

characters' dialogue. Scott and Tribble (2006:64) argue that this method produces

results which "home in on individual difference", affording useful contrasts to the

"common core" of results which typically arises from larger sets of results and

reference corpora. I evaluated the suitability of both approaches for my data, starting

with the dialogue from all the plays as a reference corpus.

As Table 2 (in S.3.2.3) showed, the data file of all play-texts in my corpus is about

five and a half times the size of the aggregated female data file, but not even twice the

size of the aggregated male data file since there is so much more male dialogue than

female dialogue in the plays. Scott's (2006) findings suggest that the size of reference

67

corpora can be smaller than Berber Sardinha's (2004) recommendation of five times the size of the text under consideration without producing poor results. This does depend on the content of the corpus, however. When I tested the male and female data files from all the plays with a reference corpus of all dialogue from all plays, only the female data file produced any key clusters of three words or more (see my discussion of cluster length in S.3.3.3 below). The male dialogue constitutes too much of the overall dialogue to contain any clusters which differ significantly.

Hota et al. (2006), mentioned in Ss. 2.4 and 3.1, used texts from equal numbers of male and female characters from each play with amounts of dialogue above a certain threshold, which solves the problem of the unequal male/female data file sizes. Inevitably, though, this means excluding a great many characters, especially male characters, since they substantially outnumber the female characters as shown in Table 1, (S.3.2.3). I rejected this option because it would have countered my aim of basing the study on all the dialogue in the plays, stated in S.2.1, and also because the dialogue of characters who say relatively little might nevertheless feature linguistic forms that are part of a bigger trend among women (or men) in the plays.

Culpeper's (2002) method of using a reference corpus of all other dialogue apart from the text(s) under consideration effectively meant comparing the male dialogue against the female dialogue and vice versa in my study. A test produced insufficient results for the male dialogue, and similar results for the female dialogue as when using the whole corpus (because the male dialogue is the main constituent of the corpus). When the single-sex data files were broken down by genre, the results using all the **other** dialogue compared to using all the plays as a reference corpus were again very similar.

The only difference was that the keyness values were higher across all the results when the text(s) under consideration were excluded, i.e. when all other dialogue was used as the reference corpus. In effect, the "common core" of results also noted by Scott and Tribble (2006:64), mentioned above, emerged regardless of whether or not the data under consideration was included in the reference corpus. I rejected the temptation to use all other dialogue as a reference corpus simply because it produced results of greater statistical significance (higher keyness value) across the board. That in itself would not make the same set of key cluster results any more useful than if they were obtained using a reference corpus of all plays; as Scott and Tribble (2006:64) point out, they will still arise as key for the same reasons, some of which may be more stylistically interesting than others.

I could see the potential benefit of using all other dialogue as a reference corpus when doing close-up analysis of a few characters, as Culpeper (2002) did, but felt it was less suitable for my study involving many characters and data files of uneven sizes. Therefore, I opted to follow Scott and Tribble (2006) and use all the plays as a reference corpus. My experiments suggested that it would be better to concentrate not on individual keyness values of the results, but on the contents of the clusters themselves with particular scrutiny of any showing a high keyness value **relative to the other results**.

This did of course mean accepting the regrettable lack of a set of results from the aggregated male dialogue as a limitation of the methodology. However, as stated in S.3.2.3, I also prepared data files of male and female dialogue broken down by genre (tragedy, comedy and history), and when these were tested with a reference corpus of

all the plays, satisfactory results for both sexes were produced. As shown in my

analyses in Chapter 4, the male results from the genre breakdown provided plenty of

context in which to discuss the female results (which is the orientation of my study, as

stated in S.1.1). Furthermore, the contents of the aggregated male data file, whilst not

producing any statistical results, could usefully be searched to determine whether

some clusters in the female data are also used by male characters.

A potential drawback in breaking down the dialogue of the plays by genre, or on some

other pre-determined basis (e.g. plays written earlier or later), is that it anticipates

differences in those categories at the start. However, the results from the aggregated

female dialogue provided an overview of key clusters in the whole corpus of plays,

from which I could identify patterns and concentrations in results which superseded

genre, or which seemed to group according to other variables (e.g. social role). These

are discussed in Chapter 4. It would have been possible, and worthwhile (as I suggest

in S.5.3), to compare the data files of male and female dialogue in Shakespeare's plays

with a reference corpus of other contemporary texts, but in this study I would not have

had sufficient discussion space to do justice to the results.

Once the decision was made about the reference corpus, I was then able to compile

wordlists in *WordSmith* and carry out tests to determine the optimum settings for

obtaining the key cluster results.

3.3.2   Statistical tests, p value and minimum frequency

The choice of statistical test, p value (the probability threshold for results occurring

due to chance; see the *WordSmith* Help Menu and Rayson et al., 2004) and minimum

frequency contribute to the quality of key results. At just over 800,000 words my

corpus is fairly small, compared to, e.g., the 1.3 million-word *A Corpus of English*

Dialogues, 1560-1760[10] ("*CED*"), despite containing all the available dialogue. The

breakdowns of the plays by sex and genre further reduced the sizes of the datasets (see

Table 2, S.3.2.3), so I anticipated fairly small numbers of results from each. Rayson et

al. (2004:928) argue that Dunning's (1993) log-likelihood statistical test is more

reliable than the chi-square test for expected frequencies below 5. As I expected low

frequencies from some of my smaller data files, I selected this option in *WordSmith*. I

then ran a series of tests using different permutations of minimum frequency of

occurrence and p value, and I considered the experiences of other corpus linguists.

Although Scott and Tribble (2006) and Culpeper (2002) both use the log-likelihood

test, they apply widely differing p values. Scott and Tribble (2006) use the lowest

possible p value (p = 0.000001) whereas Culpeper (2002) uses a p value of 0.05,

which he points out is conventional in social science studies. However, Rayson et

al.'s (2004) experiments with statistical tests in corpora indicate that a p value of

0.0001 produces more reliable results where expected frequencies are below 8.

Unfortunately though, my tests showed that I could not balance the desirable p value

of 0.0001 with a suitable minimum frequency, as I now explain.

I followed Culpeper (2002) in setting a minimum frequency of 5, to eliminate single

occurrences of key clusters, and those which would be more likely to be localised to

one or two plays due to the topic. To further minimise the problem of local or topical

results, I also adopted Culpeper and Kytö's (forthcoming) principle of excluding

results which occurred in fewer than three different plays, and for my data from all

---

[10] *A Corpus of English Dialogues, 1560-1760*, compiled by Jonathan Culpeper (Lancaster University)
and Merja Kytö (Uppsala University) (see Kytö and Walker, 2006)

plays I imposed a further criterion of occurrence in more than one genre. This information was obtained using the *WordSmith* Concord function, which pinpoints where each result occurs. The reason for the minimum genre criterion was that although the history plays can be considered to stand separately, they also form two "tetralogies", i.e. sets of plays containing many of the same characters and continuation of topics (see e.g. Kastan, 2001:174-5). This produced a satisfactory number of results distributed across at least three plays, but only when combined with a p value of 0.05. Lower p values produced results with lower frequency (below 5), and on examination these proved to be local to one or two plays, i.e. they arise from what the play is about, not from a wider function. These would be less useful in tracing characterisation effects, and therefore to answering my research questions in S.1.4.

Faced with a trade-off between p value and minimum frequency, I opted for the lower strength of statistical significance (p value = 0.05) and a minimum frequency of 5, since this produced a wider distribution of results which appeared to have non-localised functions in the plays. I felt this was an appropriate decision, bearing in mind:

(i)   Scott and Tribble's (2006:63) caution about choosing the most interesting results to analyse (in S.2.4); and

(ii)   Rayson et al.'s (2004:933) emphasis that practical factors have to be considered in corpus analysis, as well as statistical significance.

Although a p value of 0.05 has a lower cut-off critical value of 3.84, I used only those with a critical value of 6.0 and above, since this afforded a sufficient number of the strongest (most key) results to analyse.

3.3.3   Cluster length

With the statistical test, p value and minimum frequency determined, I used the *WordSmith* KeyWords function to test the presence of key clusters of different lengths in my data. A manageable number of 3-word clusters occurred in nearly all the datasets and sub-sets. There were very few clusters of 4 words or longer – insufficient to make for useful analysis on their own – and 2-word clusters did not provide sufficient details of the pragmatic phenomena I wanted to include (e.g. speech acts). This was also found by Culpeper and Kytö (forthcoming), and I follow their decision to focus on 3-word clusters. I obtained key cluster results for the aggregated female dialogue, and breakdowns by sex and genre of play (using the data files in Table 2, S.3.2.3). My tests had indicated that some 3-word cluster results were part of longer sequences (i.e. they overlap). Accordingly, I also obtained key 4-word cluster results from all data files in which these occurred, in order to say as much as possible about what is going on in the dialogue in Chapter 4. There were no key clusters longer than 4 words in my data. During the course of my analysis, I later obtained key 2-word clusters in order to provide supplementary information relevant to the key 3-word cluster 'I pray you' (see S.4.3.1).

Once I had obtained the quantitative results, which are tabled further on in S.4.1, I grouped and classified them according to their different functions in the play-texts. I now explain how this was done.

3.4   A framework for analysing the quantitative results

I had the option of devising a completely new set of functional categories to suit my

data, or adapting a framework used in other studies. As emphasised in S.2.2, this was

an important but difficult process. The method of grouping my results had to provide

a pathway to locating potential differences in the ways in which male and female

characters use language, and in what circumstances, in order to answer my research

questions in S.1.4. As I wished to spend as much time as possible on the qualitative

analysis of my data, adapting an existing framework seemed a practical choice.

However, I emphatically wished to avoid manipulating the results into a pre-selected

framework, since this would severely limit the possibilities for original analysis. I

therefore obtained and investigated the cluster results thoroughly to see what they

contained, then I evaluated existing functional classification systems to see if one

could be suitably adapted to my data. I discuss existing systems in S.3.4.1, explaining

the rationale for my eventual choice to adapt one. In S.3.4.2 I detail the adaptations I

made and the resulting functional categories into which the key cluster results were

placed.

3.4.1   Evaluating my results in view of existing systems of functional classification

I assessed the co-text and context of my key word cluster results using the *WordSmith*

Concord option, which enabled me to see more clearly what function they perform in

the dialogue of the plays. Without the information provided by the wider surrounding

text it would not have been possible to discern the full contribution of the clusters to

the meaning of the dialogue in my data. For example, the cluster 'let me go' mainly

functions as part of directive speech acts, but this can only be confirmed through

74

examining the preceding text and noting the absence of a first or third person pronoun

and the presence of the implied second person "you", as in example 4:

  (4)   Julia [to   Then <u>let me go</u> and hinder not my course:
            Lucetta]:   *Two Gentlemen of Verona*, II:vii

Examining the cluster results in context showed that they had an assortment of lexical,

grammatical, semantic, pragmatic and discoursal functions in the play-texts, as I had

cautiously anticipated from my pilot study (see S.3.1). For example, the function of

the cluster 'I will not' (analysed in detail in S.4.3.2) is sometimes simply to convey the

grammatical future tense, i.e. intention. More often, though, it serves to inform the

audience of the speaker's wider attitudes and desires, i.e. volition. Both possible uses

in ShE are confirmed by Hope (2003:145). The cluster 'I pray you' (analysed in

S.4.3.1) clearly has a function of politeness. Most clusters had more than one function

depending on the contexts in which they were located, and I discuss the way this was

handled later in this section.

Having established the nature of my results and their wide-ranging functions, I was

better placed to evaluate the suitability of classification systems devised by other

scholars. In S.2.3 I stated that these included Altenberg (1998), Moon (1998), Biber,

Conrad and Cortes (2004) and Culpeper and Kytö (forthcoming). Altenberg's (1998)

study is of phraseology, and his categories are more oriented towards the grammatical

functions of his RWC results. My results did not lend themselves readily to

grammatical classification, since many could at best be put into a single category

described as incomplete clauses (e.g. 'leave me to', 'but now I'). Grammatical

classification would have offered a limited perspective on why the clusters are important in the text.

Functional classification, on the other hand, enabled me to take my analysis further, showing what the clusters are doing over and above performing their grammatical role in the plays. As indicated in S.2.3, Mahlberg (2007a, 2007b) takes a functional approach to her cluster results in Dickens's novels, but she concentrates on very localised functions in a "bottom-up fashion" (2007a:13). This successfully enables her to show how Dickens defined his characters in quite distinctive ways. Mahlberg's (2007a:18) categories would not work as well in drama as in prose fiction, however, because they are necessarily oriented more to narrative than to interactional dialogue. Mahlberg includes functions which concern the way the author describes the fictitious world, whereas I need functions which concern the pragmatic aspects of social interaction, argued by Culpeper (2001:57-70) as being inherent in dramatic characterisation (see S.2.1). Mahlberg's "Labels" function encompasses characterisation, and her "Speech" function includes character interaction, but these are quite general categories which would not take my pragmatics analysis very far. All my cluster results function as speech, and characterisation is a potential function of any of them.

Busse (2006), Culpeper and Kytö (forthcoming) and Moon (1998) have all used functional category systems derived from Halliday's (e.g. 1994:179) three "metafunctions" of language:

- "interpersonal" functions (establishing speaker-addressee relations);

- "textual" functions (organising the message);

- "ideational" functions (making reference to something).

Busse (2006:137) analyses her data on vocatives in Shakespeare's plays using Halliday's (e.g. 1994) concepts of 'Systemic Functional Grammar' and 'grammatical metaphor', but as with Mahlberg's (2007a, 2007b) functional categories discussed above, Busse's are very localised and would not easily facilitate description of the pragmatic and discoursal functions evident in my key word clusters.

In her study of fixed expressions and idioms ("FEIs"), Moon (1998:217-218) classifies her data into five sub-categories ("informational", "evaluative", "situational", "modalizing" and "organizational"), derived from the sub-categories of Halliday's three metafunctions. However, Moon's results were complete semantic units which were manually identified in the texts, whereas mine are frequency-based units identified automatically (see S.2.2), and as such they are fragments which do not fit neatly into categories at any one language level.

Culpeper and Kytö (forthcoming) also take Halliday's ideational, textual and interactional metafunctions as a starting point, relating these to the "macro-categories" proposed by Biber, Conrad and Cortes (2003) and incorporating their sub-categories. These are more numerous than Moon's sub-categories, and they offer more specific classifications (e.g. the possibility of classifying deictic references to time or place). This greater choice of sub-categories was attractive since it enabled me to classify my clusters with greater accuracy, affording a greater likelihood of pinpointing interesting patterns in the Shakespearean dialogue. However, my results were not as numerous as

those of Biber, Conrad and Cortes or Culpeper and Kytö, and I envisaged ending up

with one or two results in every sub-category and no discernible patterns at all in the

play-texts. Accordingly, I decided to adapt Culpeper and Kytö's (forthcoming)

classification framework by combining some sub-categories where appropriate to the

number and type of results I had (explained fully in S.3.4.2 below). Adapting

Culpeper and Kytö's (forthcoming) framework had in fact proved suitable in my

earlier pilot study of lexical bundles, but it was important to investigate the results in

this study on their own merits, and not simply re-apply the same functional categories

used previously. Indeed, in the present study I had to accommodate a few clusters

with a narrative function, which had not occurred in those in the pilot study (although

as anticipated above, this was still not a major language function in the present data).

Adapting an existing classification framework was a practical option given the time

available, although not one I would have applied at the expense of interpreting the full

range of functions which seemed to be present in my data. In deciding that the

adaptation of an existing set of categories offered a sufficiently objective place to

begin analysing my results, I am effectively forging a compromise between two

arguments mentioned in S.2.2:

(i)   Culpeper and Kytö's (forthcoming) view that the functional categories
         must be based on what is in the data; and

(ii)   Stubbs's (2005:6) claim that starting with some known aspects of language
         would be inevitable.

78

Whilst the language functions I identify in my results are not new linguistic description, demonstrating the ones which are prominent in the dialogue of male and female Shakespearean characters, and my discussions and conclusions about these (in Chapters 4 and 5), will be original. In the next section I discuss the adaptation of Culpeper and Kytö's (forthcoming) categories to suit my data, and some issues that arose during this process.

3.4.2 Functional classification of my results

A few clusters in my data coincided with lexical bundles in Culpeper and Kytö's (forthcoming) data, but since these occurred in different contexts to my results I did not assume them to have the same functions in my texts as in theirs. I classified the functions of all the key clusters in my study according to the evidence from the *WordSmith* concordance data mentioned in S.3.4.1 above, assisted by the information on ShE in Blake (2002), Hope (2003), Onions (1982) and Crystal and Crystal (2002).

Culpeper and Kytö's (forthcoming) data is from the *CED* (see S.3.3.2), which contains text types from several genres. In contrast, my corpus contains texts by a single author from one genre, so it is fairly unsurprising that some functions identified in their data do not occur in mine. For example, conveying factual information, a function of some results in their courtroom trial data, does not occur in my dramatic dialogue data. In adapting their system, I excluded this and others of Culpeper and Kytö's functional categories for which my Shakespearean data contained no results. I combined some of their sub-categories for which my data showed very few results, for example Textual: Narrative-related functions. However, for clarity I retained their three distinct sub-categories of Textual function, although my data contains only one type of cluster in

each. Whereas Culpeper and Kytö identified discourse acts functioning as answers in their data, as well as questions, I could not satisfactorily do so in mine.   Their Answers category featured elicited responses to questions in courtroom trial data, and it seemed to me that this constituted a discrete function associated with this particular text type. In comparison, the responses to questions in my dramatic dialogue were more accurately described as replies. However, "Replies" could not successfully be applied as a functional category since it would have obscured the other functions of replies to questions (e.g. providing information to the audience, see S.4.4.1), which were more directly useful in answering my research questions.

Table 3 on the next page shows the complete set of functional categories which I applied to my data, with my own explanatory notes in italics. There are three levels of function. The main functions (Interpersonal, Textual and Ideational) each have sub-functions (shown in the left hand columns), and each of these has one or more sub-sub-functions (shown in the right-hand columns).

Table 3. Functional classifications of key clusters in this study

INTERPERSONAL FUNCTIONS	
Speech act-related: *(carrying some force or purpose beyond the words themselves)*	Directive *(having the purpose of telling or asking another character to do something, however strongly or weakly; includes e.g. requests, commands)* Assertion Expressive Sincerity device Vocative Thanking
Modalizing: *(showing degrees of a character's attitude or opinion)*	Volition *(a character's desire for an outcome, or negated: a desire for something <u>not</u> to happen)* Intention *(a character's plans to do/not do something or to achieve/avoid an outcome)* Ability Prediction Downtowners/amplifiers
TEXTUAL FUNCTIONS	
Discoursal: *(a communicative act at a higher discourse level than a speech act)*	Question
Narrative-related:	Reporting/reported clause fragments *(includes reports of speech, thought, writing and events or actions)*
Organisational: *(arrangement of the message)*	Informational elaboration
IDEATIONAL FUNCTIONS	
Topical:	People Informational specificity States *(including physical or attitudinal states, literal and metaphorical states)*
Circumstantial:	Place Directional
Categories adapted from Culpeper and Kytö (forthcoming)	

Culpeper and Kytö (forthcoming) point out some inherent problems in classifying the functions of RWCs, including:

(i)   the fact that language **in use** operates on a number of levels and performs several functions at the same time;

(ii)   the question of how much the surrounding context should influence the classification of the contents of the cluster itself; and

(iii)   the fact that RWCs can have different functions in different contexts.

These issues also arose and were addressed in the categorisation of my data, as I now explain.

The key clusters in my data often contributed to several discourse functions embedded in layers in the text. For example, a cluster which is a fragment of a question performs a lexical or grammatical function, within a speech act, which is part of an overall discourse act. These discourse functions exist below the ultimate level of dramatic function, i.e. the communication between playwright and audience (see the discussion of discourse levels in drama, in S.2.1), but they help illuminate how the dramatic effects such as characterisation work in the plays, as I demonstrate in Chapter 4. Culpeper and Kytö (forthcoming) adopted the principle of classifying lexical bundles according to the highest discourse level at which the bundles clearly had a function, the top tier being discourse acts, the next tier speech acts, and the lowest tier lexical or grammatical items. This seemed a good approach in principle, since the order prioritises functions which would appear most likely to lead into the most interesting effects in the dialogue, which would in turn better serve my research aims. When

applied in practice, though, it was difficult to maintain a consistent approach. This was because in using the context and co-text (from the *WordSmith* concordance data) to help identify the function of the clusters, inevitably they exert some influence over the categorisation decision. It was sometimes hard to judge the highest level of discourse at which the cluster itself (rather than its co-text) could be said to be having an effect. The strength of influence of the surrounding co-text on the function of the clusters themselves actually seemed to vary from instance to instance. Verbs immediately preceding key clusters seemed to have a strong effect on the function. For example, 'what I can' is often preceded by the verb "do". Once this is known, it is clear that the cluster helps convey the speaker's ability to act in some way. However, without the preceding "do" verb, the cluster could have a wider referential meaning. I argued in S.3.4.1 above that it is essential to consider the co-text and context of the clusters in order to determine their actual function.

The above example demonstrates that although the key clusters are the statistically significant results, their effect upon the audience (via the dialogue) does not occur in isolation. Nevertheless, the functional classifications must be primarily rooted in the quantitative results, not the words surrounding them, in order to maintain the empirical basis argued as an advantage of this study (e.g. in S.2.1). In cases where the co-text seemed to exert a major influence over the categorisation decision, it helped to consider what would be the effect of removing the key cluster altogether, to get a clearer picture of its true role. I have applied Culpeper and Kytö's (forthcoming) principle of categorising at the highest discourse level as consistently as possible, given the realities of my data, but I acknowledge that other interpretations might be possible.

As indicated above, key clusters often had different functions in different situational contexts. For example, 'my lord and' was used sometimes as a term of address and sometimes as a term of reference. Once I had worked out the complete set of categories in Table 3 above, I adopted Culpeper and Kytö's (forthcoming) method of dealing with multiple functions. I assigned a functional category to a key cluster only if it applied to at least 50% of occurrences. I placed those for which no sole function amounted to half the occurrences in a "mixed" category, and the few for which the function was really unclear in an "unclassified" category. It would also have been possible to identify a main function for each cluster then exclude occurrences with other functions using the *WordSmith* Zap facility, then recalculate the key results. However, excluding data for quantitative convenience would also have excluded potentially interesting qualitative information. Since my general principles in this study lean more towards working with any imbalances which truly reflect the nature of the dialogue in the plays (e.g. in S.3.3.1), I retained instances with different, less frequent functions.

The functions of all the clusters in the female dialogue overall, and nearly all those in the breakdown by sex and genre, were successfully classified using the category system detailed above. Those in the sex/genre breakdown with mixed or unclassified functions are presented in Table 8, with the other results for completeness (see S.4.1), but since they do not help trace any stylistic effects I do not analyse them further. Clusters including "house of" and "realm of" were difficult to classify. They are clearly Topical, but sub-categorising them further proved problematic. Part of their function is to identify a group of people (a family or a nation) but the meaning of

house and *realm* extends beyond people to institutions and kingdoms. Therefore, it did not seem appropriate to include clusters such as 'the house of Lancaster' with other clusters in the Topical: People sub-category, and no other category seemed appropriate. The fact that the vast majority of my results were satisfactorily classified indicates that the chosen framework for analysis was suitable for the project.

So few corpus stylistics studies of Shakespeare's plays currently exist (argued in S.2.4) that my methodological experiences are in themselves a contribution to the research area, in addition to the outcomes of my analyses in Chapter 4. Therefore, they have been worth charting in detail, and I give a brief summary below:

- Scott's corpus was the most suitable one for the study, because of its accessibility, existing annotation and data manipulation possibilities (argued in S.3.1). Nevertheless, adapting it to investigate male and female dialogue took considerable time and effort and required careful planning to ensure the desired results would be available (explained in Ss. 3.2 to 3.4).

- The EME spelling regulariser *VARD* can be applied to Shakespearean texts to improve the accuracy of results when using corpus software which relies on orthographic matching, even when a "modernised" edition of the plays is used (shown in S.3.2.5).

- Limitations of the methodology include the fact that determining the sex of the characters and applying the functional categories are somewhat interpretive processes (explained in Ss. 3.2.2 and 3.2.4, respectively). The male and female data files are of unequal sizes (S.3.2.3), but the consequent absence of results from the aggregated male data file will not significantly hamper the aims of the

85

study. Results from the aggregated female dialogue can be obtained, as well as comparative results between male and female dialogue when this is broken down by genre (explained in S.3.3.1).

- Assessing the contents of my results in view of existing frameworks of functional categories showed that Culpeper and Kytö's (forthcoming) system was the most suitable for adaptation (argued in S.3.4.1), since it accommodates similar types of results to those I obtained. This practical decision will maximise the resources of time and space available for original analyses and discussion.

In this chapter I have explained that my methodology combines some previously tried-and-tested processes (adapting an existing corpus of the plays, and using a simple annotation system) with some newer aspects (*VARD* and the functional categorisation system of Culpeper and Kytö, forthcoming). Having discussed thoroughly the reasons for choosing this methodology, and the limitations and problems faced, I will now go on to present and analyse my results.

4.   RECOUNTING THE TALE: ANALYSIS AND DISCUSSION OF RESULTS

This chapter begins with the complete set of results from the quantitative analysis, preceded by some important notes on their interpretation (in S.4.1). In S.4.2 I briefly comment on the results in general, and explain my choice of those to analyse in depth (since space does not permit discussion of them all). My analyses then follow in Ss. 4.3 to 4.5. In S.4.6 I discuss in detail what the results of the study show.

4.1   The key cluster results

As explained fully in S.3.3.3, my data includes 3- and 4-word key cluster results from the female dialogue overall in the plays (the aggregated female dataset), and from the breakdowns of male and female dialogue by genre of play. These are generated independently, i.e. the breakdowns are not derived from the overall results, so some results occur in the sex/genre breakdown but not in the aggregated dataset, and vice versa. I discuss both sets of results in my analyses, since it is important not to miss any interesting features which supersede genre, whilst nevertheless looking at the influence of the genre when seeking to explain the results (see also S.3.3.1).

To conserve space, the tables of results exclude functional categories for which no results occur in the dataset (for the complete set of functional categories see Table 3, S.3.4.2). As stated in S.3.3.3, the few 4-word clusters in the data often overlap with one or more 3-word clusters. Since these usually have coinciding functions, it is convenient to table them together in this section, and subsequently to discuss them together. The frequency of occurrence of each cluster is shown next to it in brackets. Nearly all the results are positively key; negatively key results are indicated by a

minus sign next to the frequency of occurrence, e.g. (13-). See S.1.5 for definitions of

positive and negative keyness.

Table 4 shows the results from the female dialogue overall (aggregated) in the plays.

This is followed by a breakdown by sex and genre of key clusters with Interpersonal

functions (Table 5), Textual functions (Table 6) and Ideational functions (Table 7).

Finally, results with mixed or unclassified functions are shown in Table 8 (about

which see S.3.4.2).

Table 4. Key clusters in aggregated female dialogue

INTERPERSONAL FUNCTIONS		TEXTUAL FUNCTIONS		IDEATIONAL FUNCTIONS	
Speech-act related: Directive	'I pray you' (76) 'let me go' (11) 'tell me how' (8)	**Discoursal: Question**	'will you go' (15) 'why do you' (13) 'is it that' (12) 'who is it' (10) 'who is it that' (7) 'what is your' (18) 'is your will' (8) 'what is your will' (8) 'do you speak' (6) 'what have I' (6) 'what shall I do' (8)	**Topical: People**	'my lord and' (11) 'my lord is' (6) 'is my husband' (6) 'my lord of' (7-)
Speech-act related: Expressive	'alas the day' (7) 'hope he is' (5) 'not such a' (7)	**Organisational: Informational elaboration**	'me how to' (8) 'leave me to' (5)	**Topical: Informational specificity**	'in your ear' (7)
Modalizing: Volition	'I will not' (64) 'I would have had (5)			**Topical: States**	'it is not' (30) 'to be your' (10) 'if you were' (9) 'am a woman' (5) 'I am a woman' (5)

Table 5. Key clusters with Interpersonal functions in male and female dialogue, broken down by genre

	COMEDIES		TRAGEDIES		HISTORIES	
	Female	Male	Female	Male	Female	Male
Speech-act related: Directive	'I pray you' (58) 'let me go' (8)		'go with us' (5)			
Speech-act related: Assertion	'I warrant you' (13)			'I know thee' (14)		
Speech-act related: Expressive					for God's sake (7)	'and Saint George' (8)
Speech-act related: Sincerity device						'by the Lord' (13) 'in God's name' (10)
Speech-act related: Vocative						'my Lord of' (78) 'my Lord of Winchester' (8) 'my Lord of Warwick' (8) 'Lord of Westmoreland' (14) 'Lord of York' (13) 'my Lord of York' (11) 'my Lord Protector' (14) 'my gracious lord' (22) 'your Grace to' (15)
Speech-act related: Thanking	'thank you for' (8)					

Table 5 cont. Key clusters with Interpersonal functions in male and female dialogue, broken down by genre

	COMEDIES		TRAGEDIES		HISTORIES	
	Female	Male	Female	Male	Female	Male
Modalizing: Volition	'which I would' (6) 'I would have had' (5)		'I will not' (22)	'I will not' (28-)	'be a queen' (7) 'I will not'(16)	
Modalizing: Intention	'do what I' (5)					
Modalizing: Ability	'what I can' (8)				'that I may' (7)	
Modalizing: Prediction						'the king will' (10)
Modalizing: Downtoners/ amplifiers					'and all the' (8)	'ain't please your' (15) 'please your Grace' (14)

Table 6. Key clusters with Textual functions in male and female dialogue, broken down by genre

	COMEDIES		TRAGEDIES		HISTORIES	
	Female	Male	Female	Male	Female	Male
Discoursal: Question	'what is your' (14) 'what is your will' (6) 'will you go' (12) 'will you not' (7) 'who is it' (7) 'who is it that' (5) 'how say you' (7) 'is your will' (6) 'what shall I do' (6)		'where is my' (7)			
Narrative-related: Reporting/ Reported Clause Fragments	'have heard him' (6) 'I have heard him' (4) 'I have heard of' (4)			'I have seen' (26)		
Organisational: Informational elaboration	'for such a' (7) 'now I see' (5)	'as much as' (27)		'what you have' (15)	'that thou hast' (6)	

Table 7. Key clusters with Ideational functions in male and female dialogue, broken down by genre

	COMEDIES		TRAGEDIES		HISTORIES	
	Female	**Male**	**Female**	**Male**	**Female**	**Male**
Topical: People	'king my father' (5) 'the king my father' (5)		'my lord and' (6)			'the Duke of' (90) 'the Duke of York' (20) 'the Duke of Norfolk' (11) 'to the Duke of' (7) 'the Earl of' (31) 'Prince of Wales' (21) 'Duke of Norfolk' (18) 'of the king' (30) 'to the king' (40) 'the king is' (27) 'Duke of Gloucester' (15) 'Henry the Fifth' (13) 'from the king' (16) 'the king hath' (24) 'Duke of Suffolk' (13) 'John of Gaunt' (10) 'Duke of Lancaster' (10) 'the noble duke' (11) 'Lord of Winchester' (9) 'Earl of March' (9) 'Duke of Clarence' (10) 'of the French' (10) 'noble Lord of' (8) 'John of Lancaster' (8)
Topical: Informational specificity	'is one of' (5) 'in your ear' (5)					'a cup of' (14) 'cup of sack' (9) 'a cup of sack' (9) 'in the king's (9) 'of the realm' (10)

93

Table 7 cont. Key clusters with Ideational functions in male and female dialogue, broken down by genre

Topical: States	'I am not' (25) 'I am your' (8) 'I am yours' (7) 'I see you' (7) 'I like the' (5) 'I was born' (6) 'am a maid' (6) 'I am a maid' (4) 'it is not' (19) 'it is not so' (7) 'if you were' (7) 'you be not' (6) 'if you be not' (5) 'he is in' (7) 'that he did' (7)			'I am a' (19-)		'to the crown' (22) 'the crown of' (9) 'King of England' (16)
Circumstantial: Place				'to the Capitol' (15) 'the market-place' (16) 'to the senate' (8)		'to the Tower' (17) 'in the field' (19)
Circumstantial: Directional	'look on me' (5)					

Table 8. Key clusters with mixed and unclassified functions in male and female dialogue, broken down by genre

	COMEDIES		TRAGEDIES		HISTORIES	
	Female	**Male**	**Female**	**Male**	**Female**	**Male**
Mixed	'draw the curtain' (5)			'he has done' (7)		'the house of Lancaster' (12) 'the house of York' (10) 'the realm of France' (8)
Unclassified						'the house of' (22) 'house of Lancaster' (12) 'house of York' (10) 'the realm of' (9) 'realm of France' (8)

4.2   <u>**Selecting results for in-depth analysis**</u>

As noted in S.2.4, Scott and Tribble (2006:63) caution that not all key results point to evidence of interesting stylistic features. I must identify and focus upon those that do, in order to answer my research questions (in S.1.4) about the characterisation of males and females in the plays through the functions of language they use relatively frequently. Accordingly, I made an initial assessment of all the results tabled in S.4.1, to determine some inclusion criteria. One or more of the following factors qualify results for detailed analysis:

(i)   Functional categories which show a concentration of 10 clusters or more in a single sub-category, but only where the results do not obviously arise mainly from the topic of the text. I do have a functional category of Ideational: Topical in my analysis framework (see Table 3, S.3.4.2, and the results in Table 7, S.4.1), and I acknowledge that non-topical Topical results may sound like a contradiction in terms. To clarify: I discuss some results in this category in S.4.5.1 below, but only those which clearly have an effect on characterisation or dramatisation of the play (over and above their localised function of arising from the subject matter of the dialogue).

(ii)   A clear contrast between female and male dialogue in the sex/genre breakdowns, either in terms of the absence of results with a particular function in the dataset of one sex or the other, or where males and females appear to use different kinds of linguistic forms for a particular function. Although my study is primarily concerned with the female dialogue, and more discussion space is devoted to this than to the male dialogue, the female results of course arise from comparisons with male data (directly in

the sex/genre breakdowns, and indirectly in the aggregated female data

since the reference corpus of all the plays is dominated by male dialogue,

see S.3.3.1). Comparisons with male dialogue, where possible, help put the

female dialogue into perspective, as emphasised in S.1.1.

(iii)   A relatively high keyness value, since although the keyness value on its

own does not correspond to any particular stylistic effect (see S.3.3.1),

comparatively high or low keyness values can indicate results which are

substantially more important than others within a dataset. Despite slightly

fewer clusters occurring with Interpersonal functions than with Textual or

Ideational functions, the two most key results in all the female dialogue are

Interpersonal clusters. These are the Speech-act related: Directive 'I pray

you' and the Modalizing: Volition cluster 'I will not', which have positive

keyness values of 15.9 and 12.3 respectively, i.e. more than twice the

minimum keyness threshold for inclusion in the study, (6.0, stated in

S.3.3.2). I discuss their importance in creating impressions of women in

the plays in Ss.4.3.1 and 4.3.2 respectively.

(iv)   An unconventional use of a cluster when compared to the majority of other

instances of that cluster (based on evidence from the *WordSmith*

concordance data). In my analyses in Ss. 4.2 to 4.5 I show a number of

cases of interesting individual character impressions captured in this way.

I will now discuss selected results with Interpersonal, Textual and Ideational functions

in Ss. 4.3, 4.4 and 4.5, respectively, bearing in mind the above rationale. At the start of

each section, I briefly point out notable results in the aggregated female data and the

sex/genre breakdowns. The analyses which follow are illustrated with examples from

the play-texts, including some longer detailed examples which take the form of case studies. I end each section with a summary of my findings. As I stated in Chapter 1, I focus primarily on the language of female characters, given the available resources and discussion space, and this is reflected in this chapter and in my conclusions in Chapter 5.

4.3   Key cluster results with Interpersonal functions

As indicated in Table 3 (S.3.4.2), Interpersonal functions are sub-categorised into Speech-act related and Modalizing. I discuss results with these functions in Ss.4.3.1 and 4.3.2 respectively.

4.3.1   Speech-act related clusters

In the female dialogue overall (Table 4, S.4.1), Speech-act related clusters were spread evenly between two sub-functions: Directive and Expressive. Although the breakdown (Table 5, S.4.1) shows there are very few Speech-act related clusters in the female dialogue in any single genre, they do have a greater presence in comedies. In addition to two Directive clusters, comedies also feature one Assertion ('I warrant you') and one Thanking cluster ('thank you for') in the female dialogue. As indicated in S.4.2, the Directive 'I pray you' is of particular interest for its outstandingly high keyness value in all the female data. This justifies a detailed analysis as a case study, below. In some contrast to the female dialogue, the male dialogue features no Directive clusters in any genre, although male history dialogue includes clusters which function as Sincerity devices ('by the Lord' and 'in God's name'), and different key clusters with Expressive functions occur in both the male and female dialogue in histories. The male tragedy data features the Assertion cluster 'I know thee', which provides an interesting

contrast to the single Assertion 'I warrant you' in the female comedy dialogue, so I also discuss these two clusters further below. The high concentration of clusters with a Vocative function in the male history dialogue is largely attributable to the topics of the plays, i.e. men with noble titles. The Vocative results therefore confirm what might reasonably be expected, and although this is reassuring is not worth investigating further here.

As stated above and in S.4.2, 'I pray you' is the most key cluster in the female dialogue overall, with a keyness value nearly twice as high as that of most other clusters. It is also most key in the female comedy dialogue, with a keyness value more than double that of any other cluster. It occurs very frequently in the plays, and although the results show it is much more significant in the female dialogue, male characters also use it.

According to Blake (2002:93) and Busse (2002:187-212) 'I pray you' is one of several formulaic phrases including *pray*, all of which have politeness functions approximating *please*, a more modern polite request formula. Blake (2002) also categorises 'I pray you' as a discourse marker. Blake and Busse indicate that *pray* formulae mitigate requests by adding a note of sincerity, urgency or invitation, and Onions (1982:165) mentions an invitational quality.

Culpeper and Archer (2008) discuss *pray* formulae in their recent study of requests in EME drama and courtroom trial proceedings, building on Blum-Kulka et al.'s (1989:277) concept of "support moves", which strengthen or soften the impact of a

request. Culpeper and Archer (2008:73-75) classify *I pray you* as a "pre-support

move" (i.e. one which precedes the "head act"), and they say:

> … note how *I pray you* works in its full, semantically unbleached form: it is an act of supplication, marking the relative power of the interlocutor. (Culpeper and Archer, 2008:76)

Busse's (2002) corpus-based study of all Shakespeare's plays includes an entire

chapter on these formulaic *pray* variants. He states (2002:189) that *I pray you* was the

most frequently-occurring *pray* formula in his corpus of the plays, and also that it

occurred more in comedies than in the other genres. My results (in Tables 4 and 5)

support Busse's findings and take them a little further by showing that the prevalence

of *I pray you*, particularly in comedies, is located significantly in the female dialogue

of the plays. Busse argues that there is evidence for the grammaticalization of *pray*

variants over the period in which the Shakespeare canon was written:

> The co-existence of different forms ranging from the full form with subject and object (*I pray you/thee*), to shortened forms featuring either subject or object (*I pray, pray you/thee*), to single *pray* and cliticised *prithee* illustrates the process of ongoing grammaticalization. (Busse, 2002:289)

Busse (2002:194-195) argues that the evidence for variance in level of politeness

conveyed by different *pray* formulae, based on his corpus results from all the plays, is

not as great as was claimed by Brown and Gilman (1989) in their study of just four

plays. Busse states that the *pray* variants all "signal politeness or urgency, especially

in requests and questions" (2002:289) and further that they "assert the sincerity of the

speaker" when making requests (2002:290).

Busse's research clearly indicates some general overall functions encompassing all the

pray variants, despite the possibility of language change during the timespan of the

100

Shakespeare canon, and some differences in politeness conveyed by the variant forms.

This being so, I briefly investigated the key 2-word clusters in the plays (using the

settings detailed in S.3.3.3, and the same reference corpus as for the 3- and 4-word

clusters, see S.3.3.2). This revealed that 'pray you', 'I pray' and 'I prithee' are very

positively key in the female dialogue overall, and in the female comedy dialogue. Of

course, some instances of 'pray you' and 'I pray' are embedded in the 3-word cluster

'I pray you', but there are also many non-embedded examples too, which are the

contracted forms mentioned by Busse (2002:289) (above).

These results suggest that women in the plays use a range of *pray* variants more than

men, particularly in comedies. Bearing in mind the abovementioned arguments of

Busse (2002), Blake (2002), Culpeper and Archer (2008) and Onions (1982) about the

functions and meanings of *pray* forms, this indicates that Shakespearean dialogue

represents women as more often mitigating requests than men, by:

- adding sincerity or urgency;
- humbling themselves as in an act of supplication; or
- reducing imposition through an invitational sense.

It is difficult to say whether this is because women in the plays are actually more often

in situations which justify this kind of pragmatics behaviour, however, or whether they

are simply represented as being more polite and/or more emotional than men in

comparable circumstances. It would be unwise to make general conclusions about the

relative politeness of male and female characters purely on the statistically high

presence of one variant, *I pray you*, which happens to occur as the most significant 3-

word cluster in my data. It would be worth conducting a separate study which also takes in:

- the contracted *pray* forms

- the possibility of linguistic change between early and later plays; and

- comparisons with *pray* formulae in male and female dialogue from other contemporary genres.

Given the comparatively lower power status of women in early modern England (argued in S.2.5), the results from my data might well be part of a wider trend in EME.

Since 'I pray you' is just the first of many other clusters meriting analysis in the limited space available, I will confine the remaining discussion of it to some general aspects of use by female characters shown in my data. I end with a single detailed example showing a female and male character using *pray* formulae in a politeness-based encounter, which demonstrates the greater use of them by the woman. I suggest that this helps explain the significance of 'I pray you', and other *pray* variants, in the female dialogue of the plays.

Examples 5, 6 and 7 show that 'I pray you' is used by women of varying social status (noblewoman to courtesan), and to addressees of lower, higher or equal social rank:

(5) Courtezan to Antipholus Syracuse [low to high social rank]:

I pray you, sir, my ring, or else the chain:
> *The Comedy of Errors*, IV:iii

(6) Rosalind to Orlando [equal social rank]:

I pray you, what is it o'clock?
> *As You Like It*, III:ii

(7) Countess to Clown [high   By what observance, <u>I pray you</u>?
       to low social rank]:   *All's Well That Ends Well*, III:ii

Interestingly, though, two thirds of the instances of 'I pray you' in the female comedy dialogue were addressed to men (as are examples 5 to 7). I considered whether this may simply reflect the greater numbers of male characters in the plays, which would increase the likelihood of women speaking to a male addressee. (This assumes women talk to men and not only to other women, as is generally evident in the plays.) However, if this were the only explanation I would expect 'I pray you' to occur as key in the non-comedy genres instead, since the ratio of male to female characters is far greater in these than in comedies (see Table 1, S.3.2.3).

Therefore, the significant presence of 'I pray you' in female comedy dialogue must reflect an importance that is somehow linked to the genre. It seems reasonable:

(i)   to suppose that this derives from its function as a polite discourse marker; and therefore

(ii)   to suggest that politeness is likely to be a more important conversational strategy in comedy plays because the plots often feature courtship and banter; and

(iii)   to assume these are negotiated through face-saving/face-threatening politeness strategies.

However, this line of thought is not supported by the findings of Kopytko (1993, 1995, cited in Busse 2002:23-24), which show that the tragedy genre features many more instances of politeness than the comedy genre. Busse (2002:202) argues that the type

103

of dialogue typically found in comedies is more likely to be an explanatory factor for

linguistic phenomena than plot type. Busse concludes that *pray* formulae in comedies

are "an artistic means to render the normal speech of common people" (ibid.:210).

Busse draws on the claims of Taavitsainen (1995) and Salmon (1965, 1967), who state

that the dialogue in comedy plays is more representative of natural spoken language

used by middle-class EME speakers than that in the non-comedy genres. Using

Salmon's term, comedy dialogue is more "colloquial". Since colloquiality can to some

extent be quantified (e.g. by examining syntax), this does seem a more reliable way of

assessing the dialogue in different genres of plays, since plot type is impossible to

capture empirically. However, I feel there is a relationship between the two, since

colloquial dialogue (more than formal speech) seems more likely to be the vehicle for

the romantic negotiations and banter which do often occur in comedy plots.

If we accept Busse's (2002:210) conclusion, above, that comedy dialogue is more like

natural spoken EME than non-comedy dialogue, it would then appear that the

prevalence of 'I pray you' (and other *pray* variants) in my female data is some kind of

representation of a sex-based habit of the time. Though there could be some dramatic

artifice in this, I do not think it accounts for the widespread high incidence of 'I pray

you' throughout the plays:

- 76 occurrences in the female dialogue overall;
- 56 of which are in comedies.

My examination of these led me to the conclusion that in most instances 'I pray you'

does not have a stylistic effect in the female dialogue, but seems simply to fulfil a

conventional conversational need as a polite discourse marker. In other words, it can be "discounted" stylistically in the dialogue of individual characters (Kelley, 1972, cited in Culpeper, 2001:113-115, see S.2.1). Notwithstanding this, the cumulative effect from all that data is that women are characterised as being more polite than men. This can be considered a "global" effect from the results across the collection of plays.

When I examined the concordance data I found that through their relatively high use of 'I pray you', female characters seem to be showing more politeness to men than actually seems justified by circumstances. This would support the idea (above in this section) that women mitigated requests to a greater extent than men because women had lower social status. Comedy plays, whether by virtue of the nature of the plot and/or relatively colloquial language, discussed above, often feature lengthy sections of conversation in which nothing much actually happens other than an exchange of politeness (or impoliteness). (See e.g. Blake, 2002:320-325 for more about politeness in ShE.) These seem to provide fertile ground for women to show apparently excess mitigation through repetition of 'I pray you' (and other *pray* forms), going some way to account for why it is so key in comedies. I illustrate this with an extract from *The Merry Wives of Windsor* in example 8 on the next page.

The characters Anne and Slender are not acquainted, but a proposal of marriage between them has just been negotiated by Slender's friends and Anne's parents. In early modern England, it would have been quite usual for a marriage to be arranged by a couple's parents, so Anne's situation would seem acceptable to a contemporary audience. Slender's situation, though, is rather unusual in that his friends have intervened because he is simply too passive and shy to find himself a wife (though he

does want to marry). The deal having been struck, it is now expected that they will all

have dinner together with Anne's family, but Slender is embarrassed and nervous and

doesn't want to join them. Anne and her father attempt to persuade him.

(8)   Anne: I pray you, sir, walk in.
         Slender: I had rather walk here, I thank you. […]
                  [Mr Page enters]
         Page: Come, gentle Master Slender, come; we stay for you.
         Slender: I'll eat nothing, I thank you, sir.
         Page: By cock and pie, you shall not choose, sir! come,
                  come.
         Slender: Nay, pray you, lead the way.
         Page: Come on, sir.
         Slender: Mistress Anne, yourself shall go first.
         Anne: Not I, sir; pray you, keep on.
         Slender: Truly, I will not go first: truly, la! I will not do you that
                  wrong.
         Anne: I pray you, sir.
         Slender: I'll rather be unmannerly than troublesome.
                                      The Merry Wives of Windsor, I:i

I have omitted several lines in which Slender offers various excuses for not joining the

company, but the lengthiness of this extract successfully shows the way 'I pray you'

and 'pray you' (which I identified as a key 2-word cluster earlier) occur several times

in quick succession in what is essentially an exercise in politeness. Indeed, Slender, in

the last line, claims manners as his reason for prevaricating. Though Anne and Slender

are in disagreement during this negotiation, Anne uses the polite *pray* formulae 'I pray

you'/'pray you' four times compared to Slender's single use of 'pray you' in return.

Neither Anne nor Slender wish to offend each other, and neither seems obviously

more powerful in the situation, though Anne is on home ground and might therefore

be supposed as feeling more at ease than Slender. In example 8, Anne's repeated use

of the *pray* formulae gives her dialogue a sense of sincerity and supplication that

seems questionable, given that she and Slender aren't acquainted and she is not in love

with him. Her personal character motivation for sincerity seems rather low, suggesting

that her greater use of *pray* formulae is more of an approximation of conventional female behaviour for the time. In other words, it is part of the global effect of 'I pray you' in the data overall, an idea conceived above. This is partly because *The Merry Wives of Windsor* is argued as giving the most authentic portrayal of Elizabethan social life out of all the plays (e.g. by Boyce, 1990:424-425).

It is worth noting the presence of Mr. Page in example 8, who has more power than either Anne or Slender, since it is his home and his dinner and he has the authority to give or withhold Anne's hand in marriage. He adopts a strategy of heartiness rather than politeness, using blunt Expressives ("by cock and pie") and Directives ("come", "come on, sir"), which lack the pseudo-invitational quality of 'I pray you' used by Anne. The difference in politeness between Anne and her father towards Slender therefore seems more likely to be power-based than sex-based.

Example 8 certainly demonstrates how the presence of 'I pray you' (and the 2-word cluster 'pray you') arises significantly in the dialogue of women as a social group, particularly in comedy plays. I argued above that it seems conventional for women to mitigate requests more than men, though since Culpeper and Archer (2008) note that 'I pray you' is a pre-support move in requests, it is also possible that its presence simply follows the distribution of requests in the dialogue. Verifying this by counting up all the requests would have been an unfeasibly large task. In S.4.4.1 I discuss the notably high concentration of clusters with a Question function in female comedy dialogue, a trend which lends support to this idea because:

- women seem more likely to use a *pray* formula to mitigate a request;

- requests often take the form of questions; and

- there is a greater incidence of questions in comedies than in other genres.

Also, the keyness of a Thanking cluster ('thank you for') in the female comedy dialogue only (shown in Table 5 and mentioned at the start of this section), adds further weight; more requests made by female characters would correspond with more opportunities for them to express thanks.

In addition to the wider effects of 'I pray you' in female dialogue, both overall and in comedies, there are clearly also some narrower effects creating individual character impressions. The loading of Anne's dialogue with the *pray* formulae tends to cast her in the role of pressuring Slender to do something against his will, making him seem more the victim than the would-be lover. This undoubtedly adds to the characterisation of Slender as rather vague and weak, and can be considered an "individual" effect, in comparison to the global effect of Anne's politeness seeming conventional for women at the time.

I also found some evidence that the absence of 'I pray you' creates an effect of reduced politeness in the dialogue of some female characters in non-comedy genres. Cleopatra (*Antony and Cleopatra*) and Lady Macbeth (*Macbeth*) are both prominent female characters in tragedies who have a lot of dialogue, but in contrast to prominent females in comedies, Cleopatra does not use 'I pray you' at all and Lady Macbeth uses it only once. Arguably, two women whose dialogue lacks a politeness marker commonly used by women of all social ranks (see examples 5-7) are constructed as

108

different or deviant from female normality. In effect, their dialogue is more like that of male characters. This seems to correspond with the exceptionally high levels of power they have in the plays, which is also more like that of male characters. Interestingly, the power wielded by Lady Macbeth and Cleopatra does not derive primarily from their high social rank. It derives from their personal influence over powerful men (Macbeth and Antony, respectively), which arguably brings about the downfall of both men in the plays. It would be worth comparing a range of *pray* formulae in the dialogue of Lady Macbeth and Cleopatra and a selection of other individual female characters to see whether an absence of politeness formulae established as occurring frequently (by the rest of my data, and by Busse, 2002) can be linked to:

- the relative colloquiality of dialogue;

- the plot;

- the genre (dialogue in tragedies is argued as being less colloquial, cf. Busse, 2002:202);

- the roles or personalities of individual characters; and/or

- a combination of factors.

I now move on to the two contrasting Speech-act related clusters with an Assertion function mentioned at the start of this section. I look first at 'I warrant you' in the female comedy dialogue, then at 'I know thee' in the male tragedy dialogue. There is an obvious grammatical contrast between the two clusters in the use of *you* and *thee*, and unfortunately I lack space here to document their usage in ShE. However, this is discussed by e.g. Crystal and Crystal (2002:450), and Culpeper (2001:195-199), from which it is clear that *thee* and *you* sometimes simply reflect conventional use

surrounding the relatively different social status between characters, and at other times

the pronouns convey an attitude of increased or reduced politeness or deference from

one character to another. This aspect of the clusters is worth keeping in mind.

Culpeper (2001:201) states that *warrant* was a colloquial word in EME, arguing that *I

warrant thee* is "roughly equivalent to the present-day 'I bet you'". Allowing for some

potential difference in politeness conveyed by the second person pronouns *you* and

thee, a similar meaning can be assumed for 'I warrant you' in ShE. My data shows that

it is particularly associated with one female comedy character, the housekeeper

Mistress Quickly in *The Merry Wives of Windsor.* According to Blake (2002:297), this

play features very colloquial language, even compared to other comedies (cf. Busse,

2002:202, mentioned above), but my data indicates that Mistress Quickly's

predilection for 'I warrant you' is excessive compared to other female characters. In

example 9 below she includes no fewer than five instances of 'I warrant' you in just

two (long) sentences addressed to Sir John Falstaff:

(9) Mistress   […] yet there has been knights, and lords, and
     Quickly: gentlemen, with their coaches, <u>I warrant you</u>, coach
                 after coach, letter after letter, gift after gift; smelling so
                 sweetly-all musk, and so rushling, <u>I warrant you</u>, in silk
                 and gold; and in such elegant terms; and in such wine
                 and sugar of the best and the fairest, that would have
                 won any woman's heart; and, <u>I warrant you</u>, they could
                 never get an eye-wink of her. I had myself twenty
                 angels given me this morning; but I defy all angels, in
                 any such sort, as they say, but in the way of honesty:
                 and, <u>I warrant you</u>, they could never get her so much as
                 sip on a cup with the proudest of them all; and yet there
                 has been earls, nay, which is more, pensioners; but, <u>I
                 warrant you</u>, all is one with her.
                                        The Merry Wives of Windsor, II:i

As example 9 shows, Mistress Quickly uses 'I warrant you' to assert a range of beliefs

and opinions. This finding first of all supports Blake's (2002:298) claim that Mistress

Quickly has an "effusive style" which is partly constructed through her use of many discourse markers (of which 'I warrant you' is one, ibid.:294). Over and above this, however, the address of such a voluble and colloquial speech to an unfamiliar social superior characterises her as one who either does not know or does not embrace the deferential formality which would be expected when a servant encounters a nobleman (though she does employ the socially appropriate *you*, rather than *thee*). The audience knows Falstaff is actually quite a disreputable aristocrat, who perhaps does not merit the usual level of respect, but Mistress Quickly does not since she has not met him before (though the possibility that she holds a prior opinion of him by reputation cannot be ruled out). It is most likely that its frequency in Mistress Quickly's dialogue accounts for the keyness of 'I warrant you' in the female comedy dialogue.

I now turn briefly to the Assertion cluster 'I know thee' which occurs as key in male tragedy dialogue. My concordance data shows that men in tragedies sometimes use 'I know thee' simply to convey the message that they recognise or are familiar with an addressee of lower social status than themselves. Where the addressee is clearly of lower rank, this would be a conventional use and not an example of impoliteness (cf. Crystal and Crystal, 2002:450 and Culpeper, 2001:195-199, above). What is very interesting, however, is that the data shows that apart from one notable exception 'I know thee' is always addressed to a male character. A search of the whole corpus confirms that there are no instances of it being addressed to a woman by a man in other genres, and that it is rarely used by women in any of the plays. 'I know thee' therefore seems to be habitually used by men to other men. Consequently, the single instance in which it is addressed to a woman stands out as unusual. In *Titus Andronicus*, Titus uses it to Tamora, the Goth queen who has orchestrated acts of rape,

mutilation and murder against his family earlier in the play, including the severing of

Titus's own hand. In example 10, Tamora is disguised and has been trying to convince

Titus that she is actually the spirit of Revenge. Titus responds thus:

(10)   Titus: I am not mad; <u>I know thee</u> well enough:
                      Witness this wretched stump, witness these crimson
                      lines;
                      Witness these trenches made by grief and care;
                      Witness the tiring day and heavy night;
                      Witness all sorrow, that <u>I know thee</u> well
                      For our proud empress, mighty Tamora.
                      Is not thy coming for my other hand?
                                             Titus Andronicus, V:ii

Titus's use of the Assertion 'I know thee' works on three levels here:

(i)   it conveys the message that he recognises Tamora despite the disguise;

(ii)   the pronoun *thee* (noted above as a possible way of insulting someone)
          indicates his contempt and disregard for her social status – not surprising,
          since he must by now bear unimaginable hatred towards her; and

(iii)   Titus's use of an assertion which the rest of the data shows as habitually
          used to men only seems to indicate that he is rejecting Tamora's actual
          status of being a woman.

Point (iii) above is most important, showing that Titus's language to Tamora

effectively constructs her as male. This seems consistent with her act of mercilessly

inciting her sons to rape Titus's daughter. The argument that the act itself characterises

Tamora as more male than female is not of course a new one (see, e.g., Findlay,

1999:76). However, my study has added some evidence that this is also achieved

linguistically, i.e. through other-presentation by a male character. In effectively

rejecting Tamora's status as a woman by using a typical male-to-male insulting

112

formula, Titus's exceptional level of hatred toward Tamora is successfully conveyed. Though other female characters in Shakespeare's plays commit or condone murder (e.g. Lady Macbeth), Tamora is arguably in a class of evil all her own because of her deliberate orchestration of a sexual crime against a member of her own sex.

Any sex-based social constraints attached to 'I know thee' would have formed part of the schematic sociolinguistic knowledge of an early modern audience (as would the significance of using *thee* not *you*, in point (ii) above). However, to a present-day audience 'I know thee' would probably be absorbed into a general schema of archaic language forms. The sense of recognition it conveys (point (i) above) is still evident today, but the apparent gender-based significance of its use by Titus to Tamora is lost to change over time. This has now been revealed through the application of modern corpus technology combined with some fine-grained qualitative analysis.

Having established 'I know thee' as being generally used by men towards other men, there is still the question of its keyness occurring in only the tragedy genre, and this apparently lies in the possible ways it can be used as an insult. My data shows that in all three genres, 'I know thee' potentially conveys insult in two ways:

(i)   through the status comparison implied by *thee*, and/or

(ii)   through whether or not the speaker acknowledges recognition of the addressee.

A well-known example of these combined insulting aspects occurs in the history play *Henry IV Part II*. In example 11, Henry, now King, summarily cancels out his earlier close friendship with the knight Falstaff:

(11) Falstaff: My king! my Jove! I speak to thee, my heart!
King Henry: <u>I know thee</u> not, old man: fall to thy prayers;
Henry IV Part II, V:v

(Further relevant discussion on the extract in example 11 can be found in Culpeper (2001:195), and see also Bousfield's (2007) analysis of the earlier relationship between Falstaff and Henry.)

In tragedies, however, the data shows that 'I know thee' also sometimes precedes a derogatory claim about the addressee, as shown in a particularly glorious example from *King Lear*:

(12) Kent: Fellow, <u>I know thee</u>.
Oswald: What dost thou know me for?
Kent: A knave, a rascal, an eater of broken meats; a base, proud, shallow, beggarly, three-suited, hundred-pound, filthy, worsted-stocking knave; a lily-liver'd, action-taking knave; a whoreson, glass-gazing, superserviceable, finical rogue; one-trunk-inheriting slave; one that wouldst be a bawd, in way of good service, and art nothing but the composition of a knave, beggar, coward, pandar, and the son and heir of a mongrel bitch: one whom I will beat into clamorous whining if thou deniest the least syllable of thy addition.
King Lear, II:ii

The obvious excess of the claims made by Kent in example 12 clearly create humour. It is interesting that this particular way of creating an insult – through claims made about the addressee preceded by 'I know thee' – is key in tragedies rather than comedies, although of course the tragic and the comic are not mutually exclusive (see

114

e.g. Boyce, 1990:654 about tragicomedy). There are not many examples, even in tragedies, but enough to account for the keyness of 'I know thee' in this genre.

As with the Directive 'I pray you', discussed earlier, the Assertion clusters in the Interpersonal: Speech-act related results have proved interesting at global and individual levels in the canon of plays. They help to characterise the habitual behaviour of:

(i)   groups of characters, through recurrent similar occurrences ('I know thee' is generally a male in-group assertion); and

(ii)   individual characters, through comparatively unusual uses within the rest of the data (Mistress Quickly's excessive use of 'I warrant you' compared to other female characters, and the exceptional address of 'I know thee' by a man to a woman, in the case of Titus and Tamora).

Having discussed the most interesting aspects of the Speech-act related cluster results, I now look at those with an Interpersonal: Modalizing function.

4.3.2   Modalizing clusters

Tables 4 and 5 (S.4.1) show that few Modalizing clusters occur in the data. Of these, 'I will not', expressing volition (see S.3.4.1), merits discussion as a case study because of its relative keyness value, noted in S.4.2 as being the second highest in the female data. Since 'I will not' expresses opposition to something or someone, investigation of the significance behind this is important. It may support Findlay's (1999:87-126) argument of an EME dramatic theme which concerned the rebellion of women against

a lack of social freedom, and the constraints of doing the will of others rather than their own.

The sex/genre breakdown (Table 5, S.4.1) shows that 'I will not' is positively key in the female history and tragedy dialogue, and negatively key in the male tragedy dialogue, a contrast also justifying discussion. Women in non-comedy genres clearly talk about what they don't want much more significantly than do men, especially men in tragedies. That is not to say 'I will not' is absent from female comedy dialogue; actually, the female character who uses 'I will not' more than any other is Cressida, in the comedy *Troilus and Cressida*. However, as it is not a key cluster in comedies, there is no significant disparity in its use by males and females, and Table 5 shows that the only key Volition clusters in comedies involve women talking about what they **do** want.

I examined the *WordSmith* concordance data for all 64 instances of 'I will not' in the aggregated female data. From this I was able to discern two recurrent situations which taken together account for half the instances:

(i)   women resisting commands to be silent or to speak (about certain things or to other characters); and

(ii)   women resisting commands to leave or stay.

I then looked closely at these situations in the tragedy and history dialogue, where the keyness of 'I will not' was shown to be located, as explained above. In order to see what kinds of risks the female characters take by showing opposition, and under what

116

circumstances, I conducted a wide-ranging examination taking in both the linguistic consequences of using 'I will not', such as impoliteness, and the non-linguistic consequences, like physical danger. I found that 'I will not' is a multifunctional cluster, since the opposition it conveys works in several very different ways. I will demonstrate these, beginning with two examples of wives who directly oppose their husbands' instructions to be quiet and/or to leave:

(13) Duke of York: Peace, foolish woman.
      Duchess of   I will not peace.
      York: *Richard II*, V:ii

(14) Iago: Go to, charm your tongue.
      Emilia: I will not charm my tongue; I am bound to
                      speak.
                      My mistress here lies murdered in her bed.
      Iago: […]
      Emilia: Be wise, and get you home.
                      I will not.
                                                    Othello, V:ii

In the above situations, husband and wife are of equal social rank, but the impoliteness of the wife in example 14 is much greater that that in example 13, and as a consequence Emilia puts herself at much greater risk of harm than the duchess. In example 13, the duke and duchess are having a private family conversation with their son. The situation is serious and urgent (their son has apparently committed treason) and they are united in their wish to solve the problem. These factors mitigate the lack of politeness shown between them, since the conversation continues in similar vein, and the Yorks' relationship appears undamaged. In this situation the duchess does not place herself at any risk in opposing her husband's instruction to be quiet. Example 14 is part of a public conversation, in which Emilia proclaims the innocence of her mistress Desdemona, who has been killed by Othello as a result of Iago's manipulation. Emilia's refusal to keep quiet not only publicly demonstrates a

preferred loyalty to her mistress over her husband (a lack of appropriate deference,

although the audience knows that her mistress is innocent and her husband is

treacherous), but threatens to reveal Iago's role in bringing about Desdemona's murder

(threatening his freedom). In contrast to the duchess in example 13, Emilia places

herself in mortal danger in refusing to comply with her husband's commands to be

quiet and to leave, and he kills her at the end of the scene.

The next two examples show contrasting uses of 'I will not' by women in situations

where their co-operation in staying or leaving is important to the plot, and to their

personal wellbeing or freedom:

(15) King Henry: Call her again.
     Crier: Katharine Queen of England, come into the court.
     Griffith: Madam, you are called back.
     Katharine: […]
                  I will not tarry; no, nor ever more
                  Upon this business my appearance make
                  In any of their courts.
                  [she leaves]
                                                    Henry VIII, II:iv

(16) Othello: Out of my sight!
     Desdemona: I will not stay to offend you.
                  [she leaves]
     Lodovico: Truly, an obedient lady;
                  I do beseech your lordship, call her back.
     Othello: Mistress!
                                                    Othello, IV:i

Example 15 takes place in a public context: the divorce trial of King Henry VIII and

Queen Katharine. The queen has already walked out, without being excused, and the

king directs that she should return. Although in the dialogue the queen addresses

herself to the court crier and her usher, Griffith, her refusal to go back is actually

directed at the king in front of his entire court. This is a brave but risky move, since

such public defiance towards the king shows a severe lack of deference. However,

118

although the queen has less power than the king in this situation, she has powerful

supporters who oppose the king's wish for a divorce, including the Roman Catholic

Church. It is preferable for the king to get Katharine's agreement to the divorce rather

than impose it on her. Since she eventually agrees to it later in the play, it is possible

that even at this point she suspects its inevitability. In this situation, the queen takes a

risk in refusing to comply with the king's wish for her to return to the court, but the

consequences are more likely to be the loss of her position (as wife and queen) rather

than her life. Her gesture of opposition here is a very public demonstration that she is

not going to go quietly.

Example 16 also takes place in public. Desdemona uses 'I will not' to comply with

Othello's angry command that she leave. In expressing this in a negated way (i.e. by

stating what she **doesn't** want to do), she implies that she will subordinate her own

wish to stay to Othello's wish for her to leave. Effectively, the opposition conveyed

by 'I will not' is self-directed. As example 16 shows, this is judged by Lodovico as

extremely polite and evident of her obedience, so much so that it persuades Othello to

retract his command. In this situation, Desdemona's use of 'I will not' to defer her

own volition in preference to Othello's restores her favour a little (although this is

subsequently and tragically lost later in the play). Lodovico's evaluation of

Desdemona as "obedient" in example 16 strengthens this impression of her through

other-presentation. Since Lodovico is the voice of "normal society" in the play

(according to Boyce, 1990:371) his assessment of Desdemona would have high

validity. (See Culpeper, 2001:167-172 for more on factors which affect the validity of

self- and other-presentation.)

119

Instances of women using 'I will not' in declining to speak occur only in tragedies.

These situations do not involve women refusing to talk at all, but are actually

metacommentary: the women talk about what they will not say, or who they will not

speak with. The following examples show two female characters concerned with

confidentiality:

(17) Doctor: […] what, at any time, have you heard her say?
      Gentlewoman: That, sir, which <u>I will not</u> report after her.
                                                          Macbeth, V:i

(18) Portia: Tell me your counsels, <u>I will not</u> disclose 'em.
      Brutus: O ye gods! Render me worthy of this noble wife.
                 […]
                 Portia, go in awhile;
                 And by and by thy bosom shall partake
                 The secrets of my heart.
                                                    Julius Caesar, II:i

In example 17, the gentlewoman directly resists the doctor's invitation to speak, out of

loyalty to her mistress Lady Macbeth. As a member of the court, the gentlewomen's

social rank is above the doctor's, and the risk of offending him would be relatively

low. Nevertheless, she wants to maintain his co-operation in looking after her

mistress's welfare, so she mitigates the impoliteness of her refusal to elaborate by

addressing him as "sir". In example 18, Portia assures her husband of her discretion

by expressing this in terms of what she will not do (repeat what he says to others).

Brutus, though, has not suggested she would be indiscreet, so it seems that Portia is

self-presenting as one who does not conform to the implied or assumed societal norm

that women can't be trusted with confidential information. This view can be explained

by Habermann (2003:8), who states that in EME drama women are represented much

more than men as "engag[ing] in gossip and an informal negotiation of people's

120

characters and relations", both as victims and perpetrators. Habermann sums up the

situation thus:

> Women are depicted as "brokers of oral reputation" who possess
> intimate knowledge of other people and are prepared to use it to
> further their own ends. (Habermann, 2003:8, quoting Gowing,
> 1996:123, 124.)

In light of this, Portia's assurance that she will not reveal what Brutus tells her makes

sense. As with Desdemona in example 16 above, Portia is praised for resisting the

implied natural female behaviour, and is rewarded by her husband (who promises his

confidences later, since they have been interrupted).

In examples 16 and 18 I argued that Desdemona (*Othello*) and Portia (*Julius Caesar*)

use 'I will not' to express compliance, but that the sense of opposition created by this

cluster was still present, directed either to self or to some implied schematic norm or

assumption (which might be held by other characters and the audience). An implied

sense of opposition is also present in the use of 'I will not' by women airing their

thoughts aloud in the absence of any other on-stage characters, one in a tragedy and

one in a history play, shown below.

(19)	Duchess of Gloucester:	And, being a woman, <u>I will not</u> be slack To play my part in Fortune's pageant.

Henry VI Part II, I:ii

(20)	Juliet:	<u>I will not</u> entertain so bad a thought. How if, when I am laid into the tomb, I wake before the time that Romeo Come to redeem me?

Romeo and Juliet, IV:iii

Example 19 occurs in what appears to be a short soliloquy, spoken after the duchess's

husband and the messenger leave the stage but immediately before the priest Hume

enters (though she calls out to ask whether he is there later in the same speech, making it slightly unclear at what point she realises she may not be alone). The duchess refers to her readiness to act, but in terms of not being slack. As with Portia's reference to not being indiscreet (in example 18), this implies an assumption that someone else is slack. Since the duchess attributes her lack of slackness to her biological sex, the implication is clearly that she assumes the opposite sex to be slack in the circumstances she describes. Culpeper (2001:168-9) states that the presence or absence of other characters affects the truth value of self-presentation, arguing that characters are more likely to tell the truth about themselves (as they see it) in soliloquies than when other characters are co-present. As the duchess's remark appears to be self-spoken, there seems no reason for its presence in the dialogue except to reveal to the audience her private view that women are superior in doing what is necessary. This supports Findlay's (1999:185) claim that the character of the duchess "highlights a feminist angle on the operations of power" in the play, especially since my data shows that the duchess's private self-presentation of her own sex as superior contrasts markedly with the public self-presentation of womanhood as inferior by other female characters (see S.4.5.1).

Example 20, also from a soliloquy, occurs in Juliet's final speech in the play, immediately prior to her suicide. She debates opposing choices of thought which will determine her next action, and she voices these rather than leaving the audience to infer her meaning. Juliet's use of 'I will not' is rather paradoxical, since in announcing the thoughts which she doesn't want to have, she nevertheless explains to the audience the inner mental anguish leading to her tragic decision at the end of the speech. Examples 19 and 20 demonstrate how soliloquies contribute to the

construction of a discourse level between the playwright and the audience, argued in

S.2.1 using Short's (1996:169-172) model of discourse structure. Audiences are given

access to the characters' inner thoughts through self-spoken dialogue in the

soliloquies, which other onstage characters in the play do not have.

I argued above in this section that 'I will not' is a significantly key cluster in the

female data due to its key presence in the non-comedy genres. Examples 13 to 20

have illustrated the importance of its function of expressing volition in female

dialogue, and that this takes several forms. Predominantly it expresses opposition to

the volition of another character, and very often this is over being quiet or

coming/going (e.g. Emilia, example 14). The level of linguistic consequences

(impoliteness, or at least a lack of conventional deference) and non-linguistic

consequences (risk of personal danger) seem to depend on the relationship between the

speaker and addressee, and on whether the conversation is private or public. However,

when self-directed, the sense of opposition created by 'I will not' also sometimes

functions to express compliance (e.g. Desdemona, example 16) or is directed at an

apparent schematic belief about women (e.g. Portia, example 18); this type of use is

judged as polite by male characters and rewarded. Finally, 'I will not' is used as a

means of self-presentation, to reveal the characters' opinions, attitudes and inner

mental states to the audience (e.g. Juliet, example 20).

To summarise my case study of this single Interpersonal: Modalizing: Volition cluster,

I have shown that 'I will not' essentially conveys opposition, but with different effects

depending on circumstances such as the addressee's identity and the co-presence of

others. It does not always function in direct opposition to another character's wishes,

though often this is the case, and there is a recurrent theme of women refusing to leave/stay or to be silent/speak when told to do so. However, 'I will not' also functions to reveal the inner thoughts of the characters to the audience, particularly in tragedies and histories, explaining the plot and the characters' actions (e.g. why Iago kills Emilia in example 14, and why Juliet decides to kill herself in example 20).

It is important to reiterate that although 'I will not' is used significantly more by female characters, particularly in tragedies, male characters also use it, and the effects I have illustrated in this section are not exclusively confined to female dialogue. However, they are associated more with women in the plays than with men, which chimes with what Findlay (1999:87-126) has said about women in EME drama pushing against the boundaries of social restraint (mentioned above). However, whilst my data clearly shows women resisting, they tend to be women in tragedies and histories, often with unhappy consequences: Emilia (example 14) is killed for speaking out on the side of justice, and Katharine (example 15) still has to agree to divorce the king.

This concludes my analysis of clusters in my data with Interpersonal functions, and I now move on to those with Textual functions.

4.4   Key cluster results with Textual functions

Textual functions are sub-categorised into Discoursal, Narrative-related and Organisational (see Table 3, S.3.4.2), which I discuss in turn in Ss. 4.4.1, 4.4.2 and 4.4.3 below.

4.4.1   Discoursal clusters

As Tables 4 and 6 (S.4.1) show, the only results with Textual: Discoursal functions

were fragments of questions. Question clusters constitute the biggest functional sub-

category of results in all my female data. Table 6 shows that Question clusters are

concentrated in the female comedy dialogue, but from my data I could not determine a

satisfactory genre-based reason for this (e.g. a link to plot type or relatively colloquial

dialogue, discussed in S.4.3.1, or the higher proportion of female characters in comedy

plays, mentioned in S.3.2.3). It is conceivable that the number of questions is linked to

the number of requests in the dialogue, suggested in S.4.3.1 as a possible contributory

factor to the significance of 'I pray you' in female dialogue. The concordance data

shows evidence of questions functioning in similar ways across all three genres.

Furthermore, I found that the question fragments which occur as key in the female

data are also present in the male dialogue, and used in similar ways. What is important

is that these questions are more significant in female speech than in male speech in the

plays, and therefore that they must have an important function in organising the

message communicated by the play-texts (cf. Halliday, e.g. 1994:179, see S.3.4.1).

My investigations of the Question clusters revealed several general dramatic functions

achieved through questions in the dialogue, which contribute to the successful

construction of the play-text to achieve its purpose (highlighted in S.1.2). These

include:

- communicating important information to the audience to engage them with the
  plot;
- revealing the thoughts, attitudes and personalities of characters; and

- facilitating the movement of characters and activities onstage.

My data contained numerous examples of Question clusters working in these ways, from which I have selected just a few (from all three genres) to illustrate the distinctive functions they have in organising the play-text and presenting it to the audience in a coherent way. Following these, I conduct a case study of the overlapping 3- and 4-word Question clusters 'what is your' and 'what is your will' in the data, which show some particularly interesting stylistic effects through comparatively unusual uses.

I begin with a wh-question requesting information, featuring the cluster 'what is your':

(21)	Katharine:	But in this changing <u>what is your</u> intent?
	Princess:	The effect of my intent is, to cross theirs:
		Love's Labour's Lost, V:ii

In example 21 Katharine's question to the Princess elicits an onstage explanation of the process by which their male admirers will be deliberately deceived. This question helps set up the two distinctive discourse levels in the play (discussed in S.2.1 and noted in the self-spoken examples 19 and 20 in S.4.3.2). The question in example 21 brings the audience in on the joke, since it gives them information which the onstage admirers will not have, equipping the audience to derive amusement from the admirers' confusion later in the play.

Example 22 shows a wh-question which includes the cluster 'is it that', occurring four lines into a 28-line speech (most of which I have omitted to conserve space):

126

(22)	Lady Percy [to Hotspur]:	Tell me, sweet lord, what <u>is it that</u> takes from thee Thy stomach, pleasure, and thy golden sleep? […]

Henry IV Part I, II:iii

Lady Percy asks her husband what is disturbing him, but instead of pausing for his reply she continues telling him how worried she is about him for a further 23 lines. Rather than eliciting information from another character, the question in example 22 is a launchpad for a long self-revelatory speech. The question raises a concern which Lady Percy then expands on, revealing her state of mind. This helps shape her personality and also the (harmonious) relationship between the Percys which is further demonstrated later in the scene.

Example 23 also features a wh-questioning cluster, 'why do you':

(23)	Lady Macbeth [to Macbeth]:	<u>Why do you</u> make such faces?

Macbeth, III:iv

Lady Macbeth cannot see the ghost of murdered Banquo which is unnerving her husband. Interestingly, she asks specifically about his facial behaviour rather than making a general enquiry about what is the matter. The plays contain other similar instances of a 'why do you' question making specific reference to a behaviour manifested by another character, often one which is non-verbal or muffled in some way (e.g. weeping, whispering). Examples include:

- "why do you weep" (Desdemona, *Othello*, IV:ii and Marina, *Pericles*, V:i);

- "why do you wring your hands" (Clarence's daughter Margaret, *Richard III*, II:ii); and

- "why do you speak so faintly" (Desdmona, *Othello*, III:iii).

There seem to be two probable reasons for this kind of questioning. Culpeper and McIntyre (2006:773) argue that Shakespearean audiences talked, ate and were otherwise often distracted during performances, and that playwrights used various strategies to help ensure that audiences did not lose the plot. Foregrounding certain important physical behaviours through the verbal 'why do you' questions which occur in my data would firstly reduce the chance of audiences missing these if they weren't attending to the stage. Secondly, these questions seem likely to be stage directions embedded into the dialogue to cue the performance of certain behaviours by the actors. Aston and Savona (1991:75), cited in Culpeper and McIntyre (2006:776-777), argue that these "intra-dialogic" stage directions were common in EME drama, since "extra-dialogic" stage directions (i.e. those written separately on the script, not to be spoken aloud) were fewer. In example 23, Lady Macbeth's question cues the actor playing Macbeth to make facial contortions, the performance of which helps signal Macbeth's worsening mental state.

The final example showing the way questions contribute to the organisation of the play-texts is an invitation, one of several similar instances including the cluster 'will you go' which occur near the end of scenes:

(24) Bawd [to   Pray you, <u>will you go</u> with us?
     Marina and   *Pericles*, IV:ii
     Boult]:

In example 24, one character invites one or more others to go somewhere (or to go and do something), which provides a convenient way for some or all onstage characters to depart. Then, the action can move on, usually by ending one scene and starting

another (although sometimes in the same scene with a revised set of characters). This Question cluster is a formula for moving the onstage activity forward.

I have now discussed the ways Question clusters contribute to the general textual organisation of the plays. As indicated at the start of this section, I will now present a case study showing individual characterisation which arises from stylistically unusual uses of Question clusters in comparison to the norms established by the majority of instances of them in the data.

The contextual evidence from the instances of 'what is your will' in my data strongly suggests it to be a conventional politeness formula used in response to a greeting or summons (summonses, greetings and responses were among the conversational strategies in the plays identified by Blake, 2002:283-290, see S.1.3). This is also true for "what is your pleasure", a question which accounts for several instances of the key Question cluster 'what is your'. Both formulae imply the acknowledgement of a duty of obedience from speaker to addressee, and I refer to them together in this analysis as 'what is your will [pleasure]'. I could not find any other research confirming general EME usage of these expressions. I can, however, establish the norms of their use in Shakespeare's plays, as I explain below, then pinpoint some interesting ways in which these are breached for stylistic effect.

To gauge the norms of use in the plays overall, I examined the contextual circumstances of all instances of 'what is your will [pleasure]' in all three genres, although its keyness is confined to comedies (shown in Table 6, S.4.1). Where no constraints (such as conflict or subterfuge) seem to be in effect which might have a

129

distorting effect on conventional language behaviour, I judged the usage to be within the boundaries of normal social expectations. Typical usage includes adult daughters responding with 'what is your will' to greetings or summonses from mothers and fathers. Female characters often use it in response to greetings or summonses from male acquaintances of equal or higher social rank. This includes women responding to would-be suitors (i.e. men with whom they have a relationship but to whom they are not married). It is **not** generally used by married women in response to their husbands, but two instances in the dialogue of married women stand out as unusual (discussed below). Though it is used significantly more by female characters, 'what is your will' is also sometimes used by male characters, always to other males and usually (though not always) to someone of higher social rank. Men do not use "what is your pleasure", apart from one instance addressed to a female character which stands out as unusual, also discussed below.

In the first place, therefore, the clusters 'what is your will' and 'what is your' in the comedy results (Table 6) seem to be fragments of related language formulae ('what is your will [pleasure]'). Their prominence in the female comedy dataset is interesting, since women in comedies are clearly proffering more responses which acknowledge a duty of obedience – apparently consistent with the lower social status of women at that time, argued in S.2.5. I discuss one example from a comedy below, but also two from a tragedy. The statistical significance of the clusters in the comedy results proved a useful starting point, but in this case the qualitative analysis shows that the most remarkable effects were not confined to comedy plays. This bears out the argument in S.1.2 that not everything which is interesting will occur as key in a corpus study. Since my main aim is to investigate sex-based rather than genre-based differences in

the language of the plays, the inclusion of examples which best achieves this is justified.

As discussed in S.2.1, Culpeper (2000:293-295, 2001:86-88) argues that audiences interpret characters in drama by bringing to bear schematic knowledge of dramatic character types, for example a set of known or expected behaviours that typify a "villain" or "hero". He argues that this is similar to the ways people form impressions of other people in real life:

- as individuals;
- as holders of a particular social role; or
- as members of a group (e.g. a particular sex).

In S.1.3 I mentioned that the social role of wife would be worth investigating in Shakespeare's plays, although not as the main focus of this study. This is a social role for which early modern (and present-day) audiences would be likely to have a schema-based interpretation. I now show how the use of 'what is your will' by two female characters contributes to the characterisation of them specifically in their roles as wives.

As outlined above, the data suggests that using the formulae 'what is your will [pleasure]' to parents and male acquaintants is a conventionally polite or deferential way for women to respond, though not generally for wives to husbands. Boyce (1990:626) and Culpeper (2000), amongst others, argue that EME society widely held the belief that wives were subordinate to husbands (rooted in Biblical interpretation).

131

Despite this, the fact that nearly all the wives in the plays don't respond to their husbands with these formulae suggests that the two who do are showing unconventionally high levels of respect or deference.

Example 25 shows Katharina responding to her husband Petruchio using 'what is your will':

  (25)   Katharina   <u>What is your will</u>, sir, that you send for me?
                                         The Taming of the Shrew, V:ii

Prior to Katharina's marriage to Petruchio (her reluctance to which is central to the play), she has been portrayed as a difficult and disobedient woman. The extent to which her language reflects her change of character to one of obedience and submission, particularly in this scene, has been much discussed (see e.g. Culpeper, 2000; Cooper, 1998; Boyce, 1990:623-627), and space does not permit me to recount this here. However, Katharina's act of responding to Petruchio's summons in this scene is the subject of a crucial bet as to whose wife is the most obedient, and in using 'what is your will' she shows exceptional deference, outside that conventional for wife to husband based on the rest of the data.

Boyce (1990:627) argues that Katharina's revised attitude of submission and affection to Petruchio would have seemed in order to Shakespearean audiences. He dismisses the idea that Katharina might have appeared enslaved to Petruchio rather than in love with him as an erroneous present-day interpretation of historical dialogue. However, my examination of the data in the plays overall suggests that Katharina's use of 'what is your will' to her husband would have seemed an unconventionally deferential response from a wife at the time the plays were written. This also chimes with

132

Culpeper's (2000:310) argument that Katharina uses excessive force in promising to

obey Petruchio slightly earlier in the play: again, she would be going over the top in

the eyes of the audience in displaying obedience. This might have contributed to an

impression of her as broken rather than wooed into obedience, although it could also

perhaps be an ironic use of an unusually deferential form. It might simply foreground

the strength of Katharina's change of heart.

Culpeper (2000:312) argues that Katharina is a multifaceted character who evokes

more than one schema of the social role of wife (i.e. a shrewish, difficult wife and an

exemplary, obedient one), in whom Shakespeare makes a point about changing social

attitudes to women. Additionally, as mentioned in S.2.1, Cooper (1998) argues that the

interpretation of Katharina's behaviour depends upon what one assumes as being the

social norms of the period (after comparing the contrasting claims of Kahn, 1977 and

Heilman, 1972). If nothing more, my evidence for the unconventional use of 'what is

your will' supports these two theorists' broad conclusion that there is more than one

plausible interpretation of Katharina's change of attitude towards her husband. It is

this scope for possibility that allows Shakespeare's plays to be interpreted and re-

interpreted by successive generations of critics, a phenomenon pointed out in S.1.1.

The other wife using 'what is your will [pleasure]' in this case study is Desdemona,

responding to her husband Othello, shown in example 26:

(26) Desdemona: My lord, <u>what is your will</u>?
     Othello: Pray, chuck, come hither.
     Desdemona: <u>What is your</u> pleasure?

<div align="right">Othello, IV:ii</div>

Again, some background is necessary to help see why these apparently unconventional

formulae are there. Othello has been deliberately misled by Iago into thinking that

Desdemona has been unfaithful. Othello has not actually confronted her with this, but

in the previous scene he has been angry and violent with her and told her to go away.

Her submissive response to this earned her praise for obedience by Lodovico, which I

argued as other-presentation of Desdemona's character (in S.4.3.1, see example 16).

Othello now summons Desdemona after being assured of her innocence by Emilia

(Iago's wife). Desdemona's response, in example 26 above, doubly acknowledges a

duty of obedience to him. This exceptionally deferential response reinforces the

impression of her as obedient which was constructed through other-presentation in

example 16 (S.4.3.2) – this time, through self-presentation. There seems no possibility

of irony here, as with Katharina in example 25; instead, Desdemona's exceptional

obedience among wives in the plays elevates her to an idealised level, which would

make the gradual dissipation of Othello's faith in her, culminating eventually in her

death, appear all the more tragic.

It is worth adding that although the evidence from my data shows Desdemona being

characterised as unusually obedient, Jardine (1983:119-120) claims that elsewhere in

the play she displays "vulgar" behaviour, arguing with Iago about "womanish wiles"

(Act II:i). Jardine (1983:184) argues that the characterisation of Desdemona alters to

one of virtue as the play proceeds, coinciding with Culpeper's (2001: e.g. 284-285)

argument that schemas of characters do not necessarily remain static. This suggests

that a complete picture of a character cannot be determined solely on the basis of what

is statistically significant in the sum total of her/his dialogue.

The final untypical use of the 'what is your will [pleasure]' response formulae in my

data is the use of "what is your pleasure" by Iago to Desdemona, shown in example 27

below. Following the extract in example 26 (above), Othello accuses Desdemona of being a "whore" and a "strumpet"; Desdemona denies this and Othello leaves. Desdemona, seeking support, asks Emilia to fetch Iago. They return together, and Iago greets Desdemona with the following:

(27) Iago: <u>What is your</u> pleasure, madam? How is it with you?
                                                  Othello, IV:ii

As argued above in this section, the formulaic response "what is your pleasure" is not used by any other men in the plays. As with the Assertion cluster 'I know thee' addressed to a woman (in example 10, S.4.3.1), it therefore breaks with habitual sex-based use as evidenced by the rest of the data. Though it may not reflect wider EME speech, it is still intriguing, particularly considering that examples 25 and 26 show these formulae signalling unusual deference from wives to husbands. If Iago had been speaking to his own wife, this might have been an ironic reversal of a recurrent stylistic device, but instead he is addressing Othello's wife, whose marriage he is deliberately undermining in the play. Since Iago is not co-present during the exchange between Othello and Desdemona in example 26, he cannot be slyly mocking Desdemona's recent display of obedience by parroting her words (unless he has been eavesdropping, but there is no evidence of this in the stage directions). Iago appears to be showing Desdemona a display of unusual deference through responding to her summons in this way, which the audience (although not Desdemona) knows to be false since he is the instigator of her troubles. The cruelty of this would reinforce the impression of him as an evil character. Lying and cruelty would seem typical behaviour of a villain, remembering Culpeper's (e.g. 2000:293-295) claim that audiences make schema-based assumptions about dramatic types (see S.2.1). Also, if "what is your pleasure" was not a formula conventionally used by men to women in

EME, this would have stood out to a contemporary audience and characterised Iago as deviant, perhaps socially malign. There is no conclusive explanation for the sole presence of this cluster in Iago's dialogue, but any deviant, non-male impression seems swiftly cancelled out by his next enquiry (also shown in example 27), since a quick search of the sex-specific data files reveals that "how is it with you" is a general formulaic enquiry made exclusively by men in the plays.

This concludes the case study, and I now summarise the Discoursal clusters, all of which had a Question function as noted at the start of this section. These have been shown to have general functions of textual organisation which are more prominent in the female dialogue, illustrated in examples 21 to 24, in addition to contributing to individual stylistic effects in the plays as do clusters with other functions in the data.

Question clusters are used:

(i)   to achieve the multiple discourse levels which enable the audience to benefit from a greater knowledge of what is going on than the on-stage characters (demonstrated by example 21);

(ii)   to provide a way for characters to reveal their inner mental states to the audience, conveying background information about their relationships, feelings and attitudes (as in example 22); and

(iii) to fulfil practical purposes such as ensuring that the audience notices important character behaviour (example 23) and facilitating the forward motion of onstage activity through the division of the play into scenes (example 24).

These functions are important to the success of the play, and my analyses have shown more about how the playwright achieves this by building formulaic devices set around questions into the dialogue. Though male dialogue also features these questions working in similar ways (stated earlier in this section), their prominence in the female data shows that the successful communication of the play is supported slightly differently by the dialogue of characters of different sex. It is important to remember that questions are not the only way in which the contents of play-texts are structured and delivered to the audience. Soliloquies constitute another method, as mentioned during my discussions of the Interpersonal: Modalizing cluster 'I will not' in S.4.3.2, and it is likely that male dialogue performs a greater role in conveying information in this way. Unfortunately, I located no data to support this, but it would make for useful further study.

The case study has demonstrated the usefulness of having results based on an entire collection of plays, since these can be analysed to establish norms or habitual uses of language formulae. The norms I identified for 'what is your will [pleasure]' can be considered global effects arising from my data, as I argued with regard to 'I pray you' in S.4.3.1. Against these, stylistically remarkable exceptions can be compared to show individual effects in the dialogue of particular characters, as in examples 25 to 27. As noted at the start of this section, the Discoursal: Question clusters formed the largest category of results in my data. There has been a lot to say about them in comparison to the other types of Textual cluster, and I now move on to briefer discussions of these.

4.4.2   Narrative-related clusters

My data includes very few clusters with Narrative-related functions (i.e. those in which characters report speech, thought, writing or actions; see Table 3, S.3.4.2). None occur in the female dialogue overall (Table 4, S.4.1). Those that do occur in the sex/genre breakdown (Table 6, S.4.1) show an interesting contrast between the verbs used in first person narrative reports of male and female characters:

- female characters (in comedies) report what they have **heard** significantly more ('I have heard him', 'have heard him', 'I have heard of');

- male characters (in tragedies) report what they have **seen** ('I have seen').

The contrast occurs across different genres, but as with the prevalence of Question clusters in comedies in the previous section, I could find nothing obvious about the genres to explain this, so it seems to be a sex-based contrast. The concordance data for the female comedy dialogue shows that the clusters introduce hearsay or overheard dialogue. Three examples are given below:

(28)   Celia [to   O! I have heard him speak of that same brother;
       Oliver]:   And he did render him the most unnatural
                       That lived amongst men.
                                                As You Like It, IV:iii

(29)   Viola [as   I have heard of some kind of men that put
       Cesario, to Sir quarrels purposely on others to taste their valour;
       Toby Belch]:   belike this is a man of that quirk.
                                                Twelfth Night, III:iv

(30)   Isabella to   I have heard of the lady, and good words went
       Duke]:   with her name.
                                                Measure for Measure, III:i

Examples 28 to 30 show women presenting hearsay as evidence to justify their opinion or attitude about another character. In about half the instances in my data,

138

claims made by women in this manner are promptly disputed by another character. In example 29, Viola/Cesario's disparaging assumption about the "knight" invented by Sir Toby Belch, based on what she has heard about men in general, is immediately contradicted by Sir Toby. Since Viola is disguised as a man at this point in the play, these remarks would seem highly amusing to the audience (who have greater knowledge of the situation than the onstage characters).

The importance of hearsay in claims about reputation seems more than a comedy device, however. Habermann (2003:1) argues that slander, or damaging someone's reputation by verbal means, was "a conspicuous phenomenon" during the early modern period, and her discussions include many examples from Shakespeare's plays. Habermann's (2003:8) claim about the significance of gossip in the representation of female characters in EME drama, also mentioned in S.4.3.2 above, seems broadly supported by my evidence that female characters report hearsay significantly more than do male characters. Although the results occur as key only in the comedies, this could be because female characters are proportionally slightly better represented in comedies than in tragedies or histories (see S.3.2.3). My data does not show that the hearsay is associated with outright slander, though; in fact, as example 30 shows, some women use hearsay to justify favourable views which they hold about other characters. Hearsay seems more to be a device through which women show how they acquired the knowledge which underlies their motivations.

Turning now to the male tragedy dialogue, the concordance data shows that men use the cluster 'I have seen' to provide evidence for their beliefs and opinions, much as do

the female comedy characters using 'I have heard of/him'. Three examples are given

below.

(31)	Casca [to Cicero]:	<u>I have seen</u> tempests, when the scolding winds Have riv'd the knotty oaks; and <u>I have seen</u> The ambitious ocean swell and rage and foam, To be exalted with the threatening clouds: But never till tonight, never till now, Did I go through a tempest dropping fire. *Julius Caesar*, I:iii
(32)	Antonius:	Octavius, <u>I have seen</u> more days than you: *Julius Caesar*, IV:i
(33)	Iago:	Alas, alas! It is not honesty in me to speak What <u>I have seen</u> and known. *Othello*, IV:i

In example 31, Casca believes the thunderstorm is a bad omen, but Cicero is sceptical.

In trying to convince him, Casca bolsters his position by claiming that his

interpretation of the weather is based on knowledge from what he has seen (of other

storms in the past). In example 32, Antonius is attempting to strengthen his argument

with Octavius, concerning the merits of another character (Lepidus). Antonius claims

greater age and superior knowledge or experience through the metaphor of having

"seen" more. The strength of equating visual evidence with knowledge is actually

made explicit in example 33. Iago claims to have "seen and known", implying that the

latter is a consequence of the former.

Whilst the significance of hearsay/what is overheard in my female comedy data does

not seem to contradict Habermann's (2003) argument that EME women in drama are

associated with gossip (above), my investigations actually reveal a direct contrast in

the way women and men are represented as citing evidence of what they know and

believe. What is particularly striking is the variance in **disputability** of the reports

presented through female hearsay/overhearing and male observation in the plays, when these are counted up overall. As stated above, in the female comedy dialogue about half the instances reporting hearsay or overheard speech are disputed by other characters, and sometimes this is clearly a device for humour as in example 29. In contrast, none of the instances in which men report what they have seen (in the tragedies) are disputed, though in reality these could surely be disputed as easily as a report of what has been (over)heard.

This raises the intriguing question of whether or not there is any deliberate irony in the contrast in the ways women and men present their claims through hearing and seeing in the plays, and/or in the tendency for women's claims to be evaluated as less reliable. In discussions of her cluster data from Dickens's *Great Expectations*, Mahlberg (2007b:29-30) mentions results which contribute to a theme of "seeing is knowing" in the narrative of the character Pip. Mahlberg argues that her findings build upon Quirk's (1961:23) claim of the novel's irony in fact being that "seeing is not knowing". In contrast to Mahlberg's data (which is of course from fiction not drama, and from a later historical period), the pattern in mine seems to supersede individual characters and indeed plays, only becoming clear as a sex-linked theme when the plays are examined collectively. If there is any irony in the ways men and women are presented by Shakespeare as seeing/knowing and hearing, it would therefore have to be recognisable to the audience in isolated instances in the plays (perhaps picking up on known stereotypes of male/female language at the time). Unfortunately, I could find no evidence which would confirm or refute this idea, but it would make for interesting wider investigation in EME drama and other contemporary genres.

In summary, though the Narrative-related clusters in the data are scarce, they have revealed an intriguing difference in the way men and women are represented in the plays. For male characters, seeing is believing, and this is indisputable; for female characters, hearing is believing, and this is easily disputed.

This brings me to the final set of results with Textual functions: the Organisational clusters.

4.4.3   Organisational clusters

The few clusters with Organisational functions which occur in my data are those sub-categorised as Informational elaboration, shown in Tables 4 and 6 (S.4.1). Put simply, they provide dialogic opportunities to add information to what has already been said, and in general this helps explain the background to character motivations and to events in the play. One especially interesting strategy for this appears to be exclusive to female characters, so I discuss this below.

The cluster 'me how to' from the overall female data (Table 4) is always used in the context of teaching or learning. Women talk about what they have been taught (by other characters, both male and female, or from experience), or what they want to learn (from another character). The cluster 'now I see' in the female comedy data (Table 6) is not dissimilar, since it always introduces some information which the female speakers have become aware of or recently understood; it reveals the acquisition of some new or updated knowledge. An example of each cluster follows:

   (34)   Lady Anne [to   But since you teach <u>me how to</u> flatter you,
          Richard]:   Imagine I have said farewell already.
                                                    Richard III, I:ii

142

(35) Widow [to   Now I see
      Helena]:   The bottom of your purpose.
                                        All's Well That Ends Well, III:vii

I considered the possibility that the prominence of teaching, learning, understanding

and realising in female dialogue might have the effect of characterising women as

relatively less knowledgeable to start with, or able to be manipulated (by being told

things). Given that women have a lower social status than men at this time (argued in

S.2.5), and the evidence that they are represented as believing disputable things which

they have heard (in the previous section), this would not seem surprising. Certainly

Lady Anne, in example 34 above, is beguiled into believing that Richard murdered her

husband because he (Richard) was in love with her. However, the rest of my data

indicates that in general women are not falsely taken in by what they have learned.

Instead, in talking about teaching, learning and understanding, the female characters

also voice what the audience has learned or understood about what is going on in the

play. In example 34, Lady Anne says that Richard has taught her flattery, but in doing

so Richard has also taught the audience what a skilful and evil manipulator he is. In

example 35, the audience as well as the widow is now aware of Helena's plan to get

her husband Bertram back.

As with the Discoursal: Question clusters in S.4.4.1, the Organisational: Informational

elaboration results have revealed another strategy by which the female dialogue

contributes to the successful communication of the play to the audience.

I summarised the Discoursal and Narrative-related results at the end of Ss. 4.1.1 and

4.1.2 respectively, so I will not do so again here. As with the Interpersonal clusters in

S.4.1, the Textual clusters have shown some interesting global and individual effects in the data, i.e. overall impressions of male and female characters across the canon of plays, with exceptional uses (within the canon) leading to specific character impressions. Over and above that, their role in organising and communicating the play-text itself has been shown to be built in to the male and female dialogue slightly differently, in that women make more use of certain types of question strategies (S.4.4.1) and women use discourses of teaching and learning to convey information at the playwright-audience discourse level.

This brings me to the third main type of results: those with Ideational functions.

4.5   Key cluster results with Ideational functions

Ideational results are sub-categorised into Topical and Circumstantial functions (see Table 3, S.3.4.2), which I discuss in Ss. 4.5.1 and 4.5.2 respectively.

4.5.1   Topical clusters

As explained in S.3.4.2, Ideational: Topical clusters are those identified as having a function of referring to a topic, and they are sub-categorised into those concerning people, states (including states of affairs, literal and metaphorical states, and physical or attitudinal states), and those providing "informational specificity" (by narrowing down the topic). As Topical clusters arise from the topics of the dialogue, I anticipated that many results in this category would not point to stylistically interesting phenomena in the language of the plays (Scott and Tribble, 2006:63, see S.2.4), and would therefore not warrant much further discussion. As with the Interpersonal: Speech-act related: Vocative clusters in S.4.3.1, most of the numerous Topical: People

clusters in the male history dialogue (shown in Table 7, S.4.1) arise simply because the plays are about royal and noble men.

However, as indicated in S.4.2, the concordance data indicates that some patterns of Topical results also have an effect on characterisation or dramatisation of the play, over and above their function of conveying the topic of the dialogue. These are People clusters in the female dialogue overall (Table 4, S.4.1) and States clusters in the female comedy dialogue (Table 7). Both types reveal contrasts in the male and female dialogue which justify discussion.

The People clusters in the aggregated female data (Table 4) all contain terms of reference for men, or more precisely, male social roles. As there is no corresponding aggregated male dialogue (explained in S.3.3.1), it is not possible to confirm whether women talk more about men than do men about women (though this seems highly likely, since many more men feature in the plays; see Table 1, S.3.2.3). However the data does show an interesting contrast in the ways women and men talk about men, revealed through the positively key clusters 'my lord and', 'my lord is', 'is my husband' and the negatively key cluster 'my lord of' (Table 4).

The concordance data for these positively key clusters shows that they introduce information about the background or character of men, as shown in examples 36 and 37.

(36)	Nurse [to Lady Capulet]:	<u>My lord and</u> you were then at Mantua. *Romeo and Juliet*, I:iii
(37)	Queen Margaret [to Cardinal]:	Henry <u>my lord is</u> cold in great affairs, *Henry VI Part II*, III:i

In example 36, the nurse provides some background information about Juliet's family to the audience, and Queen Margaret's remark about her husband in example 37 tells us something about his character. Both use the conventional honorific term "my lord", without adding the rest of the man's title. In contrast, the negatively key cluster 'my lord of' is an honorific term which always precedes a man's title, and is truly limited to topical reference, as shown in example 38.

(38)	Goneril [to Cornwall]:	Where is my lord of Gloucester? *King Lear*, II:iv

All we learn from this is which man Goneril is talking about; nothing of his life or character is revealed. That this kind of use occurs less than would be expected in female dialogue suggests that men are more likely to talk about what other men are doing in the plays, whereas women tend to talk about what the men are like, i.e. to talk more intimately about them (from the evidence of the positively key clusters).

Turning to the States clusters, those in the female comedy dialogue are striking (in Table 7, S.4.1) firstly because there are a substantial number of them. Most notably, however, nearly half of them begin with the first-person singular pronoun *I* followed by the copular verb *be* ('I am not', 'I am your', 'I am yours', 'I was born', and '[I] am a maid'). This pattern suggests that women in comedies talk significantly more about states which affect or involve them than do men. In marked contrast, the cluster 'I am a' is negatively key in the male tragedy data, suggesting that men talk less about themselves than would be expected in this genre. In S.4.3.2 I mentioned the ways characters present themselves and other characters through their dialogue, citing Culpeper's (2001:167-172) discussions of self-presentation and other-presentation in drama. From the prevalence of *I* clusters in female comedy dialogue, it is clear that

146

self-presentation is particularly significant here. The concordance data shows that through self-revelatory talk, women in comedies present information about their backgrounds and motivations. As with previous discussions of trends which are prominent in the female comedy data (e.g. in S.4.3.1), there is the possibility that this is due to a higher proportion of female characters. However, it is also plausible that more background information is included for audiences to understand comedy plots, since these tend to involve multi-stranded interpersonal relationships requiring introduction and explanation. History plays and some of the tragedies follow real events, to some degree, with which early modern and indeed present-day audiences would probably have some existing awareness (see the discussions of *Henry VIII* later in this section).

What is interesting, though, is that as with the Question clusters (in S.4.4.1), this particular explanatory strategy seems to be located more in the female dialogue. One particular *I* cluster stands out in my data, however, leading to an interesting comparison between the self-presentation of male and female characters. Table 4 shows that in the female dialogue overall, only one *I* cluster supersedes genre: the 4-word cluster 'I am a woman'. The statistical significance of 'I am a woman' in the female dialogue is of course unsurprising, given that female characters are far more likely to say this than male characters. A search of the aggregated male data file (the male sub-corpus of dialogue) confirms that male characters use the corresponding formula "I am a man", and if it had been possible to obtain key results from this dataset (see S.3.3.1) this formula would almost inevitably have occurred as key. The important point here is that although both male and female characters explicitly self-construct gender by identifying themselves as men and women, they evaluate their

biological sex in contrastive ways. This is revealed by the concordance data for 'I am a woman' in the female cluster results, and through a manual search of instances of its counterpart in the aggregated male dialogue, as I explain in the following case study.

As made clear in Chapter 1, and discussed further in S.2.5, my aim is to see what I can add to the existing vast commentary on language and gender in the plays through an empirically-based investigation of the male and female dialogue. The scope for flexible, different and possibly contentious interpretations of gender and relationships between the sexes in Shakespeare's plays, argued by writers such as Findlay (1999) and Jardine (1983), is important to bear in mind when analysing the self-referential cluster 'I am a woman' in the female dialogue, and its counterpart "I am a man" in the male dialogue.

The concordance data shows that female characters who use 'I am a woman' always frame it as an acknowledgement of some kind of deficit, as I illustrate with three examples below, discussed in turn. Example 39 is from the speech discussed in S.4.3.1, in which Portia is trying to get her husband Brutus to tell her what is going on.

(39)	Portia [to Brutus]:	I grant <u>I am a woman</u>, but, withal, A woman that Lord Brutus took to wife; I grant <u>I am a woman</u>, but, withal, A woman well-reputed, Cato's daughter.

<div align="right">

Julius Caesar, II:i

</div>

Portia acknowledges her sex, implying it is a limitation, then claims this is mitigated through her connection to excellent men (himself and her father) through the social bonds of wifehood and daughterhood. Brutus is impressed by her speech (in which she later points out her demonstration of male-like toughness in deliberately wounding her own leg), and he agrees to confide in her. Though Portia's dialogue implies

womanhood to be an inferior status, her act of weaving it into a persuasive strategy to

achieve her goal actually shows it working to her advantage.

Example 40 is from Queen Katherine's discussion with Cardinals Wolsey and

Campeius, in which they eventually convince her to accede to Henry VIII's demand

for a divorce.

  (40) Katherine: Alas! <u>I am a woman</u>, friendless, hopeless.
                                 […]
          Campeius: You wrong your virtues
                                 With these weak women's fears:
                                 […]
          Katherine: You know <u>I am a woman</u>, lacking wit
                                 To make a seemly answer to such persons.
                                                                                 Henry VIII, III:i

In example 40 Katherine alludes to her weaker and less powerful status as a woman,

and Campeius similarly refers to the weakness of women. As with Portia's speech in

example 39, Katherine's actions in the play seem to belie her assertions that women

are weak. Katherine actually puts up quite a lot of resistance to the king, threatening

to appeal to the Pope. She is not "friendless", but has allies in the Church and

therefore political influence, which would count for more than biological sex in

determining her fate. Katherine's claim that as a woman she lacks the ability to

counter the arguments of the men does not ring true, since she has successfully done

so thus far in the play. Boyce (1990:332) argues that Katherine "concedes her

helplessness, but refuses to co-operate in her own downfall". Yet in fact Katherine

does co-operate, by verbally contributing to her own disempowerment through her

claim that her womanhood makes it impossible for her to resist.

Of course the storyline of the play cannot diverge too much from the historical events upon which it is based, so Katherine must eventually be divorced from Henry. This has to be achieved dialogically, in a convincing way. However, from Katherine's dialogue in the play, her womanhood and not her (stated) insufficient "wit" seems to be the crux of her acquiescence to the cardinals (and the king). Boyce (1990: 332-333) argues that Katherine is characterised as a "good" woman in the play, and her act of submitting to male authority essentially on the grounds of her womanhood would certainly fit with a contemporary audience's understanding that women are subordinate to men and much less powerful (based on Boyce, 1990:626 and Culpeper, 2000 mentioned in S.4.4.1).

However, in submitting to male authority, Katherine has to reject the authority of Catholicism, which does not permit divorce. The difficulty of her position, and the way her character presents this through the dialogue, would surely have been much discussed by Shakespearean theatregoers. The tumultuous upheaval of England's separation from the Catholic Church, catalysed by Henry VIII's divorce, had occurred only about a hundred years before the play was written, and the religious and social consequences were still being felt. (Shakespeare's play *Henry VIII* was created only about ten years after the death of Elizabeth I, Henry VIII's daughter; for more on this see e.g. Boyce, 1996:289 or Greenblatt et al., 1997:12-17.) Furthermore, Hillman's (1997:73-74) discussions of the effects of "secularization" on EME drama suggest that early modern audiences would have been starting to reflect on the events leading up to the Reformation in light of increasing human authority set against the diminishing authority of religion.

The final example in this series is given below.

| (41) | Rosalind [to Celia]: | Do you not know <u>I am a woman</u>? when I think, I must speak. |

As You Like It, III:ii

In example 41, Rosalind voices the idea that women can't keep their thoughts to themselves. Though this is clearly humorous, especially from a 21st century perspective, there is an underlying seriousness in the fact that relatively powerless 16th century women lived with the consequences of the socio-cultural assumption that a tendency to gossip and slander was an undesirable but inherent female quality (argued by Habermann, 2003, see S.4.2.2). For an early modern audience, humour may have arisen from the irony of a male actor playing the role of Rosalind (see S.2.5). Jardine (1983:9) argues that although this was usual at the time, the all-male casting of female roles would have affected the audience's interpretation of the plays.

My analysis of examples 39 to 41 has shown that the ways in which female characters identify themselves as women acquire further layers of meaning when filtered through an audience's likely knowledge of history, their political and religious beliefs, and their opinions about relationships between men and women – in other words, through their schemas of real life, as argued by Culpeper (2001:47-112) (see S.2.1). My analysis supports Findlay's (1999:106-108) claim that sexual and gender identities in the plays are "multilayered". There are various ways of interpreting them, as mentioned above. Though the words of the three women in examples 39 to 41 seem to reduce the status of womanhood, the outcomes for them in the plays do not necessarily show that being a woman disadvantages them. Portia gets her husband to do what she wants, Katherine's professed disempowerment in the face of men helps characterise her as virtuous, and Rosalind's remark makes her seem witty. Again, Findlay's

arguments with regard to the representation of women from a feminist critical literary

perspective seem borne out by my findings, since they support the claim that women

in the plays resort to "feminine ingenuity" to overcome the imbalance in the relative

social positions of men and women (Findlay, 1999:98, 131). Examples 39 to 40 all

take place in public contexts, and there is a clear contrast in the ways these women

talk about their sex as inferior and the private self-presentation of superior

womanhood in example 19 (from the Duchess of Gloucester's dialogue, discussed in

S.4.3.2). As stated in S.4.3.2, Culpeper (2000:169-171) argues that private self-

presentation is more likely to show the characters' true feelings than public self-

presentation, which is "inevitably oriented towards others" (ibid.:168). My findings

indicate that women in the plays are only able to voice pride in their womanhood

when in private. In public, women are shown as casting their sex in an inferior light, in

order to get what they want or be viewed positively, according to the outcomes in the

drama.

In contrast to the female characters, my data shows that male characters saying "I am a

man" never put themselves down or acknowledge any disadvantage through

identifying themselves as men. There is less to say about the three examples from the

male dialogue, so I present these below then discuss them together.

(42)	Lear [to Cordelia and Doctor]:	For, as <u>I am a man</u>, I think this lady To be my child Cordelia.
		King Lear, IV:vii
(43)	Ferdinand [to Prospero and Miranda]:	No, as <u>I am a man</u>.
		The Tempest, I:ii
(44)	Macbeth [to Lady Macbeth and Lords]:	[Banquo's ghost disappears] Why, so; being gone, <u>I am a man</u> again.
		Macbeth, III:iv

As with the examples of 'I am a woman', above, examples 42 to 44 all take place in public, mixed-sex contexts. In example 42, King Lear boosts the strength of his claim that he recognises his daughter Cordelia by saying he knows it as well as he knows he is a man. Similarly, in example 43 Ferdinand counters Prospero's accusation that he is a spy, identifying himself as a man to strengthen his argument. Both Lear and Ferdinand seek to improve their images by referring to their biological sex, as does Macbeth, in example 44. Macbeth has been deeply disturbed and frightened by the presence of Banquo's ghost, behaviour which he now implies is unmanly, since he becomes "a man again" after the ghost has gone. He says this to make himself look better in the eyes of the assembled company (his wife and the members of the court).

Examples 42 to 44 show that male characters' self-presentation as men is much less complex than the self-presentation of female characters as women, and there seem to be fewer possible layers of interpretation. Unlike women in the plays, men boost their image by publicly referring to their sex. That Shakespeare presents the explicit self-construction of gender in such contrasting ways in public suggests he is pointing out different socially-sanctioned expectations of male and female behaviour, i.e. the contrast in value which they may publicly place upon their sex.

The discussion of manhood and womanhood and the way it was represented in the plays is not new (see e.g. Hillman, 1997:264-265), but again my data adds some empirically-based findings. The private onstage claim to superior femininity made by the Duchess of Gloucester (example 19, S.4.3.2) is of course heard by the audience at the higher discourse level (see S.2.1). It is therefore clear that the female characters who publicly claim the inferiority of womanhood in the plays do so not because they

believe it to be true, but because it is the only socially-approved route available to them in furthering their aims. This is evidenced by the women in examples 39 to 41 deriving benefit one way or another from the negative evaluation of womanhood. My findings here support Findlay's (1999:98) claim that women in the plays stealthily manipulate the power balance to overcome their inferior social position. This case study has shown that through using language which overtly disempowers their own sex, women covertly empower themselves – an irony which fits with Culpeper's (2000:312) view that Shakespeare's work is part of contemporary debate on women in society (in S.1.1).

To sum up the Topical clusters in my results, though many arise because of what the play concerns (such as the concentration of People clusters in the male history dialogue), others have revealed more about the ways male and female characters talk about themselves and each other. Contrasts between the positively and negatively key People clusters show that women talk significantly more about men in personal, intimate ways which help flesh out the personalities and histories of male characters, whereas men are more likely to communicate the actions of other male characters in the plays. The notable concentrations of States clusters beginning with *I* in female comedy dialogue suggest that self-revelation contributes to what the audience needs to know to understand the characters' behaviour and motivations. As with certain types of questions (S.4.4.1), this dramatic device seems to operate more in the dialogue of female characters than male characters. Finally, the investigation of the States cluster 'I am a woman' has illuminated an important contrast in the ways men and women characterise their own sexes through public self-presentation, and confirms some irony in the portrayal of women negatively evaluating their sex to achieve some advantage.

Despite the absence of a set of key results from the aggregated male data file, I was nevertheless able to make the series of connections in the empirical data which led to this finding. This is because it depended more on careful qualitative analysis of the context and co-text of the cluster 'I am a woman' and its counterpart formula "I am a man" than on a direct statistical comparison between them.

The final results to be discussed (very briefly) are Ideational clusters with a Circumstantial function.

4.5.2   Circumstantial clusters

There were no clusters with Circumstantial functions in the aggregated female data (Table 4, S.4.1) and only one in the female dialogue in the sex/genre breakdown (Table 7, S.4.1). These are more a feature of male dialogue, particularly in histories and tragedies, where a number of clusters concerning places occur. Not surprisingly, these arise because of the locations which characters move between or make reference to during the course of the plays, so they are topical in the sense of being local to what the dialogue is about, and therefore of low interest (as explained in Ss. 4.2 and 4.5.1 above). They are classified as Circumstantial rather than Topical because their function is primarily deictic rather than descriptive, i.e. they indicate relative locations of people and activity in the play, rather than contributing to an impression of the location itself. I could not determine a reason for the keyness of Circumstantial clusters in the male tragedy and history dialogue (shown in Table 7), and it may well be due simply to the higher proportions of male characters in these genres (see S.3.2.3).

This completes my analysis of the key cluster results obtained from my Shakespeare corpus. In the last part in this chapter, I discuss the overall implications of my findings.

4.6   Discussion of the main findings

I will now discuss what the analyses in Ss. 4.3 to 4.5 add up to when compared and considered all together. The results revealed some general trends in female dialogue overall in the plays and some contrastive trends in male and female data. From these trends, it was possible to identify two types of effects in the play-texts. I noted this first at the end of S.4.3.1, where I distinguished between "global" effects and "individual" effects:

- global effects are overall impressions created by the style of language habitually used by women (and men) in the plays, based on data from the dialogue of all characters in the corpus of 37 plays;

- individual effects arise when one or a few characters use a cluster in a way which deviates from the majority of other instances in the play.

For example, a cluster used only by women in the plays creates a global effect. Against this, one male character making use of that cluster appears unusual, creating an individual effect (see e.g. S.4.4.1, example 27).

Since they all characterise people in plays, both global and individual effects can be considered stylistic in the broad sense. The distinction is that:

- global effects characterise members of a whole social category (male or female gender) as typically behaving in certain ways; and

- individual effects characterise the personalities and behaviour of one or more individual men and/or women in the plays.

However, trends which I argue as creating global effects are unlikely to have been consciously applied for stylistic effect by the playwright. Instead, they arise when we look at the ways Shakespeare – a male writer – represents men and women as tending to use language in different ways. Global effects are therefore discountable in Kelley's (1972) terms (discussed in S.2.1), although they are relevant for their historical sociolinguistic interest in the ways habitual female and male speech is portrayed at the time. They may therefore be related to language variation in EME, as discussed further in S.4.6.1. I discuss the outcomes of specific global and individual effects in S.4.6.2.

Aside from global and individual effects, when the plays were considered as instruments for transmitting a story to an audience, i.e. in terms of the discourse levels (Short, 1996:169-172) discussed in S.2.1, it was clear that the female dialogue contributes to or carries certain types of dramatic device more than the male dialogue. Effectively, female characters help tell the tales in different ways to male characters, through the different kinds of formulaic language which surface in the key clusters. I discuss this further in S.4.6.3.

4.6.1   Trends and variation in the language of female and male characters

On the whole, my investigations did not show vast differences in the ways men and women are represented as speaking in the plays. Instead, some subtle trends emerged

157

through the global effects, creating impressions of habitual language behaviour for men and women under "normal" circumstances in the plays (see S.4.4.1). It was impossible to say to what extent these tendencies accurately reflect wider EME language conventions, without supporting evidence from historical sociolinguistic studies taking in other genres of speech-like texts (much as a television soap opera today reflects only a partial snapshot of PDE language conventions and social behaviour). However, I would suggest that general trends established as norms in the canon of plays would be likely to fit in with the schematic assumptions of an early modern audience about people of the sex, age, social rank and role, occupation and other social circumstances ascribed to the characters in the plays. If they did not, the audience would not have had a reference point from which to understand the occasional departures from these norms which clearly have a stylistic purpose. The individual effects I identified – stylistically unusual uses which have a dramatic purpose – would have only worked against a background of norms which the audience shared and understood.

Whilst there was insufficient evidence to say whether or not the trends in female (and male) language hold true only for Shakespeare's plays or more widely, there was some useful supporting data from historical linguistics studies for the Interpersonal: Speech-act related: Directive cluster 'I pray you' – by far the most key result in the female data (see the case study in S.4.3.1). Based on the research of Culpeper and Archer (2008) and Busse (2002), amongst others, my results showed women in the plays making much more use of 'I pray you' as a polite request formula, i.e. tending to frame requests as supplications, which would be consistent with the lower power and social status of women than men in actual social life at that time. This was further

158

supported by the fact that over two thirds of the instances of 'I pray you' in my female data were addressed to male characters. Furthermore, my detailed analysis of an excerpt from *The Merry Wives of Windsor* (example 8, S.4.3.1) showed the way a female character is represented as using *pray* formulae more often than a male character, despite being in a situation where there was a fairly equal social power balance, if anything weighted in favour of the woman.

Even so, as pointed out in S.4.3.1, I could not discount the possibility that the keyness of 'I pray you' was influenced by the number of requests in the female dialogue in the plays to some extent, since assessing this was well beyond the scope of my study. The view that women make more requests than men, and would consequently be more likely to use 'I pray you', was supported by:

(i)   the presence of a Speech-act related: Thanking cluster in the female dialogue (S.4.3.1); and

(ii)   the significant presence of clusters which were fragments of questions (S.4.4.1).

However, the keyness of 'I pray you' was so comparatively high that I felt the number of requests would not totally explain it. Furthermore, some women who make a lot of requests, e.g. Lady Macbeth and Cleopatra, do not tend to use 'I pray you', which I argued makes them seem more like men – i.e. relatively less polite – in their speech habits (in S.4.3.1). Although this result was interesting, as I pointed out, there are many *pray* formulae (see e.g. Busse, 2002 and Blake, 2002), all of which would need

159

investigating in order to make any firm claims about sex-based variation in politeness/deference and requests in EME.

Just as 'I pray you' indicated a trend in women being more polite in making requests, the keyness of the Textual: Discoursal: Question cluster 'what is your will' (and 'what is your' followed by "pleasure", which formed a similar formula) indicated a trend in women being more submissive in responding to summonses and greetings (S.4.4.1). My data indicated that this only applied to interlocutors in certain social roles, however, and that wives did not typically respond to husbands with these formulae. Hence, several exceptions stood out, creating interesting individual effects. Unfortunately though, I found no historical sources mentioning these particular responses. This was also the case for the Interpersonal: Speech-act related: Assertion cluster 'I know thee', which appeared from my data to be conventional in male in-group language, particularly in insulting strategies, but not something which was usually said by or towards women (S.4.3.1).

To sum up the possible links between trends in my data and actual language variation in EME, I can say that the evidence from the clusters 'I pray you', 'what is your will [pleasure]' and 'I know thee' are **suggestive** of wider language use which was conventional among women and men at the time, **although not proof**. As anticipated in S.2.1, in analysing the results it has been useful to bear in mind what the schematic assumptions of a 16th or 17th century audience might have been when considering what was and was not stylistically unusual in the data, bearing in mind changes in social behaviour governing language use since the plays were written. It must also be remembered that despite Shakespeare's undoubted skill in constructing characters of

160

different gender, his own position as a male writer and his own schematic assumptions about male and female language behaviour would have influenced the variation he built into their dialogue (noted in S.2.5).

Other results in my data were more clearly related to individual stylistic effects, or to dramatic devices particularly located in the female dialogue, than to possible links with wider EME language variation, so I summarise these in Ss. 4.6.2 and 4.6.3. The most important aspect of the trends which may have a historical sociolinguistic explanation is that they would not have stood out to an early modern audience. Only the unusual uses would have stood out, when backgrounded against schematic assumptions of normal social behaviour, and it is these I address in the next section in my discussions of global and individual effects. Quite possibly neither the trends nor the exceptions to them would stand out to many present-day audience members, most of whom are likely to have a schema for historical drama and/or Shakespearean characters which comfortably bridges the gap between the language of women and men then and now. Additionally, of course, no audience is ever presented with any comparisons from other plays during the course of a performance, and is therefore never confronted with the trends which became clear in my electronic analysis of the canon of plays.

4.6.2   Global and individual effects in the male and female dialogue

Thus far I have built up the following argument:

- global effects are the habitual uses indicated through the language patterns of men and women in the plays, some of which might also approximate wider usage in EME as represented by Shakespeare; and

- individual effects stand out by breaching the norms of the global effects in some way, thereby alerting the audience to something particular about the personality or behaviour of individual characters.

I now discuss the specific global and individual effects which could be seen in the results analysed in Ss 4.3 to 4.5. I summarised the case of 'I pray you' in the previous section, so I do not do so again here except to reiterate that the greater use of it by women in the plays is also a global effect, i.e.

(i)   it is a language formula that is more characteristic of women than men, for reasons that are apparently associated with their lower social status; and

(ii)   the dialogue of women who don't tend to use it (e.g. Cleopatra and Lady Macbeth) therefore seems constructed to be more like that of men.

As mentioned in S.4.6.1 above, the Textual: Discoursal: Question cluster 'what is your will' and the related formula "what is your pleasure" in S.4.4.1 were established as polite responses made by women to a summons or greeting, from the majority of instances in the data. This created a global effect of positioning women as needing to appear more obedient, which would be consistent with their lower power status in early modern social life (as was the case with 'I pray you'). In S.4.4.1 I pointed out that wives in the plays did not generally respond to their husbands with 'what is your will [pleasure]', however, and I showed that Desdemona (example 26) and Katharina

162

(example 25) were unusual cases who did. I argued that this created an individual effect of exceptional obedience in both these characters which was instrumental in positioning Desdemona as a tragic victim, and Katharina as having radically switched from a position of extreme shrewishness to extreme obedience (which might be interpreted as genuine or ironic). I also highlighted an interesting individual effect created by Iago's use of "what is your pleasure" to Desdemona (example 27). I argued that since it is an expression used otherwise by women in the plays, Iago seems deviant and yet also clever in employing it in his efforts to subvert the marriage of Desdemona and Othello.

The data and supporting evidence discussed in S.4.3.1 established 'I warrant you' as an Interpersonal: Speech-act related: Assertion cluster which characterises informal speech, particularly in comedies. Against this global effect, an individual effect of characterisation was shown in the excessive use of 'I warrant you' by Mistress Quickly (example 9), when compared to other characters (both male and female). I argued that this contributed to an impression of her as rather garrulous, and lacking awareness of appropriate behaviour for her relatively low social position of housekeeper. The recurrence of 'I warrant you' in Mistress Quickly's dialogue almost certainly skewed the results for this cluster, inflating its presence in the female comedy dialogue. Therefore, I would argue that it is particularly associated with one character in the plays overall who happens to be female, but not otherwise associated with one sex more than the other. Since it is a colloquial expression (cf. Culpeper, 2001:201), it is more likely to be a marker of (lower) social rank than gender. Bach (2007:33), mentioned in S.2.6, is one of a number of scholars who suggest that social rank is actually more crucial to language difference than gender in Shakespeare's plays.

163

In contrast 'I know thee', the other Assertion cluster discussed in S.4.3.1 and mentioned in S.4.6.1, was shown to be much more typical of male speech to other male characters, and often (but not always) part of an insulting strategy. The global effect is that it appears to be socially appropriate in male in-group language, but not in language used by and to women. I stated that the individual effect of the single use of 'I know thee' to a woman, by Titus to Tamora (example 10), seemed to be the linguistic negation of Tamora's very status of being a woman, and therefore the construction of her as male. I argued that Tamora is marked out as exceptional by being addressed using this formulaic Assertion, which mirrors an exceptionally unwomanly act among women in the plays (organising the rape of another woman).

The analysis of 'I know thee' showcased the considerable potential for discovering stylistic effects hitherto hidden by the passage of time and changes in language and social behaviour. Since 'I know thee' is an archaic language form, its limitation to mainly male in-group language (as indicated by my data) would not be apparent to a present-day audience or reader. Its general semantic meaning is still clear, but the pragmatic effect – the insult conveyed by using it to a woman – is not. This may arguably not matter much, since Titus's feelings toward Tamora are already clear in the play, and would be made more so by non-verbal communication by the actor in a performance (I mention aspects of performance in S.1.2). However, it is interesting for what it reveals about the way Shakespeare apparently harnessed sex-restricted language to convey an appropriately singular degree of hatred from one character to another.

Several clusters in the data simply created global effects of contrasts in the ways male and female gender is constructed by the dialogue, without any individual effects associated with particular characters. The Textual: Narrative-related clusters (in S.4.4.2) revealed that women in the plays more often make claims based on what they have heard, and men on what they have seen. The women's claims were much more often disputed by interlocutors, although neither type of allegation seemed inherently more reliable. The global effect is essentially that men and women typically use different formulae to express their arguments, with women more likely to be judged as unreliable. However, this seems most likely to arise from recurrent stereotyping in the plays, rather than from an approximation of conventional variation in male/female language behaviour. It may well allude to the lower social power of women, which would have rendered them less able to question the claims of men, and more likely to have their arguments overridden. However, from my data it was impossible to say whether or not this recurrent contrast in representation would have seemed ironic to an early modern audience.

There was much clearer irony evident in the analysis of the Ideational: Topical: States cluster 'I am a woman' and its counterpart formula "I am a man", in S.4.5.1. Here, I showed that in public contexts, women in the plays claim their own sex as inferior and men claim their own sex as superior: two contrasting global effects. I suggested that Shakespeare presents female characters as unable to boost their image or strengthen their arguments through a positive evaluation of their womanhood in public, especially where they are trying to achieve some end; instead they are shown as resorting to a strategy of claiming inferiority. It is worth pointing out that female characters do not all share the same motivations and personalities, but a general insincerity in their

claims to inferior womanhood is implied by the advantages they gain in the plays through making them (either tactical or image-enhancing). This was also supported by the individual effect of the Duchess of Gloucester's private claim to female superiority (in example 19, S.4.3.2) which is heard only by the audience. The evidence from this analysis of the ways men and women explicitly self-construct gender in the plays is suggestive of Shakespeare's awareness of an inequality in the way women could or should present themselves in public, to which he alludes through the dialogue of his characters. This supports the general consensus among linguists (e.g. Culpeper, 2000:312) and feminist literary critics (e.g. Findlay, 1999:98) that, through the language and behaviour of his characters, Shakespeare involved himself in early modern discussions of women's place and power in social life (see S.2.5).

With more space, I have no doubt that further global and individual effects would have been identifiable through the corpus results, and more stylistic effects uncovered which otherwise remain hidden by changes in language and social behaviour over time. Those among my analyses in Ss. 4.3 to 4.5 clearly demonstrate that there is still more to be discovered about Shakespeare's plays through the careful qualitative analysis of electronically-derived quantitative data.

I stated at the start of S.4.6 that in addition to what my findings reveal about the ways women and men are characterised as social groups and individuals, the role of male and female dialogue in actually communicating the drama successfully to the audience is also shown to be different in some ways. Certain types of dramatic device are more associated with the female dialogue than the male dialogue, and I discuss these in the next section.

166

4.6.3   Dramatic devices operating more in the female dialogue

In S.4.4.1 I argued that my results showed female characters making greater use of certain types of questions than male characters in the plays, so much so that my functional category of Textual: Discoursal: Question clusters contained the highest number of key results in the female dialogue overall. The examples I discussed demonstrated that questions posed by female characters work in various ways to make information available to the audience in order to help them understand and enjoy the play, including:

- revealing information about other, offstage characters, which affords the audience privileged knowledge (example 21);

- providing opportunities for characters to voice their inner thoughts, which explain the background to their motivations and actions (example 22);

- helping to cue the behaviour of other characters to the actors playing them, simultaneously emphasising it to the audience (example 23); and

- enabling the action to move on from one point to another (example 24).

I stressed that an examination of my sub-corpus containing the male dialogue confirmed that male characters use questions in similar ways, and I argued that it was therefore the greater prevalence in the female dialogue that was of special interest. As I pointed out, I could not discount the possibility that it may be linked to a greater proportion of requests in female dialogue (also mentioned in S.4.6.1 above).

As well as making more use of questions, female characters made much greater use of a strategy of expressing negative volition (saying what they don't want), shown in a case study of the Speech-act related: Modalizing: Volition cluster 'I will not' in S.4.3.2. I argued that although it mainly conveys desires which conflict with those of another character, the impression this creates very much depends on the situational context. I pointed out a recurrent situation type in the plays in which female characters respond with 'I will not' when told to be quiet or to leave or stay by another character. The impression created by the woman's opposition depends upon the circumstances and particularly upon the character(s) with whom the woman is in conflict. I showed, for example, that Emilia appears brave in resisting Iago's instructions to be quiet, because the audience knows that he is an evil character and Emilia is on the side of justice (example 14). I also showed that women sometimes use 'I will not' to express inner personal conflict, i.e. it is a vehicle for self-presentation. I argued that some characters use it to counter some implied norms about female behaviour (e.g. the assumption that women are likely to gossip, cf. Habermann, 2003). These implied norms may represent Shakespeare's anticipation of the schematic beliefs of an early modern audience, and his awareness of gender stereotypes. If so, this is another way in which he brought these into focus through the dialogue in the plays, as I also argued above in S.4.6.2 with regard to the way female characters evaluate their biological sex in public contexts. This seems highly likely, but as with any play-text written for performance so long ago, it is not possible to say for certain exactly what nuances of irony the dramatist was intending to achieve.

The Textual: Organisational: Informational elaboration clusters in S.4.4.3 revealed another way of foregrounding what the audience knows or should have realised in

order to understand the play. Here, I argued that female characters talk significantly more than men about teaching and learning, and in so doing they voice what the audience has actually learned. This is a way in which Shakespeare communicates information to the audience at a discourse level above that between the onstage characters (cf. Short, 1996:169-172 discussed in S.2.1). The emphasis on character behaviour, the revelation of characters' inner thoughts prompted by questions, and the self-revelation channel 'I will not' (all mentioned above) also work by virtue of the playwright-audience discourse level.

The final dramatic device which my results showed as more prevalent in female dialogue was self-presentation through the Ideational: Topical: States clusters starting with *I* (S.4.5.1). This method of revealing information about themselves is particularly important among female characters in the comedy genre. In S.4.5.1 I also argued that women contribute to the characterisation of men in the plays through other-presentation, shown by my analysis of the Ideational: Topical: People clusters. The concordance data revealed that women talk significantly about men in intimate and personal ways which cast light on the characters of the men, rather than on the men's actions in the play. Other results from the Interpersonal: Speech-act related: Vocative clusters and Ideational: Topical: People clusters in the male history dialogue showed that men also talk significantly more about men. However, unlike the results from the female data, this is largely because they describe the actions of men in the plays. Although these results were of less stylistic interest, the contrast in the ways both female and male characters talk about men in the plays is certainly noteworthy: the female dialogue helps characterise the men, who are often the subject of the plays.

There are some qualifying comments to make about the key clusters which were evidently dramatic devices associated more with female dialogue than male dialogue summarised above. The results in Table 5 (S.4.1) showed the Interpersonal: Modalizing: Volition cluster 'I will not' as being concentrated more in the histories and tragedies than in the comedies, unlike the Textual: Discoursal: Question clusters which were located more in the comedies. Both, however, were shown by my analyses as having multiple potential effects which depend upon the requirements of the situational context of the play, rather than as creating an overall impression (i.e. a global effect) of women in the plays or in a particular genre.

I cannot be certain whether the links between global effects and genres suggested by the concentrations of key results tabled in S.4.1 are genuine, or whether these were artificially influenced to some extent by the different ratios of female to male characters in each genre (see S.3.2.3). The significance of female self-presentation through the *I* clusters (S.4.5.1) may be influenced by the relatively greater numbers of female characters in comedies compared to histories and tragedies which, as explained in S.3.2.3, could increase the likelihood that clusters would occur as key in that genre. The same linguistic forms may be present in the non-comedy genres but in insufficient numbers to stand out statistically, i.e. to occur as key clusters, when compared with the male dialogue. This remains a limitation of using all the dialogue in the plays as the basis for my results. However, it is far outweighed by the benefits which have been shown through the identification of global and individual effects, made possible by establishing majority usage of a language form across all the plays against which unusual minority uses can be usefully compared (as discussed in Ss.4.6.1 and 4.6.2 above). The nature of comedy plots arguably requiring more explanation to the

audience than those in histories or tragedies (mentioned in S.4.5.1) may also help explain the presence of dramatic devices which are particularly concentrated in comedies.

My study shows that the ways of communicating a play to the audience which are more significant in female dialogue are clearly important dramatic tactics woven into the language of women in the plays in specific ways. These are:

- other-presentation of men;

- self-presentation (in comedies);

- questions eliciting information or catalysing actions; and

- opposition to other characters or to inner thoughts.

They would, of course, be designed to go unnoticed as such by the audience whilst watching the play. It is therefore reasonable to assume that, like the global effects (S.4.6.2), they would have largely coincided with typical schematic assumptions of an early modern audience about the ways women use language, although then as now not all audience members would have had identical schemas of language and social behaviour (Short, 1996:231; Culpeper, 2001:68, see S.2.1). Only when the plays are examined all together does an overall picture emerge from which it is possible to discern differences in the ways female and male dialogue helps to achieve successful communication of the play.

It is important to make clear that my study does **not** show that female dialogue in Shakespeare's plays conveys the plot more than male dialogue. There is vastly more

male dialogue in the plays than female dialogue, and other dramatic devices for explaining the plot (e.g. self-spoken dialogue, and quite possibly other types of questions which do not occur as key clusters in my results), are likely to have a greater presence in the male dialogue. Whilst some of the patterns of results in S.4.1 are clearly related to their function of communicating the play, rather than to characterisation of groups or individuals within it, these are still valuable findings of the study:

(i)   for the norms of women's language behaviour they hint at; and

(ii)   for what they add to knowledge about the construction and communication of Shakespeare's plays.

In this section, I have put my results and their analysis (earlier in this chapter) into perspective, and I have made clear what can and cannot be claimed based upon them. This means I can now draw some conclusions about the success of the study and what has been learned from it, in my final chapter.

5.   FINALE: SUMMARY AND CONCLUSIONS

My investigation of a corpus utilising all the dialogue in the plays has provided an overall picture of what the formulaic language of female characters is like, based on statistical frequency of key word clusters in comparison to those used by male characters. In S.5.1 below I summarise the main findings and outcomes of the study (briefly, since these are detailed more fully in S.4.6). In S.5.2 I evaluate how well the methodology has served my research aims, and finally, in S.5.3, I offer some suggestions for further research.

5.1   Summary of the main findings

Through the analysis and discussion of results in Chapter 4, my study has revealed:

(i)   some ways in which Shakespeare styled men and women as habitually using language differently;

(ii)   many individual stylistic effects which contribute to the construction of particular characters in the plays (discussed in S.4.6.2); and

(iii)   that certain types of dramatic device operate more significantly in the female dialogue (discussed in S.4.6.3), affording new insight into the ways the plays are actually constructed and communicated to the audience or reader.

The following general trends in the female dialogue in Shakespeare's plays were found:

- women use the polite request formula and discourse marker 'I pray you' much more than men (S.4.3.1);

- women talk more about teaching and learning than men (S.4.4.3);

- women make greater use of certain types of questions than men (S.4.4.1); and

- women make greater use of negative volition (S.4.3.2).

Some direct contrasts between the male and female dialogue were also evident, notably that in public contexts female characters voice negative evaluations of being women, whereas male characters voice positive evaluations of being men (S.4.5.1).

Two types of effects surfaced through the analysis of results, which I termed "global" and "individual" effects (in S.4.6):

- global effects are the ways in which women and men are generally represented as using language through the majority of instances of a cluster; and

- individual effects are the unusual instances of a cluster when used by one or a minority of characters in a way which is different from the rest.

For example, 'what is your will [pleasure]' (S.4.4.1) is a response habitually made by women to certain social superiors (global effect); it is not usually used by women to their husbands, so an individual effect is created by its occurrence in the dialogue of Katharina and Desdemona (examples 25 and 26), foregrounding them as exceptionally obedient wives. Another individual effect is created when Iago uses "what is your pleasure" to Desdemona (example 27), making him appear deviant but wily.

The global effects can be said to represent norms of use within the canon of plays, but not necessarily in wider EME, since there is insufficient historical sociolinguistic evidence to support such claims (discussed in S.4.6.1). Global effects are of value for what they indicate about Shakespeare's representation of variation between typical male and female language behaviour, and the ways this constructs gender in the plays (for example, the greater use of the polite request formula 'I pray you' among women, see S.4.3.1).

Evidence from the results indicates that some types of dramatic device for communicating the plot to the audience are more prevalent in female dialogue (though not exclusive to female dialogue). This was shown:

(i)   in the multiple elicitation effects which questions asked by women have in the dialogue (S.4.4.1); and

(ii)   in the way discourses of teaching and learning highlight what the audience has also learned (in the organisational elaboration clusters, S.4.4.3).

Self-presentation and other-presentation in female dialogue was also argued as helping to:

- explain the plot;
- reveal character background and motivation; and
- emphasise important onstage behaviour.

This was demonstrated:

(i)   by the ways women talk to men (e.g. Lady Macbeth asking Macbeth about

        his facial appearance, in S.4.4.1, example 23);

(ii)   by the way they talk intimately about men (shown by the Ideational:

        Topical: States clusters, S.4.5.1); and

(iii)   by women expressing what they don't want (through the Interpersonal:

        Modalizing: Volition cluster 'I will not', S.4.3.2).

However, it is likely that other types of dramatic device are more prevalent in male

dialogue (e.g. the use of soliloquy).

In this section I have set out briefly what my investigation of the formulaic language

(shown in the key cluster results) has added to knowledge about the language in the

dialogue of men and women in Shakespeare's plays. A great deal more could no doubt

have been revealed through further qualitative analysis of the quantitative results, if

space had allowed. It is worth emphasising that Shakespeare's plays feature a rich

diversity of characters, both male and female, rather than two distinctive sex- or

gender-based groups whose members all behave the same way. Many other factors

influence the way men and women in the plays behave, for example the social ranks

and roles accorded to them, and their individual motivations.

5.2   Reflections on the success of the study

Having drawn together the main outcomes in S.5.1, I will now briefly recap the aims

of the study and discuss to what extent my research questions in S.1.4 have been

answered. The overarching goals of my study, as stated in S.1.2, were to help address

the lack of detailed corpus-based research comparing the language of women and men

in Shakespeare's plays, and to contribute some evidence about the representation of female characters based on quantitative data to the vast body of (mainly qualitative) existing commentary.

My first research question (1.4.1) concerned the identification of key word clusters in the dialogue of male and female characters in Shakespeare's plays, and was addressed successfully apart from the absence of results from an aggregated data file of male dialogue. As explained in S.3.3.1, this was due to the over-representation of male dialogue in the corpus, making it too similar to the reference corpus of all plays for any clusters to occur with statistical significance. This was of course an inherent problem in using all the dialogue, which would occur in any corpus of all Shakespeare's plays, not only the one I adapted from Mike Scott's corpus (see S.3.1). However, as discussed in S.3.3.1 and shown in S.4.1, a set of results for the aggregated female dialogue (Table 4) was successfully obtained, together with comparative sets of results for male and female dialogue when the corpus was broken down into three genres of play (Tables 5 to 8). From these I was able to set the female dialogue into the context of the male dialogue, which suited the orientation of my study towards women in the plays. In a larger study, both my male and female sub-corpora of Shakespearean dialogue could be compared to a reference corpus of other contemporary texts (see S.5.3 below), which would almost certainly be sufficiently different to produce key results for both the male and female datasets.

Research questions 1.4.2 and 1.4.3 concerned the categorisation of functions in the key cluster results, and the subsequent identification of patterns which would usefully lead to stylistic analysis of the male and female dialogue. The set of functional categories I

177

used (Table 3, S.3.4.2) successfully allowed me to deal with a range of grammatical, lexical, semantic, pragmatic and discoursal aspects of my data, all of which came through in my analyses in Chapter 4. As explained in S.3.4.1, I was able to adapt Culpeper and Kytö's (forthcoming) functional classification system, based on Halliday's (e.g. 1994:179) interpersonal, textual and ideational metafunctions of language. Adapting a suitable existing system afforded me more time to spend on the close analysis of the results in context, leading to findings which could be linked to three major areas:

(i)   historical sociolinguistic variation (discussed in S.4.6.1);

(ii)   the stylistics of characterisation: of women and men in general, and as individuals in the plays (S.4.6.2);

(iii)   the operational role of dialogue in communicating the play to the audience via certain kinds of dramatic device, via the female characters (S.4.6.3).

Much of this would probably not have been feasible if I had attempted to design a completely new system of functional categories; this would in itself have had to be a major outcome of the study.

Many of my "individual" effects (see S.4.6.2 and the summary in S.5.1 above) are actually the result of local functions of key clusters. In other words, their functions depend on local contextual factors such as relative power of speaker and addressee, their social positions, the influence of who else is co-present, etc. On reflection, it is therefore worth considering whether adopting a set of local functional categories (as was done by Mahlberg, 2007a and 2007b, mentioned in Ss.2.3 and 3.4.1), would in

178

fact have been better. I could have applied the principles but used different category labels to suit my dramatic dialogue data. However, a bottom-up approach beginning with local clusters would have made it difficult to capture trends in the ways larger groups of characters (e.g. belonging to one sex, or another social category such as rank or class) use language. My top-down approach, on the other hand, allowed me to start by identifying group trends in the functions of my results, then drill down to local effects which counter the trends in the speech of one or more individuals (the global and local effects discussed in S.4.6 and mentioned in S.5.1 above). Effectively, I was able:

(i)   to establish some norms of language use across the fictional worlds created by Shakespeare;

(ii)   to place unusual cases into this context; and

(iii)   thereby to discuss how a character behaves in relation to other characters in the playwright's universe, not just how s/he behaves in relation to the audience or reader's (probable) schematic expectations.

This is a particularly useful approach to historical drama, where not all the information about the social context of the time is available, and helps to avoid the temptation to apply 21st century politeness norms and social conventions in an inappropriate way (a risk noted in S.2.6). The disadvantage of my top-down approach is that not all interesting local stylistic effects will be numerous enough to occur as key, and some may go unnoticed in a study based on statistical frequency of occurrence in all the dialogue in a literary work or set of works.

It is also worth pausing briefly to ask if categorising the results in a formal way was necessary at all, and whether in fact the outcomes mentioned above could have been arrived at by simply examining the results and the *WordSmith* concordance data and seeing how they grouped together. Many of the sub-categories in the functional categorisation system I used were sparsely populated (see Tables 4 to 8, S.4.1), a concern highlighted in S.3.4.1. This means the data in them might reflect occasional functions specific to a contextual situation, rather than functions which feature prominently in Shakespearean male or female dialogue per se. However, I do not feel this matters, as the functional categories are a means to an end (tracing the characterisation effects which led to the outcomes mentioned above), not an end in themselves. In other words, my conclusions are not based on the strength of the presence of any particular functional categories or sub-categories, but on what they enabled me to find out about the dialogue of individuals and groups of characters (in Chapter 4). I would probably have achieved some of this without applying the functional classification system (e.g. the keyness of 'I pray you' would still have been apparent, leading to the discussions in S.4.3.1), but by no means all of it. For example, without the sub-sub-category of Interpersonal: Speech-act related: Assertion I doubt that I would have uncovered the interesting isolated instance of 'I know thee' used towards a female character (in S.4.3.1). I am therefore satisfied that a formal framework of functional categories with a solid theoretical basis led to a more systematic analysis, and was essential to the study.

I now turn to the question of how successful the examination of formulaic language (as it occurs in the form of key word clusters) has been with regard to learning more about the characterisation of women (and men) in the plays (research question 1.4.4). In Ss. 2.2 to 2.4 I presented the case for examining RWCs in language, based on

claims that language is "formulaic", i.e. stored and used at the lexical level in units outside of grammatical structures (cf. Sinclair, 1996, 1991, 2004; Wray, 2002; Hoey, 2005). Citing many studies (e.g. Stubbs, 2005; Biber et. al., 1999; Scott and Tribble, 2006; Culpeper and Kytö, forthcoming), I argued that the analysis of RWCs is now a well-established approach in corpus linguistics, and suitable for early modern texts thanks to the advantage of new technology to overcome the problem of spelling variation (*VARD*, see S.2.5).

The validity of this approach has been borne out by my results and discussions in Chapter 4. These showed numerous formulaic language strategies, e.g.:

- 'I pray you' as a polite discourse marker, in S.4.3.1;

- 'what is your will [pleasure]' as a response to summonses, in S.4.4.1;

- the self-construction of gender through 'I am a woman' and its counterpart "I am a man" in the male dialogue, in S.4.5.1.

As summarised in S.5.1, the analyses of these revealed:

(i)   possible links to variation in EME (see S.4.6.1);

(ii)   contrasts in the language men and women are represented as using typically (the global effects, see S.4.6.2);

(iii) many interesting individual effects of characterisation (S.4.6.2); and

(iv)   some evidence of the loading of female dialogue with particular strategies for communicating the play to the audience (S.4.6.3).

For a first foray into the male and female dialogue of the plays via formulaic language, these outcomes represent a considerable achievement.

I would qualify this claim, however, by acknowledging that "formulaic" language in real speech is not necessarily the same as "formulaic" dramatic dialogue (see e.g. the categories for the latter identified by Blake, 2002:283-290, in S.1.3), and with more space I could usefully have tried to clarify this further. Undoubtedly the two are related, since playwrights and dramatists presumably apply their own natural storage and retrieval systems for spoken language to the dialogue they create. My study has shone some light on the ways that Shakespeare evidently mapped his own knowledge and interpretations of formulaic language on to the dialogue of the male and female characters he created (for example, by making the female characters use 'I pray you' a great deal more than the male characters, see S.4.3.1).

My final research question (1.4.5) asked whether the statistically significant formulaic language in female and male dialogue in the plays showed evidence of traces of wider language and gender issues of the historical period (over and above stylistic effects). The answer is yes, although I cannot make any concrete claims about the extent to which this is so. Neither can I say exactly what stance Shakespeare may have been taking through the dialogue he created, since often there seemed to be several plausible explanations (as in the case of 'I pray you', in S.4.3.1). No firm claims about real speech could definitely be made from this or any study of drama, particularly when limited to a single author's works, but in S.5.3 below I suggest how further comparative studies in EME drama and other speech-type texts would build on my findings.

Whilst the variation between female and male dialogue in the plays revealed by my study does show traces of differences in contemporary EME language behaviour (discussed in S.4.6.1), these are an interesting by-product of a corpus stylistics analysis rather than an attempt at historical sociolinguistics research. Nevertheless, establishing some norms of language use in the complete works of one playwright has been immensely useful in its own right. It has enabled a comparative discussion of individuals and gendered social groups of characters against a backdrop of Shakespeare's entire character inventory, revealing new information about the ways they appear through what they tend to say relatively frequently. This bears out my decision to include **all** the dialogue in the plays, despite the limitations placed on the quantitative analysis performed by *WordSmith* due to the unequal size of the male and female data files (explained in S.3.3.1). Claims about the language of one character in a play are vastly strengthened when made in comparison to all the others created by Shakespeare, rather than to just a few of them. For example:

- the absence of 'I pray you' in the dialogue of the female characters Lady Macbeth and Cleopatra (S.4.3.1);
- the use of "what is your pleasure" by male Iago (example 27, S.4.4.1); and
- the use of 'I know thee' by male Titus to female Tamora (example 10, S.4.3.1)

can be said to be exceptional compared to the language behaviour of all other members of one sex or the other, adding weight to the claim that Shakespeare used language to construct gender atypically as part of the characterisation of individuals.

183

My research questions did not anticipate the revelation of some clear differences in the operation of dramatic devices in the female and male dialogue, discussed in S.4.6.3. However, incorporating an awareness of what is going on at different levels in the discourse structure of the plays in my analyses in Ss.4.3 to 4.5, using Short's (1996:169) model (in Figure 1, S.2.1), led to these surfacing as an added bonus. My study has shown that the dialogue of women:

- actually helps characterise men in the plays (through talking about them in personal and intimate ways, see S.4.5.1);
- is loaded more than male dialogue with certain types of questions (S.4.4.1) and discourses of teaching and learning (S.4.4.3) to help the audience grasp what is going on in the play.

This adds to what is known about the construction of the plays, and shows that as well as contributing to the plot, the female dialogue also serves a supporting role both to the male characters (onstage) and to the audience (at a higher discourse level). It is important to return to the fact that all the dialogue in the plays ultimately serves the end of telling the story of the play to the audience, as emphasised in S.1.2.

In this section I have shown that my research aim of taking a fresh, empirically-based look at the ways women are represented in Shakespeare's plays has successfully yielded much useful information about:

(i)   the stylistic characterisation of women and men in general and as individuals; and

184

(ii)   the operational role of female dialogue in communicating the play.

This is despite some limitations posed by the shape of the corpus itself and the absence of historical evidence which would have helped clarify some results. My results and discussions in Chapter 4 indicated some further research possibilities which would usefully pick up on particular aspects of variation, stylistic interest, or the construction of drama, and I end my study by mentioning these briefly.

5.3   <u>Suggestions for future research</u>

As indicated in S.5.2, the extent to which variation in the male and female dialogue in the plays (evidenced by my data) can be linked to wider EME language variation is uncertain. Results which could be further tested in a multi-genre variation study (including other contemporary EME drama and other genres of speech-like texts) include:

- the significant keyness of the polite discourse marker 'I pray you' in female dialogue (S.4.3.1);

- the apparent restriction or limitation of 'I know thee' to male in-group language (S.4.3.1); and

- the apparently stereotypical representation of women talking more about what they have heard and men about what they have seen (S.4.4.2).

For example, a multi-genre variation study of 'I pray you' and other *pray* formulae would help establish whether the greater use of 'I pray you' among women in Shakespeare's plays is actually part of a wider trend towards supplication in requests

made by early modern women (due to their relatively low power status), whether it arises from the contextual circumstances of characters in the course of the drama, or whether it is related to something more simple such as a link to the number of requests in the dialogue.

As suggested in S.5.2, a study examining both the male and female sub-corpora using a reference corpus of other EME drama would usefully enable further discussion of the male dialogue in Shakespeare's plays. A suitable reference corpus of contemporary texts could be constructed from, e.g., the *CED* (see S.3.3.2) and/or those from the relevant period of the *Helsinki Corpus*[11].

A more detailed analysis of the ways male and female characters self-construct manhood and womanhood in the plays would be valuable, since my analysis in S.4.5.1 was necessarily limited to the cluster 'I am a woman' in my data and the corresponding male formula "I am a man". This would reveal more about the ways Shakespeare's work contributed to contemporary discussions about the place and power of women at the time (cf. Culpeper, 2000:312; see Ss. 1.1 and 2.5).

Finally, I mentioned very briefly that Ideational: Topical: States clusters in my results functioned literally or metaphorically (in Table 3, S.3.4.2 and in S.4.5.1). I did not have space in this study to look at the kinds of metaphors used frequently by male and

---

[11] *The Helsinki Corpus of English Texts* (1991). Department of English, University of Helsinki. Compiled by Matti Rissanen, Merja Kytö, Leena Kahlas-Tarkka, Matti Kilpiö, Saara Nevanlinna, Irma Taavitsainen, Terttu Nevalainen and Helena Raumolin-Brunberg. See http://www.helsinki.fi/varieng/CoRD/corpora/HelsinkiCorpus (accessed 12.08.2009)

female characters, and the impressions these create, but this would undoubtedly be interesting.

The above suggestions would build on the findings of my study, taking further what I have shown here about the ways Shakespeare represents women (and men) in his plays through what they say relatively frequently in their dialogue, both as individual characters and also as members of gendered social groups, at a time when the role of women in society was beginning to change considerably. As argued in S.5.2, the outcomes of my study have relevance to several major disciplines, including corpus linguistics, stylistics and historical sociolinguistics. My study also confirms that modern computer technology has much to offer in illuminating the representation of women and men by arguably the most famous playwright in the history of English language and literature.

<u>REFERENCES</u>

Aijmer, K. (1996) *Conversational Routines in English*. London and New York:
Addison Wesley Longman Limited.

Altenberg, B. (1998) "On the Phraseology of Spoken English: The Evidence of
Recurrent Word-Combinations". In Cowie, A.P. (ed.) *Phraseology: Theory, Analysis
and Applications.* Oxford: Clarendon Press, pp. 101-122.

Archer, D. (2007) "Computer-assisted Literary Stylistics: The State of the Field". In
Lambrou, M. and Stockwell, P. (eds.), pp. 244-256.

Archer, D., McEnery, T., Rayson, P. and Hardie, A. (2003) "Developing an automated
semantic analysis system for Early Modern English". In *Proceedings of the Corpus
Linguistics 2003 conference.* [UCREL technical paper number 16]. UCREL,
Lancaster University, U.K., pp. 22-31.

Argamon, S., Chase, P., Hota, S.R., Garg, N., Levitan, S. and Whitelaw, C. (2007)
"Stylistic Text Classification Using Functional Lexical Features". In *Journal of the
American Society for Information Science and Technology* 58(6): 802-822.

Astington, J.H. (2001) "Playhouses, players, and playgoers in Shakespeare's time". In
de Grazia, M. and Wells, S. (eds.) *The Cambridge Companion to Shakespeare.*
Cambridge: Cambridge University Press, pp. 99-113.

Aston, E. and Savona, G. (1991) *Theatre as sign-system: a semiotics of text and
performance.* London: Routledge.

Austin, J.L. (1962) *How to do Things with Words.* Oxford: Oxford University Press.

Bach, R.A. (2007) *Shakespeare and Renaissance Literature Before Heterosexuality.*
Basingstoke and New York: Palgrave Macmillan.

Baker, P. (2004) "Querying Keywords. Questions of Difference, Frequency, and Sense
in Keywords Analysis". In *Journal of English Linguistics* 32(4): 346-359.

Baker, P., Hardie, A. and McEnery, T. (2006) *A Glossary of Corpus Linguistics.*
Edinburgh: Edinburgh University Press.

Baron, A. and Rayson, P. (2008) "*VARD* 2: A tool for dealing with the spelling
variation in historical corpora". In *Proceedings of the Postgraduate Conference in
Corpus Linguistics*, Aston University, Birmingham, U.K., 22 May 2008.
http://www.comp.lancs.ac.uk/~barona/*VARD*2 (accessed 20.01.2009).

Baron, A., Rayson, P. and Archer, D. (2009) "Word frequency and key word statistics
in historical corpus linguistics". In Ahrens, R. and Antor, H. (eds.) *Anglistik:
International Journal of English Studies* 20(1): 41-67.

Berber Sardinha, A.P. (2004) *Lingüística de Corpus*. Barueri, SP, Brazil: Manole.

Biber, D., Conrad, S. and Cortes, V. (2003) "Lexical bundles in speech and writing:
An initial taxonomy". In Wilson, A., Rayson, P. and McEnery, T. (eds.) *Corpus
Linguistics by the Lune: A Festschrift for Geoffrey Leech.* Frankfurt am Main: Peter
Lang, pp. 71-92.

Biber, D., Conrad, S. and Cortes, V. (2004) "*If you look at...*: Lexical bundles in
university teaching and textbooks". In *Applied Linguistics* 25(3): 371–405.

Biber, D., Johansson, S., Leech, G., Conrad, S. and Finegan, E. (1999) *Longman
Grammar of Spoken and Written English*. Harlow: Pearson Educated Limited.

Blake, N.F. (2002) *A Grammar of Shakespeare's Language*. Basingstoke and New
York: Palgrave.

Blum-Kulka, S., House, J. and Kasper, G. (eds.) (1989) *Cross-Cultural Pragmatics:
Requests and apologies.* [Vol. XXXI Advances in Discourse Processes]. Norwood,
New Jersey: Ablex.

Bousfield, D. (2007) "'Never a Truer Word Said in Jest': A Pragmastylistic Analysis
of Impoliteness as Banter in *Henry IV, Part I*". In Lambrou, M. and Stockwell, P.
(eds.), pp. 208-220.

Boyce, C. (1990) *Shakespeare A to Z. The essential reference to his plays, his poems,
his life and times, and more.* New York: Dell Publishing.

Brown, P. and Levinson, S.C. (1987) *Politeness: Some Universals in Language
Usage*. Cambridge: Cambridge University Press.

Brown, R.W. and Gilman, A. (1989) "Politeness theory and Shakespeare's four major
tragedies". In *Language in Society* (18): 159-212.

Bryson, A. (1998) *From Courtesy to Civility. Changing Codes of Conduct in Early
Modern England.* Oxford: Clarendon Press.

Busse, B. (2006) *Vocative Constructions in the Language of Shakespeare*. Amsterdam
and Philadelphia: John Benjamins.

Busse, U. (2002) *Linguistic Variation in the Shakespeare Corpus: Morpho-Syntactic
Variability of Second Person Pronouns.* Amsterdam: John Benjamins.

Cameron, D. (2005) "Language, Gender, and Sexuality: Current Issues and New
Directions". In *Applied Linguistics* 26(4): 482-502.

Coates, J. (2007) "Gender". In Llamas, C., Mullany, L. and Stockwell, P. (eds.) *The
Routledge Companion to Sociolinguistics.* Abingdon and New York: Routledge, pp.
62-68.

Cooper, M.M. (1998) "Implicature, convention and *The Taming of the Shrew*". In
Culpeper, J., Short, M. and Verdonk, P. (eds.), pp. 54-66.

Craig, W.J. (ed.) (1914/2000) *The Oxford Shakespeare*. London: Oxford University
Press and New York: Bartleby.com
http://www.bartleby.com/70/ (accessed 16.04.2008).

Craig, W.J. (ed.) (1916) *The Complete Works of William Shakespeare (The Oxford
Shakespeare)*. Oxford: Oxford University Press.
http://oll.libertyfund.org/index.php?option=com_staticxt&staticfile=show.php%3Ftitle
=1612&Itemid=28 (accessed 12.08.2009).

Crystal, D. (2003) *The Cambridge Encyclopaedia of the English Language*. (Second
Edition). Cambridge: Cambridge University Press.

Crystal, D. (2008) *"Think on my words". Exploring Shakespeare's Language.*
Cambridge: Cambridge University Press.

Crystal, D. and Crystal, B. (2002) *Shakespeare's Words. A Glossary and Language
Companion.* London: Penguin.

Culpeper, J. (1998) "(Im)politeness in dramatic dialogue". In Culpeper, J., Short, M.
and Verdonk, P. (eds.), pp. 83-95.

Culpeper, J. (2000) "A cognitive approach to characterization: Katherina in
Shakespeare's *The Taming of the Shrew*". In *Language and Literature* 9(4): 291-316.

Culpeper, J. (2001) *Language and Characterisation: People in Plays and Other Texts.*
Harlow: Pearson Education.

Culpeper, J. (2002) "Computers, language and characterisation: An analysis of six
characters in *Romeo and Juliet*". In Melander-Marttala, U., Ostman, C. and Kytö, M.
(eds.), *Conversation in Life and in Literature: Papers from the ASLA Symposium*,
Association Suedoise de Linguistique Appliquée (ASLA), 15. Universitetstryckeriet:
Uppsala, pp.11-30.

Culpeper, J. and Archer, D. (2008) "Requests and directness in Early Modern English
trial proceedings and play-texts, 1640-1760". In Jucker, A. H. and Taavitsainen, I.
(eds.). *Speech Acts in the History of English.* Amsterdam and Philadelphia: John
Benjamins, pp. 45-84.

Culpeper, J. and Kytö, M. (2005) *Exploring speech-related Early Modern English
texts: lexical bundles re-visited*. Paper presented at the 26th Conference of ICAME
(International Computer Archive of Modern and Medieval English), University of
Michigan, U.S.A., May 2005.

Culpeper, J. and Kytö, M. (forthcoming, 2010) *Speech in Writing: Explorations in
Early Modern English Dialogues.* Cambridge: Cambridge University Press.

Culpeper, J. and McIntyre, D. (2006) "Drama: Stylistic Aspects", in Brown, K. (ed.)
Encyclopedia of Language and Linguistics. (Second Edition). Oxford: Elsevier Ltd.
Vol. 3, pp. 772-785.

Culpeper, J., Short, M. and Verdonk, P. (eds.) (1998) *Exploring the Language of Drama. From Text to Context.* London and New York: Routledge.

De Cock, S. (1999) "Repetitive phrasal chunkiness and advanced EFL speech and writing", in Mair, C. and Hundt, M. (eds.) *Corpus Linguistics and Linguistic Theory.* Papers from the Twentieth International Conference on English Language Research on Computerized Corpora (ICAME 20), Freiburg im Breisgau 1999. Amsterdam and Atlanta: Rodopi, pp. 51-68.

Demmen, J.E.J. (2007) *Key lexical bundles in Shakespeare: An analysis of their functions and stylistic effects in comedy, tragedy and history plays.* Unpublished undergraduate dissertation, Lancaster University, U.K.

Dunning, T. (1993) "Accurate Methods for the Statistics of Surprise and Coincidence". In *Computational Linguistics* 19(1): 61-74.

Dusinberre, J. (1996) *Shakespeare and the Nature of Women.* (Second Edition). Basingstoke: Macmillan Press.

Fennell, B.A. (2001) *A History of English.* Blackwell Publishing.

Findlay, A. (1999) *A Feminist Perspective on Renaissance Drama*. Oxford and Malden: Blackwell Publishers.

Firth, J.R. (1957) "A synopsis of linguistic theory 1930-1955". In *Studies in Linguistic Analysis.* [Special volume of the Philological Society, Oxford], pp. 168-205.

French, M. (1982) *Shakespeare's Division of Experience*. London: Cape.

Gilbert, A.J. (1997) *Shakespeare's Dramatic Speech*. [Studies in Renaissance Literature 15]. Lewiston, Queenston and Lampeter: Edwin Mellen Press.

Gowing, L. (1996) *Domestic Dangers: Women, Words, and Sex in Early Modern London.* [Oxford Studies in Social History]. Oxford: Clarendon Press.

Greenblatt, S., Cohen, W., Howard, J.E. and Maus, K.E. (eds.) (1997) *The Norton Shakespeare*. London and New York: W.W. Norton & Company.

Habermann, I. (2003) *Staging Slander and Gender in Early Modern England.* Aldershot, U.K. and Burlington, U.S.A.: Ashgate Publishing.

Halliday, M.A.K. (1994) *An Introduction to Functional Grammar*. (Second Edition). London and New York: Edward Arnold.

Heilman, R.B. (1972) "Introduction to *The Taming of the Shrew*". In Barnet, S. (ed.) *The Complete Signet Classic Shakespeare*. New York: Harcourt Brace Jovanovich, pp. 321-329.

Hillman, R. (1997) *Self-Speaking in Medieval and Early Modern English Drama. Subjectivity, Discourse and the Stage*. Basingstoke: Macmillan Press Ltd. and New York: St. Martin's Press, Inc.

Hoey, M. (1991) *Patterns of Lexis in Text*. Oxford: Oxford University Press.

Hoey, M. (2005) *Lexical Priming: A New Theory of Words and Language*. London: Routledge.

Hope, J. (2003) *Shakespeare's Grammar*. London: The Arden Shakespeare.

Hota, S., Argamon, S., Koppel, M. and Zigdon, I. (2006) "Performing Gender: Automatic Stylistic Analysis of Shakespeare's Characters". In *Digital Humanities*, July 2006: 81-82.
http://lingcog.iit.edu/pub_year.xml (accessed 16.04.2008).

Jardine, L. (1983) *Still Harping on Daughters: Women and Drama in the Age of Shakespeare.* (Second Edition, 1989). Sussex: Harvester Press and New Jersey: Barnes and Noble.

Kahn, C. (1977) The Taming of the Shrew: Shakespeare's Mirror of Marriage. In Diamond, A. and Edwards, L.R. (eds.) *The Authority of Experience: Essays in Feminist Criticism.* Amherst: University of Massachusetts Press, pp. 84-100.

Kahn, C. (1981) *Man's Estate: Masculine Identity in Shakespeare*. Berkeley: University of California Press.

Kastan, D.S. (2001) "Shakespeare and English history". In de Grazia, M. and Wells, S. (eds.) *The Cambridge Companion to Shakespeare.* Cambridge: Cambridge University Press, pp. 167-182.

Kelley, H.H. (1972) "Attribution in social interaction". In Jones, E.E, Kanouse, D.E., Kelley H.H., Nisbett, R.E., Valins, S. and Weiner, B. (eds.) *Attribution: Perceiving the Causes of Behavior*. Morristown, New Jersey: General Learning Press, pp. 1-26.

Klein, L.E. (2002) "Politeness and the Interpretation of the British Eighteenth Century". In *The History Journal* 45(5): 869-898.

Kopytko, R. (1993) *Polite discourse in Shakespeare's English*. [Seria Filologia Angielska 24]. Poznań: Adam Mickiewicz University Press.

Kopytko, R. (1995) "Linguistic politeness strategies in Shakespeare's plays". In Jucker, A.H. (ed.) *Historical pragmatics: Pragmatic developments in the history of English.* [Pragmatics & Beyond. New Series 35]. Amsterdam: John Benjamins, pp. 515-540.

Kytö, M. and Walker, T. (2006) *Guide to A Corpus of English Dialogues 1560-1760.* Acta Universitatis Upsaliensis. Studia Anglistica Upsaliensia 130. Uppsala: Uppsala Universitet.

Lambrou, M. and Stockwell, P. (eds.) (2007) *Contemporary Stylistics*. London and
New York: Continuum.

Lenk, U. (1999) "Stabilized expressions in spoken discourse: Worth our time?". In
Mair, C. and Hundt, M. (eds.) *Corpus Linguistics and Linguistic Theory*. Papers from
the Twentieth International Conference on English Language Research on
Computerized Corpora (ICAME 20), Freiburg im Breisgau 1999. Amsterdam and
Atlanta: Rodopi, pp. 187-200.

Livia, A. (2003) "'One Man in Two is a Woman": Linguistic Approaches to Gender in
Literary Texts". In Holmes, J. and Meyerhoff, M. (eds.) *The Handbook of Language
and Gender*. Malden, Oxford, Melbourne and Berlin: Blackwell, pp. 142-158.

Loomba, A. (1989) *Gender, race, Renaissance drama*. Manchester and New York:
Manchester University Press.

Mahlberg, M. (2007a) "Clusters, key clusters and local textual functions in Dickens".
In *Corpora* 2(1): 1-31.

Mahlberg, M. (2007b) "A Corpus Stylistic Perspective on Dickens' *Great
Expectations*". In Lambrou, M. and Stockwell, P. (eds.), pp. 19-31.

McEnery, T., Xiao, R. and Tono, Y. (2006) *Corpus-Based Language Studies: An
advanced resource book.* Abingdon and New York: Routledge.

Mills, S. (1995) *Feminist Stylistics*. London and New York: Routledge.

Moon, R. (1998) *Fixed Expressions and Idioms in English. A Corpus-Based
Approach*. Oxford: Clarendon Press.

Mueller, M. (2005) "The Nameless Shakespeare". In *TEXT Technology* (No. 1, 2005):
61-70.
http://texttechnology.mcmaster.ca/pdf/vol14_1_06.pdf (accessed 14.04.2008).

Mueller, M., Parod, W., Cousens, J., Burns, P. and Norstad, J. (2006). *WordHoard*.
Evanston, Illinois: Northwestern University.
http://wordhoard.northwestern.edu/userman/whatiswordhoard.html (accessed
14.04.2008).

Murphy, S. (2007) *Now I am alone: A corpus stylistic approach to Shakespearian
soliloquies*. Paper from the Lancaster University Postgraduate Conference in
Linguistics & Language Teaching, Vol. 1. In Gabrielatos, C., Slessor, R. and Unger, J.
(eds.) *Papers from LAEL PG 2006*.
http://www.ling.lancs.ac.uk/pgconference/v01/Murphy.pdf (accessed 14.04.2008).

Nevalainen, T. (1996) "Gender difference". In Nevalainen, T. and Raumolin-
Brunberg, H. (eds.) *Sociolinguistics and Language History. Studies based on the
Corpus of Early English Correspondence.* Amsterdam and Atlanta: Rodopi, pp. 77-91.

Nevalainen, T. and Raumolin-Brunberg, H. (2003) *Historical Sociolinguistics*.
Harlow: Pearson Education Limited.

Onions, C.T. (1982) *A Shakespeare Glossary*. (Second Edition). Oxford: Clarendon
Press.

Partington, A. and Morley, J. (2002) "From frequency to ideology: investigating word
and cluster/bundle frequency in political debate". Conference paper from *Teaching
and Language Corpora (TaLC)*, 2002.

Piao, W., Wilson, A., and McEnery, A. (2002) "*A multilingual corpus toolkit*". Paper
presented at the Fourth North American Symposium on Corpus Linguistics,
Indianapolis, Indiana, 1-3 November 2002.
http://bowland-files.lancs.ac.uk/corplang/cbls/resources.asp (accessed 11.04.2008).

Quirk, R. (1961) "Some observations on the language of Dickens". In *A Review of
English Literature* 2(3): 19-28.

Rayson, P., Archer, D., Baron, A., Culpeper, J. and Smith, N. (2007) "Tagging the
Bard: Evaluating the accuracy of a modern POS tagger on Early Modern English
corpora". In *Proceedings of Corpus Linguistics 2007*, Birmingham University, U.K.,
27-30 July 2007.
http://ucrel.lancs.ac.uk/VariantSpelling/ (accessed 12.01.2008).
http://www.comp.lancs.ac.uk/~barona/*VARD2* (accessed 24.06.2008).

Rayson, P., Archer, D., Smith, N. (2005) "*VARD* versus WORD: A comparison of the
UCREL variant detector and modern spellcheckers on English historical corpora". In
Proceedings of Corpus Linguistics 2005, Birmingham University, U.K. ,14-17 July
2005; *Proceedings from the Corpus Linguistics Conference Series* on-line e-journal,
Vol. 1, no. 1, ISSN 1747-9398.
http://ucrel.lancs.ac.uk/VariantSpelling/ (accessed 12.01.2008).

Rayson, P., Berridge, D. and Francis, B. (2004) "Extending the Cochran rule for the
comparison of word frequencies between corpora". In Purnelle G., Fairon C., Dister A.
(eds.) *Le poids des mots: Proceedings of the 7th International Conference on
Statistical analysis of textual data (JADT 2004)*, Vol. II, pp. 926-936. 10-12 March
2004, Louvain-la-Neuve, Belgium: Presses universitaires de Louvain. ISBN 2-
930344-50-4.

Rose, M.B. (1988) *The Expense of Spirit. Love and Sexuality in English Renaissance
Drama.* Ithaca, New York: Cornell University.

Salmon, V. (1965) "Sentence structures in colloquial Shakespearean English". In
Transactions of the Philological Society, pp. 105-140. [Reprinted in Salmon, V. and
Burness, E. (eds.)], pp. 265-300.

Salmon, V. (1967) "Elizabethan colloquial English in the Falstaff Plays". In *Leeds
Studies in English* 1, pp. 37-70. [Reprinted in Salmon, V. and Burness, E. (eds.)], pp.
37-70.

Salmon, V. and Burness, E. (eds.) (1987) *A reader in the language of Shakespearean drama.* [Amsterdam Studies in the Theory and History of Linguistic Science 35]. Amsterdam: John Benjamins,

Scott, M.R. (1999) *WordSmith Tools Version 3.0*. Oxford: Oxford University Press.

Scott, M. (2006) "In Search of a Bad Reference Corpus". Paper presented at the *AHRC ICT Methods Network Expert Seminar on Linguistics*, Lancaster University, U.K., 8 September 2006.
http://www.methodsnetwork.ac.uk/redist/pdf/es1_05scott.pdf (accessed 24.06.2008).

Scott, M. and Thompson, G. (eds.) (2001) *Patterns of Text in honour of Michael Hoey*. Amsterdam and Philadelphia: John Benjamins.

Scott, M. and Tribble, C. (2006) *Textual Patterns. Key words and corpus analysis in language education.* Amsterdam and Philadelphia: John Benjamins Publishing Company.

Searle, J.R. (1969) *Speech Acts: An Essay in the Philosophy of Language*. Cambridge: Cambridge University Press.

Semino, E. and Short, M. (2004) *Corpus Stylistics*. Abingdon and New York: Routledge.

Short, M. (1996) *Exploring the Language of Poems, Plays and Prose*. London and New York: Longman.

Short, M. (1998) "From dramatic text to dramatic performance". In Culpeper, J., Short, M. and Verdonk, P. (eds.), pp. 6-17.

Sinclair, J. (1966) "Beginning the study of lexis". In Bazell, C., Halliday, M.A.K., Robins, R.H. and Catford, J. (eds.) *In memory of J.R. Firth.* London: Longman, pp. 410-430.

Sinclair, J. (1991) *Corpus, Concordance, Collocation*. Oxford: Oxford University Press.

Sinclair, J.M. (2004) (edited with Ron Carter) *Trust the Text: Language, corpus and discourse*. London and New York: Taylor and Francis.

Sinclair, J.M. and Coulthard, R.M. (1975) *Towards an Analysis of Discourse*. Oxford: Oxford University Press.

Spevack, M. (1968-80) *A Complete and Systematic Concordance to the Works of Shakespeare.* 9 volumes. Hildesheim: George Ohms.

Starcke, B. (2005) "Phraseology and meaning in literary texts and corpora: how frequent phraseological units contribute both to the characterization of protagonists and places and the structural organisation of literary texts". Paper given at the pre-

conference workshop at *Corpus Linguistics 2005*, Birmingham University, U.K., 14 July 2005.

Starcke, B. (2006) "The phraseology of Jane Austen's *Persuasion*: phraseological units as carriers of meaning". In *ICAME Journal* 30: 87-104.

Stubbs, M. (1996) *Text and Corpus Analysis*. Oxford, U.K. and Malden, U.S.A.: Blackwell.

Stubbs, M. (2001) *Words and Phrases. Corpus Studies of Lexical Semantics*. Oxford, U.K. and Malden, U.S.A.: Blackwell.

Stubbs, M. (2005) "Conrad in the computer: examples of quantitative stylistic methods". In *Language and Literature* 14(1): 5-24.

Stubbs, M. and Barth, I. (2003) "Using recurrent phrases as text-type discriminators. A quantitative method and some findings". In *Functions of Language* 10(1): 61–104.

Sunderland, J. (2006) *Language and Gender. An advanced resource book*. Abingdon and New York: Routledge.

Talbot, M.M. (1998) *Language and Gender*. Cambridge: Polity Press.

Taavitsainen, I. (1995) "Interjections in Early Modern English: From imitation of spoken to conventions of written language". In Jucker, A.H. (ed.) *Historical pragmatics: Pragmatic developments in the history of English.* [Pragmatics & Beyond. New Series 35]. Amsterdam: John Benjamins, pp. 439-465.

Wales, K. (1989) *A Dictionary of Stylistics*. Harlow: Longman Group UK Limited.

Weinert, R. (1995) "The role of formulaic language in second language acquisition: A review". In *Applied Linguistics* 16(2): 180-205.

Wells, S. and Taylor, G. with Jowett, J. and Montgomery, W. (1987) *William Shakespeare. A Textual Companion*. Oxford: Clarendon Press.

Williams, R. (1976) *Keywords*. (Second Edition, 1983). London: Fontana.

Wolfreys, J. (ed.) (1999) *Literary Theories. A Reader and Guide.* Edinburgh: Edinburgh University Press.

Wray, A. (2002) *Formulaic Language and the Lexicon*. Cambridge: Cambridge University Press.

Wrightson, K. (1982) *English Society, 1580-1680*. London: Unwin Hyman.

Wynne, M. (2006) "Stylistics: Corpus Approaches". In Brown, K. (ed.) *Encyclopedia of Language and Linguistics.* (Second Edition). Oxford: Elsevier Ltd. Vol. 12, pp. 223-226.

APPENDIX I

Plays by William Shakespeare in the corpus

Comedies:
A Midsummer Night's Dream
All's Well That Ends Well
As You Like It
Cymbeline
Love's Labour's Lost
Measure for Measure
Much Ado About Nothing
Pericles
The Comedy of Errors
The Merchant of Venice
The Merry Wives of Windsor
The Taming of the Shrew
The Tempest
The Winter's Tale
Troilus and Cressida
Twelfth Night
Two Gentlemen of Verona

Tragedies:
Antony and Cleopatra
Coriolanus
Hamlet
Julius Caesar
King Lear
Macbeth
Othello
Romeo and Juliet
Timon of Athens
Titus Andronicus

Histories:
Henry the Fourth Part One
Henry the Fourth Part Two
Henry the Eighth
Henry the Fifth
Henry the Sixth Part One
Henry the Sixth Part Two
Henry the Sixth Part Three
King John
Richard the Second
Richard the Third

Source of the corpus: Dr. Mike Scott's Shakespeare corpus (see www.lexically.net), the material for which was based on Craig, W.J. (ed.) (1916) *The Complete Works of William Shakespeare (The Oxford Shakespeare)*. Oxford: Oxford University Press, obtained from the *Online Library of Liberty*, hosted by Liberty Fund, Inc. (see http://oll.libertyfund.org).