The Shape of Early Modern English:
An Interview with Jonathan Culpeper on the *Encyclopaedia of Shakespeare's Language* Project

Iolanda Plescia

1.

Iolanda Plescia (IP): First of all, let me thank you very much for taking the time to answer some questions on an area of study – Shakespeare's language – to which you have made such a significant contribution. And also let me congratulate you on leading the team that has recently been awarded substantial Arts and Humanities Research Council (AHRC) funding to create an *Encyclopaedia of Shakespeare's Language*: just a few years ago, the introduction to the collection of essays on *Stylistics and Shakespeare's Language* (2011) that you edited with Mireille Ravassat pointed out that there is still comparatively little research being done on the language of Shakespeare, and the language of his time, as opposed to literary or new historical approaches. The word 'language', however, has often been used in the past within literary approaches dealing with rhetoric and meter, for example – starting with Frank Kermode's *Shakespeare's Language* (2000)[1]. How would you draw the distinction between those approaches and the contribution that modern linguistics and linguistic investigative methods have made in more recent years? Is there a gap to be bridged, and is 'literary linguistics' the right way forward in your opinion?

Jonathan Culpeper (JC): Thanks for your congratulations! It has been a hard and exceptionally long – as I will elaborate in my answer to Question 3 – road of preparation. I am so thankful that the UK's AHRC awarded the funding at the end of the day.

---

[1] See Jonathan Culpeper and Mireille Ravassat, eds, *Stylistics and Shakespeare's Language: Transdisciplinary Approaches*, Advances in Stylistics, London, Continuum, 2011; Frank Kermode, *Shakespeare's Language*, London, Penguin, 2000.

It is still true that comparatively little research has been done specifically on the language of Shakespeare. A trip to any university library will reveal a handful of books on 'Shakespeare's language', but shelves upon shelves full of books on every aspect of Shakespeare's works, their performance, and his life and context, not to mention the evolution of those works to the present-day. I would like to strike a positive note, and say that things are changing. But that note is more like one struck on a triangle in an orchestra rather than the timpani. Nevertheless, it is there. In the last ten years, the following book-length studies have appeared: Busse (2006), Crystal (2008, 2016), Hope (2010), Johnson (2014), Kizelbach (2014) and Ravassat and Culpeper (2011)[2]. And there have been a steady flow of journal articles and book chapters. It is not the case that all these books represent a single approach, such as literary linguistics or stylistics, though all of them make a contribution to understanding the language or style of literature, and Shakespeare in particular. What they have in common is a focus on the micro linguistic detail, an approach informed by current linguistic theory, and a method that leans towards being empirical, systematic and exhaustive (Crystal 2016 is the best exemplar of this method; I will briefly mention this work further below).

Is there a gap between literary approaches dealing with rhetoric, meter and so on and linguistic approaches, as discussed in the previous paragraph? A 'gap' suggests a clear demarcation between the two things. This, in my view, is not the case. It is important to remember that a range of approaches exists in both linguistics and literary studies, more so now than ever before. Although Noam Chomsky is perhaps the best known modern linguist, and his work is influential, especially in North America, it has very little to offer any study of literature. In contrast, the more applied and the more social kinds of linguistic work have much to offer. There we find notions

---

[2]   Beatrix Busse, *Vocative Constructions in the Language of Shakespeare*, Pragmatics & Beyond New Series, Amsterdam, John Benjamins, 2006; David Crystal, *"Think on My Words": Exploring Shakespeare's Language*, Cambridge Introductions to Literature, Cambridge, Cambridge University Press, 2008; David Crystal, *The Oxford Dictionary of Original Shakespearean Pronunciation*, Oxford, Oxford University Press, 2016; Jonathan Hope, *Shakespeare and Language: Reason, Eloquence and Artifice in the Renaissance*, Arden Shakespeare Library, London, Arden Shakespeare, 2010; Keith Johnson, *Shakespeare's English: A Practical Linguistic Guide*, Abingdon-New York, Routledge, 2014; Urzsula Kizelbach, *The Pragmatics of Early Modern Politics: Power and Kingship in Shakespeare's History Plays*, New York-Amsterdam, Rodopi, 2014; Culpeper and Ravassat, eds.

such as discourse, narrative, voice, style, power, gender and so on, notions that are familiar to many literary scholars. Consider a notion like rhetoric. It may be associated with more literary approaches, but linguistic pragmatics is in many ways a modern treatment of rhetoric, even dealing with the same kinds of figures (e.g. metaphor, irony, litotes). In fact, Leech's 1983 classic *Principles of Pragmatics* used the term "rhetoric" for the pragmatic phenomena he described[3]. Methodologically, the most exciting area to have developed over the last few decades is digital humanities, and this is an area that has a foot in both linguistics and literary studies. Some of the techniques here I recognise from my work in corpus linguistics (I will expand on this in various places below). But I also know of colleagues in literary studies doing pioneering work in digital humanities, including, for instance, interactive textual editions or GIS mapping techniques.

So, now the final question: is literary linguistics the way forward for linguistic and literary synergies? On the face of it, it should be. However, I started doing literary linguistics in the mid-1980s. Since then, it has hardly provided a golden bridge. It is the developments described in the previous paragraph that seem to be the unifying forces. Of course, one should also note that literary linguistics itself has not been immune to those very developments: literary linguistics today is not what it was thirty years ago. Old dichotomies (and egos!) seem to be dissolving as people discover that collaborative enterprises, especially at intersecting points of interest, have so much potential.

2.

IP: Can you describe the aim and scope of the *Encyclopaedia of Shakespeare's Language Project*? Does it aim, for example, to define every word in the Shakespeare corpus?

JC: Describing the meanings of a word in Shakespeare may seem easy, as, one might think, we can use the premier reference for historical English language, the *Oxford English Dictionary* (OED) (and its derivatives, including some specialist dictionaries of Shakespeare).

---

3   Geoffrey Leech, *Principles of Pragmatics*, London, Longman, 1983.

However, the OED prioritises an etymological approach to language and treats it as relatively static, rather than, as we intend to do, thoroughly examining how Shakespeare actually used language, with reference to its contexts and effects. Moreover, since the entries in the OED for Shakespeare's period are in large part determined by Shakespeare's language they provide no independent perspective on it; instead, they offer circularity.

The *Encyclopaedia of Shakespeare's Language Project* focuses on Shakespeare's linguistic usage in its early modern context. The guiding principle will not be etymology or editorial intuition but frequency. It needs to be stressed that what is proposed is not a traditional concordance of Shakespeare. Matters of frequency are used to reveal patterns of meaning and usage; they are not an end in themselves. Internal comparisons will reveal how Shakespeare's language dynamically varies across his works. For example, it will reveal whether certain words, meanings, structures, etc. are peculiar to tragedies, comedies or histories, to certain social groups (e.g. men/women) and to specific periods and sites of composition/performance. External comparisons with the language of Shakespeare's contemporaries will form an even more significant and innovative part of the research. The project will deploy techniques developed within corpus linguistics to analyse vast electronic corpora of historical texts. It will compare Shakespeare's usage with 321 million words in texts across all extant genres, 1560 to 1640. This will enable the discovery not only of specific usages characteristic of Shakespeare, but also the stylistic, discoursal and attitudinal flavour of particular items (e.g. whether certain words were considered colloquial, religious, courteous, offensive, and so on).

The major output from the project will be a two-volumed encyclopaedia. The first volume will focus on words (and, yes, every word in Shakespeare), including multi-word expressions and also their grammatical parts of speech. It will be based not only on corpus-derived information, as indicated above, but also on information extracted from commentaries (e.g. lexicons) from Shakespeare's time, as well as extant present-day research (especially for words that occur with little frequency, thus making the corpus-method less effective). The second volume will focus on patterns of words, patterns that constitute idiolects, sociolects and themes; or, in other words, characters, social groups (e.g. men/women), plays and groups of plays (e.g. tragedies, comedies and histories).

3.

IP: In your article for the 2011 volume *Stylistics and Shakespeare's Language*, which draws in turn on a first article you wrote in 2007, you laid out an "immodest proposal" for a "new kind of dictionary for Shakespeare's plays" (pp. 58-83)[4]. I am assuming that that article largely articulated the theoretical and methodological foundation for the *Encyclopaedia* project? How long have you been thinking about this project, and have you changed your mind about any aspect of it over the last few years?

JC: Yes, my 2007 article does outline the theoretical and methodological foundation for the project. Using the corpus-based method implies both a particular methodology for revealing meanings, and a particular theoretical approach to meaning. There is less reliance on the vagaries and biases of editors, and a greater focus on the evidence of *actual usage*. The question "what does X mean?" is pursued through another question: "how is X used?" But more than this, the *Encyclopaedia* is comparative, revealing not just the usage of words and other linguistic units in Shakespeare but also in the general language of the period. This way, we can tap into issues such as what is distinctive about Shakespeare's language, and, more particularly, how Shakespeare's language would have been perceived by his contemporary audience. For example, the play *Henry V* contains Welsh, Irish and Scottish characters. The words *Welsh*, *Irish* and *Scottish* do not appear in any Shakespeare dictionaries, presumably because their meanings are (erroneously) assumed to be transparent. A pilot examination I conducted with Alison Findlay of the usage of those words in over 100 million words written in Shakespeare's time revealed that: (1) the Welsh barely registered on the Elizabethan consciousness, being considered a harmless in-group, only noteworthy for their curious language, (2) the Irish were *wild*, *savage*, *rebels*, viewed positively only in relation to *Irish rugs* (an important colonial import), and (3) the Scottish, whilst also *rebels*, were respected for their political power.

However, I had the idea for the project well before 2007; in fact, slightly over twenty years ago. The fact that it has taken so long to get it off the ground has much to do with method, the data and tools

---

4   Jonathan Culpeper, "A New Kind of Dictionary for Shakespeare's Plays: An Immodest Proposal", in Culpeper and Ravassat, eds, pp. 58-83.

required. The problem twenty years ago was the lack of comparative data. In the early 1990s, the leading historical corpus of English was without doubt the *Helsinki Corpus of English Texts*, completed in 1991. This corpus amounted to 1.5 million words – an impressive figure in those days! Moreover, it had been put together with great care; it was reliable. But those 1.5 million words covered the period 730 to 1710. The section contemporaneous with Shakespeare amounted to less than half a million words, and is thus far short of what is required for serious comparative work. To solve the problem, I set about, with Merja Kytö, creating the *Corpus of English Dialogues*. The reason for the focus on dialogues is that this would provide an interesting comparison for the dialogues of Shakespeare's plays. This project soaked up ten or more years, taking on a life of its own, resulting not just in the creation of the dialogues corpus but also in publishing the various insights it afforded into early modern dialogues along the way.

In more recent years, I have been overtaken – in a positive way! – by other events, notably, the advent of a fully-searchable 1.2 billion word transcribed version of *Early English Books Online* (EEBO) (i.e. EEBO-TCP). For years, EEBO, which purportedly contains all early modern printed output, had been of limited value to linguists because the texts were only available as images, and language searches relied on OCR, with all its inaccuracies. Now, however, I have a 321 million word fully searchable corpus of texts written by Shakespeare's contemporaries. In addition, solutions, or at least partial solutions, have evolved for the various problems associated with the computational analysis of historical language data. For instance, early modern spelling variation had been a major stumbling block (e.g. the word *would* could be spelt *would*, *wold*, *wolde*, *woolde*, *wuld*, *vvold*, etc.). This problem has been largely solved by the *Variant Detector* (VARD), devised by scholars at Lancaster, especially Alistair Baron. This program regularizes a variant to a single, regular form. Furthermore, software for identifying parts-of-speech has progressed. The Lancaster-developed *CLAWS part-of-speech annotation system*, which works well for present-day English, has now been adapted for Early Modern English (though some more work will be necessary).

Over this period of time, I have not changed my mind about any of the fundamentals relating to the project. However, I certainly did not predict that a transcribed version of EEBO would be available and the opportunities that that would afford.

4.

IP: One of the trends of recent linguistic scholarship on Shakespeare has been to debunk myths surrounding his use of language (cf. Crystal 2008, Hope 2010, Elliot and Valenza 2011, and others[5]), most notably with reference to the number of neologisms he supposedly coined. In a recent talk Jonathan Hope pointed out that these myths derive largely from a Romantic notion of what an author is supposed to be[6] – do you agree? What do we gain by assessing Shakespeare's language in a more realistic way, and will the *Encyclopaedia* also help us to do this?

JC: Despite work by Crystal, Hope, Elliot and Valenza, and others, I have yet to see a full account of Shakespeare's neologisms. Providing such an account is not one of the central aims of the *Encyclopaedia* project. However, there will be many spin-offs from the project, and this will be one of them. In fact, I have already undertaken some work on this with one of the project team members, Sheryl Banas. We took the words that the *Oxford English Dictionary* records as being first cited by Shakespeare – they amount to around 1,400 items. We are now checking for ante-datings in EEBO-TCP. This is not a straightforward procedure. Even with the help of VARD, the spelling regularizer mentioned above, there is no guarantee that all spelling variants are catered for. We are proceeding cautiously, searching for multiple spelling possibilities. But our problems do not end here. Another problem is: what counts as a particular word-form? For example, it seems that Shakespeare was the first to use *acerbic*, a borrowing from Latin. However, in the First Folio (1623), it is actually written as *acerb*. So, should we count that as the first recording, or attribute it to a later recording by another writer, where it is written *acerbic*? Another problem is: what counts as a new word, a neologism? Some of our items seem near enough one-off or nonce creations – hardly evidence that Shakespeare is shaping the English language! The literature on word-formation takes currency

---

[5]   Crystal, *"Think on My Words"*; Hope, *Shakespeare and Language*; Ward E. Y. Elliott and Robert J. Valenza, *"*Shakespeare's Vocabulary: Did It Dwarf All Others?", in Culpeper and Ravassat, eds, pp. 34-57.

[6]   Keynote speech at the conference "A Great Feast of Languages: Shakespeare's Language and the Language(s) of Shakespeare's Time", Mid-Term SLIN Association Symposium, Sapienza University of Rome, 27-29 October 2016, organised by Donatella Montini and Iolanda Plescia. The same argument is also put forward by Hope in his essay published in the present issue of *Memoria di Shakespeare*.

into consideration (did Shakespeare kick-start an expansion of usage?) and also institutionalization (did early dictionaries and lexicons pick up Shakespeare's word and add it to their works). We are considering all of this. At the current time, we have looked at half of the 1,400 items. If this trend in our findings continues, we will conclude that around one third of these have a solid claim to be neologisms.

Jonathan Hope's idea that the Shakespeare neologism myth, and others, derives largely from a Romantic notion of what an author is supposed to be sounds entirely plausible. This is not something that I have researched, and it is not on the project's agenda. However, I will add that simple 'myth-busting' should never be on anybody's agenda. Language myths need not only to be exposed but also understood. They can tell us about the attitudes and ideologies that reflect them (and are constructed by them). The positioning of Shakespeare in Romantic thinking is Hope's contribution here.

As for what we gain by assessing Shakespeare's language in a more realistic way, I would prefer to phrase that as a more empirical way. Note I say 'more empirical'. I don't think anything can be entirely empirical or objective, as the subjectivity of the researcher will always come in at some point. What we are trying to do in the project is to (a) be guided as much as possible by the linguistic evidence rather than pre-conceived notions, and (b) encompass new kinds of evidence afforded by computers identifying patterns in vast collections of language data (in our case, principally, Shakespeare's entire works and those of his contemporaries in EEBO-TCP). This approach, we hope, will shed new light on Shakespeare in two ways. On the one hand, it should provide evidence to substantiate what we always thought but could not quite put our finger on exactly where the thought was coming from. On the other hand, it should provide evidence to substantiate what we had not thought about, i.e. a new thought.

5.

IP: At the same time, it is hard not to feel that Shakespeare is in some sense a 'creator' of language – but perhaps this means something very different in philosophical terms (as for example in Wittgenstein's definition of Shakespeare as "Sprachschöpfer"[7]) than in strictly linguistic terms. His creativity and in-

---

[7]   See Nadia Fusini, "Shakespeare: Playwright or 'Sprachschöpfer'?", *Memoria di Shakespeare*, 8: *On Authorship*, eds Rosy Colombo and Daniela Guardamagna (2012), pp. 95-118.

ventiveness can hardly be questioned, but they are getting harder to pinpoint now that we are letting go of the word-making myths. How much of what we perceive as ground-breaking in his use of language is due to our ignorance of the possibilities of early modern English?

JC: Undoubtedly, some of the early modern resonances of Shakespeare's language use are simply lost. This can happen on different levels. Regarding pronunciation, it has obviously changed, and so some rhymes and puns are no longer easily accessible. Regarding grammar, whilst particular structures were associated with certain social groups – men/women, high/middling/low rank (cf. Nevalainen and Raumolin-Brunberg 2003[8]), and so on – it is quite wrong to import present-day prescriptive notions into the assessment of Shakespeare. For example, multiple negation was not quite yet the fully stigmatized feature that it is today. Regarding vocabulary, words have often shifted their associations, as I illustrated with the words *Welsh*, *Irish* and *Scottish* above. The main aim behind the comparative aspect of the *Encyclopaedia* project is to enable us to capture those general possibilities and resonances of early modern English. This is not to say that we can only appreciate Shakespeare's creativity and inventiveness with such knowledge; much shines through without (see the following answer for an illustration), but it will provide a fuller appreciation.

6.

IP: Where do you think Shakespeare's linguistic 'greatness', if we can call it that, his inventiveness, does lie (syntax, versification, use of figurative language…)?

JC: The question hints that Shakespeare's linguistic 'greatness' might reside in one thing. Actually, even a half-decent writer should be multi-faceted in their writing skills. Shakespeare excels in what I think of as creative layering. Let me reprise an example I discussed at the beginning of Culpeper et al. (1998)[9]. I like this example because it achieves creativity in so many layers yet is not

---

[8] Terttu Nevalainen and Helena Raumolin-Brunberg, *Historical Sociolinguistics: Language Change in Tudor and Stuart England*, London, Longman, 2003.

[9] Jonathan Culpeper, Mick Short and Peter Verdonk, *Exploring the Language of Drama: From Text to Context*, London, Routledge, 1998.

amongst the Shakespearean linguistic examples that critics dwell on ad nauseam.

At the end of *Henry IV Part 2* Hal has succeeded to the throne to become Henry V. Falstaff, his hitherto disreputable companion, is overjoyed, imagining all sorts of privileges. Outside Westminster Abbey, Falstaff meets the king:

> Falstaff
> My King! My Jove! I speak to thee, my heart!
> King
> I know thee not, old man. Fall to thy prayers. (V.v.46-47)[10]

How does Shakespeare make one of the greatest snubs in the history of literature so effective? Note the parallelism built out of a repeated form of address: *My King! My Jove!* [...] *my heart!* This reinforces Falstaff's emotional gush. Note also that Falstaff's line is metrically regular: five stressed syllables alternating with five unstressed syllables – a perfect iambic pentameter. And note that the grammatical boundaries and the punctuation boundaries coincide with the metrical units. This all works to enhance the regularity of the line and seems to enhance Falstaff's joyful enthusiasm. Semantically, Falstaff asserts *I speak to thee*. What does he mean by that? Answering that question is a pragmatic issue, a matter of meaning in context. It is obvious to one and all that Falstaff is speaking to the king. He provides unnecessary information. The exchange of information in conversation orientates to the Cooperative Principle (Grice 1975)[11]. Giving unnecessary information flouts one of the Cooperative Principle's constituent maxims, the maxim of Quantity, and triggers an inference. What are we to infer? Given Falstaff's expectations about wealth and privileges, what he means is acknowledge me, acknowledge your old friend Falstaff. This is fairly obvious stuff, but the point for creativity is that Shakespeare does not have Falstaff simply say 'acknowledge your old friend'. A further pragmatic issue concerns Falstaff's choice of referring expressions for the king. There is nothing remarkable in the choice of *king*,

---

but with *Jove* he implies, flatteringly, that the king is some kind of god, and with *heart* that the king is so dear to him that it is as if the king is his very heart. The mechanism for both these implications is again the Cooperative Principle, but this time the maxim of Quality (it is not literally true that the king is either Jove or Falstaff's heart). There are also some sociolinguistic issues here. Falstaff presumes a very close social relationship with the king; he fails to pay the respect the king's new position of power demands.

Falstaff's line should not be considered in isolation – it is part of a conversational exchange. Contrast Falstaff's line with the king's reply. The regularity is destroyed half way through the line, as we encounter two stressed syllables followed by two unstressed syllables:

> FALSTAFF:
>   x   /   x /   x /   x /   x   /
> My King! My Jove! I speak to thee, my heart!
> KING:
> x   /   x /   x   /   / x x   / x
> I know thee not, old man. Fall to thy prayers.

This change reinforces a change in mood, a difference in attitude: the king is not overjoyed to see Falstaff. Semantically, far from recognizing a close relationship, he states that he does not even know him. This is obviously untrue: he flouts the maxim of Quality. What is the implication? Hal had said in a soliloquy at the beginning of *Henry IV Part 2* that when he is king he would turn away from all his former roguish friends. Clearly, the king implies that he wishes to have absolutely nothing to do with Falstaff, as if he does not know him at all. Interestingly, although the audience can work out this inference, Falstaff fails to pick it up. After these lines we find out that he assumes that the king will speak to him more positively in private. The foundation of the snub is that they have decidedly different conceptions of their relationship. In terms of social distance for Falstaff they are familiar, for the king they are distant. The contrast is reinforced by linguistic politeness issues. Not only does the king deny that he knows Falstaff, but he exercises linguistic power, through a direct command, *Fall to thy prayers* (perhaps punning on *fall* and *Fal*staff), and chooses a potentially insulting referring expression, *old man*.

The point about this example is that all this creative linguistic layering is packed into a mere two lines. Of course, it is not as if other

writers have failed to construct moments of such layering. But, to use a musical analogy, Shakespeare is no one hit wonder. The most renowned writers – Shakespeare, Chaucer, Austen, Dickens, and so on – tend to have large bodies of surviving work. A large volume of work seems to be a necessary condition of greatness, but not a sufficient one. In addition, a high proportion of that work must be great. The contemporary composer, Howard Goodall, made this point about the Beatles in the documentary *The Beatles: Eight Days a Week* (2016). What makes the Beatles stand out from other pop groups is that a higher proportion of their works are good. This, he suggests, is what makes Mozart stand apart from Schubert: Schubert composed about 800 works, of which only around 200 are highly rated, but most of Mozart's output is highly rated. Shakespeare not only managed moments of creative linguistic layering, but repeatedly did so. That, in my view, is where a claim to his linguistic greatness lies.

7.

IP: Another recent movement in Shakespeare studies that is linguistically informed is the one focused on reconstructing Original Pronunciation, which has led to a number of OP theatrical productions as well[12]. What do you think of this trend? While I am not entirely sure about the stage appeal of such experiments, I can certainly appreciate the way OP elucidates textual cruces and editorial dilemmas, clarifies rhyme schemes, and helps understand puns that are lost on contemporary ears. Will the *Encyclopaedia* contain some form of phonetic transcription of Shakespeare's words?

JC: The Crystals – both the renowned linguist David Crystal, and his son, the professional actor cum linguist, Ben Crystal – have been prominent in leading the Original Pronunciation (OP) movement. As you say, OP can help illuminate rhymes, puns and so on. But one might argue that these things occur sporadically: would it not be more economic to just focus on these cases and elucidate their pronunciation, rather than invest a huge amount of time in reconstructing plausible original pronunciations for every word in Shakespeare? One of the arguments that I have heard the Crystals make is that

---

[12]   An interesting and informative video by David and Ben Crystal on Original Pronunciation may be viewed at https://www.youtube.com/watch?v=gPlpphT7n9s (last accessed December 2016).

OP lends a rather different quality to performances, compared with 'normal' modified-RP performances. Ben Crystal claimed OP performances were more earthy, more visceral. From his demonstration, I could sort of see what he meant. In OP, Shakespeare is free from at least some of the social attitudes that perception of present day accents triggers.

As to whether the *Encyclopaedia* will contain phonetic transcriptions of each word, that is something I considered doing, but rejected. I rejected it for two reasons. One is that the *Encyclopaedia* is centrally underpinned by the corpus-based method. Reconstructing plausible historical phonetic transcriptions involves educated guesswork, drawing on original spellings, rhymes and puns, contemporary commentaries regarding pronunciation, and also assumptions about the paths of sound change. Aside from working with original spellings (see my comments above about spelling regularization), postulating hypotheses about pronunciation from these kinds of evidence calls for very different methods. The other is that it has already been done, and done very well. The *Oxford Dictionary of Original Shakespearean Pronunciation* (Crystal 2016)[13], a monumental work in many ways, supplies a possible original pronunciation for every word in Shakespeare, along with supporting evidence.

8.

IP: It seems that the *Encyclopaedia* project has far reaching implications that potentially involve several different levels of linguistic investigation at the same time. Thinking of some of your other fields of expertise, I wonder how the new project will illuminate, for example, characterisation, or the discourse markers that Shakespeare uses to construct a play text that is made up of a certain amount of 'realistic' speech and dialogue patterns.

JC: Volume 2 of the *Encyclopaedia* focuses on patterns of words. More specifically, it describes the patterns of words that constitute characters, socially defined character groups (e.g. artisan, school teacher), plays and play genres (i.e. tragedy, history, comedy). Some Shakespearean dictionaries contain non-linguistic descriptions of characters and plot summaries. The *Encyclopaedia*, however, will provide

---

[13]   See note 2.

a description of the linguistic idiolect or thumbprint of every major character. This will be done by conducting a statistical comparison between the vocabulary of one character and that of the other characters in the same play, in order to reveal characteristic words, i.e. 'keywords'. This approach is very much a continuation of a line of work on characterisation that I began in the 1990s. More specifically, it replicates the kind of work I did on characters in *Romeo and Juliet*, first published as Culpeper 2002[14]. Romeo's most characteristic words are predictably *beauty* and *love*. For Juliet, the less predictable results are *if* and *be*. Although the results for Juliet are surprising, they can readily be explained by qualitative analysis of the text. They reflect her anxieties and worries about Romeo's intentions and welfare (e.g. "*If* he *be* married, / Our grave is like to be our wedding-bed", I.v.134-135). A reading of the play would obviously have resulted in an understanding of Juliet's anxieties and worries – it would not necessarily have led to the identification of the linguistic source of that very understanding.

As for discourse markers, the corpus approach, which the *Encyclopaedia* adopts, treats all words within a corpus or body of data (a word being defined as any character or series of characters bounded by spaces or punctuation marks). Therefore, discourse markers such as *well*, *why*, *fie* or *pish* have equal status. This is a departure from other, earlier treatments of the language, which tend to exclude items that they considered less important. And, discourse markers, which partly derive their meanings through the specifics of their contexts, are often considered less important. One particular set of discourse markers that I have been interested in is primary interjections, or what I and Merja Kytö termed "pragmatic noise"[15]. They include items such as *ah*, *oh*, *ho*, *ha* and *fie*. You may be thinking that these are of rather limited value, but let me prove their richness, by illustrating how *ah* has five distinct meanings in Shakespeare's texts:

---

14   Jonathan Culpeper, "Computers, Language and Characterisation: An Analysis of Six Characters in *Romeo and Juliet*", in *Conversation in Life and in Literature: Papers from the ASLA Symposium*, Uppsala, Association Suedoise de Linguistique Appliquée, 2002, pp. 11-30.

15   Jonathan Culpeper and Merja Kytö, *Speech in Writing: Explorations in Early Modern English Dialogues*, Cambridge, Cambridge University Press, 2010.

(1) Speaker attitude/state: sorrow, emotional distress

DESDEMONA
To whom, my lord? With whom? How am I false?
OTHELLO
Ah, Desdemona, away, away, away.
DESDEMONA
Alas the heavy day: why do you weep?
Am I the motive of these tears, my lord? (*Othello*, IV.ii.41-44)

(2) Speaker attitude/state: pity

GLOUCESTER
Canst thou blame him?
His daughters seek his death. **Ah**, that good Kent,
He said it would be thus, poor banished man.
Thou sayest the King grows mad; I'll tell thee, friend,
I am almost mad myself. (*King Lear*, III.iv.158-162)

(3) Speaker attitude/state: surprise, realisation

Enter Adriana and Luciana.
ADRIANA
Ah Luciana, did he tempt thee so? (*The Comedy of Errors*, IV.ii.1)

(4) Discourse marker: preface to the correction/rejection of the previous speaker's proposition(s), emotions or actions

MENAS
These three world-sharers, these competitors,
Are in thy vessel. Let me cut the cable,
And when we are put off, fall to their throats.
All then is thine.
POMPEY
Ah, this thou shouldst have done
And not have spoke on't. In me 'tis villainy;
In thee 't had been good service. (*Antony and Cleopatra*, II.vii.70-75)

(5) Discourse marker: reinforcing elicitation

LEONATO
All thy tediousness on me, **ah**?
DOGBERRY
Yea, and 'twere a thousand times more than 'tis, for I hear
as good exclamation on your Worship as of any man in the
city, and though I be but a poor man, I am glad to hear it.
(*Much Ado About Nothing*, III.v.22-26)

9.

IP: You are also the author of a history of English manual with Routledge[16]. As someone who teaches HEL, I very much appreciate the organization of the book around themes and problems in language development rather than mere chronological order, and so do students! Do you see the *Encyclopaedia* as becoming a useful tool in HEL teaching as well, since it changes the approach from philological-etymological to a language-in-use methodology?

JC: I certainly hope that the *Encyclopaedia* becomes a useful tool in various kinds of teaching. The *Encyclopaedia* will eventually be available in various formats, tailored to various audiences. The two-volumed paper version will incorporate most of our findings. Our challenge will be to make sure that different kinds of reader/user can navigate the contents and extract what they want. The *Encyclopaedia* will become available in electronic form, probably through Bloomsbury's Drama Online webpages. This will have the merit of being quite dynamic, allowing people to choose and combine options from drop-down menus. Finally, there will be an app, which contains a 'lite' version of volume 1, the Dictionary. This, I hope, will have wide application in classrooms. One of the things that we are currently doing is investigating exactly what university students find difficult in Shakespeare texts, so that we can tune our work to help them.

In addition, you are right to allude to our language-in-use method being more pedagogically useful than the philological-etymological approach. As our approach is thoroughly based on actual usage, it is likely to be more relevant to the reader. One particular point to mention here is that in volume 1 the senses for any particular word are ordered in terms of frequency of use. This means that the first senses a user sees listed will be the ones that he or she is more likely to have encountered.

10.

IP: I wonder also, given the current trend in authorship attribution studies to look at 'small words', or function words, if the *Encyclopaedia* will offer tools that are useful to identify expected Shakespearean 'patterns' as well. (And as a

---

16   Jonathan Culpeper, *History of English*, London, Routledge, 1997 (rev. edition 2005).

side question, since it is the hot topic of the day: have you formed an opinion of the current debate on the Marlowe-Shakespeare connection?)

JC: Let me spell out the position of our project in relation to authorship attribution studies. Corpus linguistics does share some methodological features with attribution studies. At a very basic level, both typically use computers, statistics and electronic texts. Also, attribution studies typically compare a target text (of controversial or unknown authorship) with some other texts (of known authorship). There will be a comparative aspect to the *Encyclopaedia* project: for every word in volume 1 and most themes in volume 2, we will be looking to see if the meanings and usage in Shakespeare match the writings of his contemporaries.

However, there are major differences too. Attribution studies are generally quite narrowly focused on the question of whether a certain text can be attributed to a certain author. Corpus linguistics addresses a much wider range of research questions. These include, for example, the meanings of words, expressions and grammatical structures, including their contextual associations. This is not something that authorship attribution studies typically engage in. They are looking at patterns in any formal units (typically words or groups of words, but also spellings, grammatical structures, etc.); the meaning of those units is not a concern.

I have not done any research on the issue of the Marlowe-Shakespeare connect, and so I have not much to say here. However, I do note that many of the techniques that they used (e.g. n-grams) are the bread-and-butter of Corpus Linguistics.

11.

IP: And now for a few more general, wide ranging questions. In a 2010 survey of the year's contributions to Shakespeare studies in *Shakespeare Survey*, Julie Sanders wrote: "The linguistic turn in Shakespeare has shown itself at various moments in recent years but the movement has never been sustained. The separation between linguistics and literary or performance studies modules that still pertains in many departments in UK and US universities has tended to create discrete debates which only intermittently encounter each other at conferences and in the publishing context" (p. 402)[17]. First of all, would you

---

[17] Julie Sanders, "This Year's Contribution to Shakespeare Studies 1 – Critical Studies", *Shakespeare Survey*, 63: *Shakespeare's English Histories and Their Afterlives*, ed. Peter Holland (2010), pp. 388-405.

accept or challenge this statement? Do you think the AHRC grant can be taken as a sign of a definitive "linguistic turn" in Shakespeare studies? I find it significant that the award marks, in a sense, the quatercentenary year in a very material, tangible way, by asking us to pay attention to the material conditions of language (just as we have learnt to pay attention to the material conditions of the Elizabethan stage, etc.).

JC: Depressing though it may be, I would largely accept Sander's statement. Of course, one should ask: should linguistics and literary form studies always be welded together in the pursuit of Shakespeare? The answer is most certainly no. Shakespeare's language and the modern linguistic study of it is just one facet of the whole. But the point I was making earlier in answer to your first question is that this one linguistic facet seems to have been seriously overshadowed by everything else. It would be nice to think that the AHRC grant is a sign of a "linguistic turn". However, one grant, large though it is, is unlikely to result in a paradigm shift. I am more optimistic about this being achieved by some of the developments, some of the bridges, I alluded to in my answer to your first question, specifically, pragmatics/discourse analysis and digital humanities. In this respect, it might be worth noting that the *Encyclopaedia* project partly belongs to the digital humanities revolution.

12.

IP: I have a wonderful memory of a seminar on "Shakespeare's Language and Style" that you convened at Lancaster University in 2012, where a really diverse range of approaches to Shakespeare's language were all welcome. A key word of your co-edited 2011 collection is "transdisciplinarity". How do you view collaborative work across literary and linguistic disciplines and would you set any ground rules, for example?

JC: In the social sciences, transdisciplinarity (or what is often referred to as interdisciplinarity or multidisciplinarity) has become extremely common. In fact, the key UK government funding agency for the social sciences, the Economic and Social Research Council (ESRC), actively favours such research. My sense is that transdisciplinarity in humanities research lags somewhat behind that of the social sciences. It is, however, gathering steam, largely thanks to the kind of bridging developments I have already referred to. Specifically with respect to transdisciplinarity and the *Encyclopaedia* project, a key person is Ali-

son Findlay. Alison is a co-investigator. She is based in the English Literature department, and is a notable expert on Shakespeare. There are numerous occasions where Alison has provided additional insights to linguistic issues and results, especially with respect to interpretation.

As for ground rules, I don't know whether I would call them ground rules, but I would certainly have two hopes. One would be that people make time to hear what people from another discipline have to say. Ultimately, it may not be one's cup of tea and one doesn't pursue it any further, but at least one should hear them out (and one might be pleasantly surprised, interested and so on!). The other is that ego should be set aside, as I already hinted in my answer to question one. Transdisciplinary work is not a matter of convincing the other side that your approach is the 'right' one, but an appreciation that two or more parts make a greater whole.

13.

IP: And finally, when can we expect the *Encyclopaedia* to be published?

JC: The project finishes in May 2019, and the manuscript will be submitted one year after that. Allowing time for small delays and a year for production, we can reasonably expect publication of the *Encyclopaedia* in 2021 or 2022.