PART I

METHODOLOGY

Chapter 1

Introduction

Authorship studies in the early Modern period

There is a great number of Elizabethan and Jacobean plays which are either anonymous, wrongly ascribed, or thought to be the work of more than one writer. This situation is a result of the particular context in which early Modern plays were written, acted, and published – and an understanding of this context explains many of the most common authorship problems which arise (Bentley 1971, and Wells *et al.* 1987:1–68 provide excellent accounts of these issues in more detail than can be given here). Such an understanding can also help in explaining the shortcomings and limitations of much previous authorship work.

The single most important factor in this context is that early Modern plays were only very rarely regarded as 'literature' in a sense recognisable today. They are better regarded as raw material fuelling the profitable entertainment industry of early Modern London, much as film scripts are the raw material of today's film industry. Like film scripts, they were bought from writers by acting companies and, just as today, once a script was sold, the writer lost control over it.

The demand of the theatres for new material, especially in the later sixteenth and early seventeenth centuries, was insatiable, and, again as often in the film industry, this favoured collaboration as a means of producing material of the required standard in as short a time as possible. Writers might specialise in certain types of writing – opening scenes, closing scenes, comic scenes, love scenes; and unsatisfactory scripts might be touted round various authors who would add to and cut the script in the hope of producing something playable.

Furthermore, the fact that the ability and right to produce a play lay in the ownership of a licensed copy was a disincentive to publication. Acting companies had an interest in protecting their investment in a popular play by not publishing it. Plays might thus be in repertory, but unpublished, for a considerable time, during which time they would naturally be adapted for each new production, possibly with topical additions by different dramatists. When plays *were* released for publication (significantly often in times of

3

plague, when the theatres were closed, and companies had to find alternative income), the process of transmission of texts produced effects which are highly relevant to authorship studies (again, Wells *et al.* 1987:9–52, provide a detailed account). Most important is the recognition that certain aspects of the author's original manuscript – spelling, punctuation, even some word forms – would not necessarily have been transmitted faithfully by those who prepared subsequent versions of the text: scribes preparing fair copies to sell to acting companies; bookkeepers producing prompt texts; compositors setting type from handwritten pages, these may all have introduced changes to the original. The printing process is particularly relevant here, not only because many authors seem to have punctuated manuscripts very lightly, if at all, in the expectation that pointing would be supplied in the printing house if the text were ever printed, but because the practicalities of fitting text onto lines and pages often introduced expansions and contractions (see Wells *et al.* 1987:42, 44–5, for illustrations). In many cases, the effects of these factors are unpredictable: we know, for example, that some scribes regularly preserve original spelling, others do so intermittently, while others rarely do so.

This context has numerous consequences for authorship studies. The most theoretically challenging is that our very desire to fix the authorship of early Modern play-texts is something of an anachronism, stemming largely from what Stillinger calls the 'romantic myth of the author as solitary genius' (1991:203). Dramatic texts are inherently collaborative, and there comes a point at which we have to admit that we can never know who wrote which word. Even so, early Modern dramatic collaboration was a physical process, which left physical evidence in the texts it produced. Various processes have intervened to corrupt that evidence, but often it is still detectable and interpretable – if not to the extent we might wish.

In a more concrete way, this context of production means that we have to be alive to the nature of the evidence an early Modern play-text presents us with. We need to know which features of a text are stable, and which are not. We need to know that spelling, punctuation, lineation, contractions, sentence length, oaths, stage directions are all subject to being changed by hands which are not those of the 'author' of the play. If we are to use these features as evidence for the authorship of a text (and they all have been), we need to know the likelihood of their being changed, and the consequences if they were. We also need to know that any play in the repertory of a London company for any length of time might have been cut, revised, or had material added for a particular performance, a tour, or to bring it up to date. This work would probably have been done by the company's contracted playwright, whoever the original author of the play had been.

The consequence of this is that any expectation of textual integrity, or

purity, in early Modern play-texts is misplaced. Modern authorship studies, like modern editing, is not about returning us to some 'original' text, neatly doled out to collaborators, or ascribed to a single author. What it tries to do – what I have tried to do in this book – is look at the evidence present in the texts we have, understand how that evidence got there, and reveal what that evidence tells us about the history of how the text came to be like it is.

Early Modern English and socio-historical linguistic evidence

I approach the question of the authorship of early Modern English play-texts via the historical linguistic context in which they were written. Early Modern English is conventionally dated from 1500–1700, and while linguistic change is a continuous process rather than a collection of discrete events, these dates are generally accepted as being meaningful by historical linguists, encompassing as they do a period in which written texts in England can be seen to show progressively less variation in a number of linguistic features as a written Standard language emerges (as we await Lass (ed.) forthcoming, the best book on early Modern English remains Barber 1976).

The key to socio-historical linguistic evidence is the fact that linguistic change is a process: when a change is in progress, the alternative forms will co-exist in the language. Changes in progress in early Modern English offered alternatives to writers and speakers where Present-day Standard English offers none – for example, a choice between 'you' and 'ye' in the second person pronoun, and '-th' and '-s' as an ending for the third person singular present tense of verbs (e.g. 'hath' versus 'has').

Alternatives such as these have been used as evidence in authorship studies (notably in the work of Hoy on the Fletcher canon) – unfortunately, these features are textually unstable: scribes and compositors could, and did, alter them when transmitting texts. This means that any attempt to use such evidence as an indicator of authorship is likely to be beset by questions of possible interference: results which do not conform to the subjective prejudices of the researcher can be explained away, and the tendency will always be to dismiss or discount inconvenient evidence (see chapter 5, and my discussion of Hoy's work on *Henry VIII*).

However, there are linguistic changes in progress in the early Modern period whose alternates (or more accurately variants) are not regarded as interchangeable by scribes and compositors, and which are therefore textually stable. Furthermore, by drawing on modern advances in linguistics, it is possible to elucidate, and in some cases predict, differences in the usages of variants by writers.

Socio-linguistics, as developed by Labov (1972), offers techniques for analysing and explaining language variation and change, in relation to factors such as age, sex, and class of language users. These techniques depend on the abstract notion of the linguistic variable; that is, an element of language which has two or more possible realisations (variants), which vary within the usage of a group of speakers of a language. In classical socio-linguistics, linguistic variables are phonological (for example pronunciation or non-pronunciation of /r/ in words such as 'fourth' and 'floor'), and their use can be related to social factors such as age, class, and sex. Socio-historical linguistics (Romaine 1982) has attempted, with success, to apply a methodology developed for the study of contemporary speech, to historically distant writing, specifically treating relative marker choice as a linguistic variable.

Crucial for the value of socio-linguistics to authorship studies is the fact that quantitative socio-linguistic methods are explicitly designed to deal with variation in linguistic forms which is not categorical (either form A or form B appears in a text), but is a matter of more or less (forms A and B both occur in texts, but in differing proportions, depending on factors such as the age or class of the producer of the text, and the situation in which it was produced). Quantitative socio-linguistics can therefore enable the researcher to make statistical comparisons between texts on the basis of the usages of linguistic variants within them, as Romaine notes.

> Anyone can observe that two samples of speech or writing are different. Sociolinguistic analysis can show that these differences are objectively measurable, and that there are patterns in the choices which a speaker/writer does make, on the one hand, and can make, on the other.
>
> (Romaine 1982:13)

Socio-historical linguistic evidence attempts to use the predictable patterning of incoming and recessive variants during language change in order to detect the hand of a chosen author. The key to this is the 'S' curve model of linguistic change (Aitchison 1991:83–5).

To illustrate this, let us assume that a change is taking place within a linguistic feature of a language. This change involves an older variant (A) and a newer variant (B), which is replacing the older one. In the course of this change, B forms will not appear randomly in the language, nor will A forms disappear overnight. Socio-linguistic models of change would lead us to expect that the shift from A to B within the language will be patterned in time, within the speech community, and within the usage of individual members of the speech community.

If we were to draw a graph of this change, we would expect it to have a distinctive 'S' curve – B forms would appear first in certain restricted
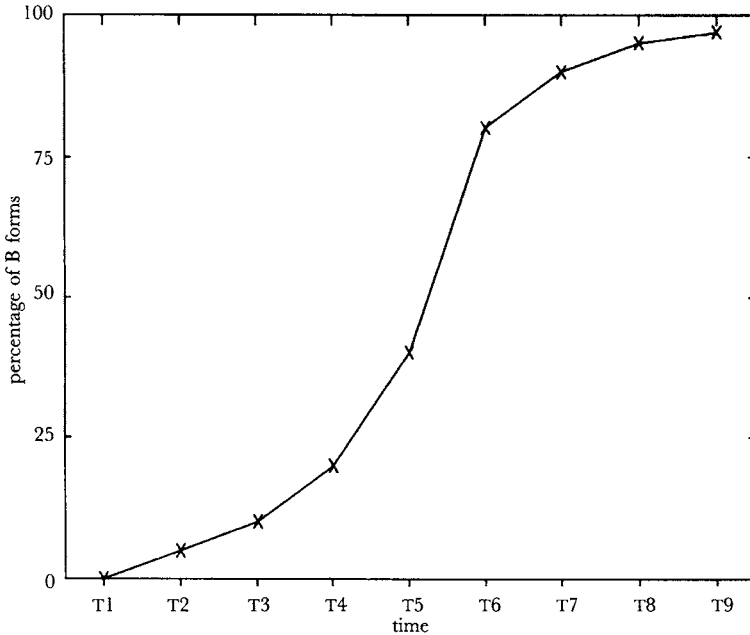
Graph 1.1   'S' curve of linguistic change (after Aitchison 1991:84)

contexts (restricted grammatically, geographically, and by individual) and would remain in a small proportion relative to A forms for some time. Then, as the change took off, B forms would increase exponentially, 'invading' contexts in which A had previously been the only variant found. Finally a third stage is reached, with B numerically dominant, but A retained in a few specific contexts (see graph 1.1, where T1 . . . T9 represent successive sampling points).

An example of such a change in Standard English would be the replacement of 'thou' by 'you' as the second person singular pronoun form. For this change, the 'S' curve begins with a few instances of singular 'you' in letters in the thirteenth century, has its exponential phase in the sixteenth and seventeenth centuries, and settles down from the eighteenth century to the present, with 'thou' retained only in highly specialised poetic and liturgical contexts.

The important point here is that during the exponential phase, an early Modern English speaker's frequency of use of 'thou' will be determined by a number of factors. We know that in-coming prestige variants like 'you' are used more frequently by younger, more educated, more urban members of the speech community. Because linguistic usages tend to be fixed early in life, two writers writing in the same year (even in the same text) who have

different dates of birth, will have differing rates of usage of 'thou' corresponding to the different points at which their birthdates intersect the curve in the proportional change from 'thou' to 'you'.

The basic premise of socio-historical linguistic evidence is that early Modern English writers will show differences in their usages of certain early Modern English variables, according to the influence of the factors identified by socio-linguistics as playing a role in patterning linguistic variation and change. These usages will be identifiable and distinguishable statistically, and will be explicable in terms of socio-linguistic theory.

For example, John Fletcher was born in 1579 in the south-east and brought up in an upper-class, urban environment. His father became bishop of London in 1594, but the family had probably been resident in the capital since 1589. His uncle was a distinguished Elizabethan diplomat, and (despite Taunton 1990) the evidence favours the view that Fletcher attended Cambridge University (see Collinson 1983, Berry 1960 and 1964). This contrasts with Shakespeare's birth, fifteen years earlier in the rural south-west midlands, his lower class status, and lack of higher education. All of these factors would tend to suggest that Fletcher will use more in-coming prestige variants than Shakespeare.

In this ability to explicate, and even predict the differences it studies between authors, socio-historical linguistic evidence differs crucially from the most successful current approach to authorship studies: stylometry.

Stylometric tests typically involve the computational analysis of texts on the basis of factors such as average length of word, and the frequency of the appearance of certain words at the start of sentences. A wide range of other tests has been developed following the initial work in this area by A.Q. Morton (see Morton 1978; Kenny 1982 presents a more critical view of the claims of Morton and his followers). Leading workers in the field, whose studies are referenced in the Bibliography, include M. W. Smith, Morton, Stratil and Oakley, and Merriam. The work of Jackson, Lake, and Holdsworth at times draws on stylometric tools, as does that of the Shakespeare Clinic (see references to articles by Elliot and Valenza).

Stylometric tests can appear impressive, especially to those not familiar with the statistical tests used. Wells *et al.* however (1987:80), note difficulties which arise when Morton's techniques are applied to early Modern texts (not the least of which is how to define the sentence in texts punctuated in the printing house), as well as potential statistical problems (these are taken up in greater detail in various articles by M. W. Smith). Non-statisticians are also likely to suspect a fundamental methodological flaw in the basis of stylometric tests, which Furbank and Owens term the 'spectre of the meaningless' (1988:181). This is an unease with stylometric evidence based on its focus on apparently arbitrary factors, and the inability of its

practitioners to explain either what exactly they observe and measure, or why authors should be different on these grounds. As A. Q. Morton puts it,

> the question which now arises naturally is, if it is not the grammatical, linguistic or philological role of words which is being investigated, what is it that is being studied? To this reasonable request the only reasonable reply is that we do not know.

> (Morton 1976:15)

This lack of understanding within stylometry of the nature of the object being measured must raise potential doubts about the methods used to measure it: how can the stylometricians be sure that their tests are appropriate?

In fact, I suspect that this difficulty is one of public relations rather than a fundamental methodological flaw. Literary scholars and non-statisticians are unwilling to accept the conclusions of work which they cannot understand, and which appears to be arbitrarily based. None the less, it is a major strength of socio-historical linguistic evidence that the differences it detects between the linguistic usages of authors are explicable, and that it does not rely on complex statistical tests for validation.

Applying socio-historical linguistic evidence

The application of socio-historical linguistic evidence is essentially a two-stage process. In the first stage, the usages of the candidates for the authorship of a play, or number of scenes, are established from a comparison sample of their unaided work. In the second stage, results from the disputed texts are compared to those findings to indicate the possibility or otherwise of authorship by the suggested candidate. It will immediately be apparent that the application of the technique depends on the existence of a suitable comparison group of non-controversial texts.

The comparison samples analysed in chapters 2, 3, and 4 were selected on the basis of one over-riding factor: plays chosen had to be the unaided work of the named dramatist. This is easier for some writers than for others, and, as has been indicated, the very notion of an 'unaided' text at this time may be inappropriate.

When this study was begun, with the specific intention of separating Shakespeare and Fletcher's hands in *Henry VIII* and *The Two Noble Kinsmen*, an arbitrary decision was taken to include ten plays from each author in the samples. As work progressed, it became clear that this size of sample was not only time consuming to count by hand; it was unnecessary. As results in chapters 2 and 3 show, five plays give a more than adequate sample.

After initial work on auxiliary 'do' in Shakespeare and Fletcher therefore, comparison samples were reduced to five plays, and a second group of Shakespeare plays was selected from the early part of his career to allow

assessment of the extent to which linguistic usages might change over an
author's career. The selection of playwrights other than Shakespeare and
Fletcher was done on the basis of those who had been suggested as possible
authors of plays in the apocrypha. Suspicions of collaboration meant that in
the cases of Marlowe and Dekker it was not possible to select five plays for
the sample. (A list of the plays selected for each author can be found under
Abbreviations at the front of the book.)

Chapter 2

The auxiliary 'do'

Background

In Present-day Standard English, the auxiliary verb 'do' is used mainly in the formation of questions ('*Did* you go home?') and negatives ('I *didn't* go home'). Such usages are usually referred to as being 'regulated' (Barber 1976:263–7), which means that use of 'do' is obligatory in certain sentence types (negative declaratives, positive and negative questions) and absent from others (positive declaratives). The use of auxiliary 'do' in positive declaratives today automatically carries emphasis ('I *DID* go home').

In early Modern English however, the use of the auxiliary was optional in all of these cases – for example questions could be formed by inversion *or* by the use of auxiliary 'do': 'Went you home?' versus 'Did you go home?'. Positive declarative sentences in early Modern English could use the auxiliary without automatic implication of emphasis. Effectively early Modern speakers had a choice of two constructions whenever they formed any one of the four sentence types mentioned above. Constructions conforming to present-day usage are termed 'regulated', while those which would be unacceptable in Present-day Standard English are termed 'unregulated'. Table 2.1 illustrates the two systems (see also table 2.2 for actual examples of the sentence types from early Modern texts). Inevitably it has been necessary to simplify this account of one of the most researched and most contentious areas of historical English syntax. Fuller accounts of the modern and historical systems can be found in Tieken 1987, 31–4; Quirk *et al.* 1985 sections 2.49–51, 3.36–7, 12.21–6; Stein 1990.

The regulated system of auxiliary 'do' usage constitutes a rising variant over the period 1400–1700, conforming to the 'S' curve of linguistic change predicted in the Introduction. In 1400, auxiliary 'do' is hardly used, in the period 1500–1600 it is available in all sentence types, by 1700 it has been regulated in virtually all contexts: again, as stated in the Introduction, it can be assumed that younger, higher class, more educated, more urban individuals will be at the forefront of the change. For example, the assumption is that if regulation is the rising variant (as it will eventually be established as the standard), then Fletcher will show a more highly regulated usage than

11