The artful science of cognitive literary study

Peter Stockwell

The ghosts in the world

More people in the world have abilities in two or more languages than are monolingual. Of those who are literate, this means that they have an ability to see the literatures of more than one culture, and so for the majority of the world’s readers, the observation that literature is a matter of being able to process language is not a controversial point. Further, the notion that the proper scholarly study of literature as an institution and a practice would be necessarily better if the literary critic had a decent training in linguistics should also be uncontroversial. It is unfortunate, then, that the university study of literature has been most well-funded and most developed in those monoglot countries of Europe and especially the United States where it is common to find literary critics and applied linguists in separate and distinct departments. The level of discourse about language in the literature departments and the scholarly community of the West is woefully amateurish, poorly-informed and predicated on philosophical fashion rather than any proper engagement with the arts and sciences of language.

When literary critics discuss language, they generally rely on an understanding that is completely outdated, for the most part deriving from misapplications of early 20th century language study. Most of the insights of modern linguistics (from the late 1950s onwards) have found little usage in the broad field of literary scholarship. Literary critics prefer to speak of Language in the abstract, or language as a lazy term when they actually mean something more specific such as register, semantic or syntactic choice, or style. When it comes to the contemporary linguistics of the last couple of decades, the field is even more bleak: there is barely a passing understanding of the great insights brought about in sociolinguistics, psycholinguistics, corpus linguistics, cognitive linguistics, and the innovations developed by language teachers working broadly in applied linguistics across the world. Those scholars who have been aware of these changes have generally discussed literature in the margins: in language classes for the most part, in Education schools and departments of applied linguistics; developing within literature institutions only under the general term of stylistics.

Meanwhile in departments of literature across the world and in all cultures, literary study has largely come to mean the study of literary history. Following Jameson’s (1981) dictat to ‘Always historicize!’ , the pattern of practice in literature departments from the US to China has been to fall into a paradigm of historiography. Literary scholars tend now to define themselves and are appointed to university posts in terms of their ‘period’: you are a Renaissance scholar, a medievalist, a Romanticist, working in late 19th century novels, and so on. It is rare, these days, for anyone to define themselves by method or approach. The literary work itself is often either lost in the general discourse of cultural studies or is treated mechanistically as the archaeological trace of authorship, editorialising, transmission and the history of writing. Both of these remove the practice of literary scholarship from anything that resembles the practice of literary reading in the population at large. The academy is interested in professional readings rather than readings in the wild by natural readers. As a result, the practices of literary scholars are simply incomprehensible to the vast majority of the literate world.
In spite of the millions of living, breathing readers amongst our fellow humanity, literary scholarship is more concerned with ghosts: authors long dead, the cultures that sustained and produced them. In fact, authors are doubly ghosted, treated not as human voices communicating across time but fossilised out of texts and rendered as historical curiosities. Like many others, I have had my photograph taken on Juliet’s balcony in Verona, read the Decameron on Boccaccio’s hilltop in Tuscany, had a beer in Hemingway’s café in Paris, sat at Dickens’ desk and walked around the English country garden of Thomas Hardy’s cottage. But when I was doing these things, I recognised that I was a tourist on holiday, not engaging in literary scholarship. Too much that passes for research in literature departments is little different from this sort of tourism – concerned with the minutiae of authors’ lives, their presumed motivations, and their biographical activities around their writing. And the discussion of this cultural study generally operates at a non-technical and rather conversational level of discourse. This sometimes makes for interesting history, but it is not genuine literary scholarship, because this time it is the current reader who has become the ghost intruding into the past.

Instead, we should be looking at what literary works do in the world. To be sure, some of that will involve exploring readerly responses over history, but we should put history in its proper perspective. It is perverse that historicism prefers to historicise every period except our own present, and it historicises every other critical approach without noticing its own historical place as an ephemeral and rather odd blip in the thousands of years of the analysis of literature. From Aristotle of Greece to Liu Hsieh of China, older traditions have emphasised the close analysis of the language of literature, with matters of meaning, aesthetics and the ethics of authors, characters and readers not allowed to be separated (see Hogan 1996). After a brief hiatus of a few decades, it seems that literary scholarship is beginning to burn out its narrow reliance on mere history and is ripe once more for a discussion of the humanity of literary reading.

However, we have been somewhere else in the intervening years, and we cannot simply pick up pre-Theoretically where we left off. The major change in the state of human knowledge in recent decades has been our advances in understanding our own minds. Though still in its infancy, modern cognitive science has given us more precise insights into the processes, effects and experience of being a conscious person in society than in all the previous millennia. In literary scholarship both old and new, readers as well as authors have for too long been treated as ghosts: idealised, implied, abstracted, rarefied into geniuses, spirited away as books and ideas, or left vague, ill-defined and unexplored. Those times are changing.

The range of cognitive poetics

Cognitive poetics is the application of cognitive science to literary reading. This simple statement covers a whole range of activities and different emphases. My own work, for example, comes out of a British and European tradition of stylistics, which is primarily concerned not only with the style of literary texts but with what we might once have called the contextual matters of situation, readerly experience and purpose. The field owes its origins to Russian formalism in the 1920s, French structuralism and American New Criticism and German hermeneutics and stylistics in the 1950s. The revolution in modern linguistics from the 1960s onwards gave stylistics its first toolkit for the systematic analysis of literary texts. Over the decades, it has expanded this toolkit as developments appeared in pragmatics, sociolinguistics and discourse analysis: each of these offered a wider possibility for the
The stylistic tradition is central to cognitive poetics in my work. Colleagues have even developed the term cognitive stylistics to capture this emphasis (Semino and Culpeper 2002). Cognitive poetic work that includes a close attention to the linguistic patterning of textuality nevertheless treats literature as a heteronomous object (in Ingarden’s (1973) terms): that is, as an object that consists not simply of the autonomous existence of a text but which arises from the interaction with an observing consciousness. In other words, text is important, but literature does not exist unless it is read. This emphasis means (again in Ingarden’s terms) it is the literary work rather than the literary text alone that is the object of investigation.

Accounting for the interactive and invisible process of a reader reading a text and the effects that are involved is not an easily definable object. For cognitive poetics, this places the question of textual evidence at the heart of the discipline. The different approaches within the cognitive account of literature can be traced to the different forms of evidence that they each pursue. In the stylistic tradition of cognitive poetics, for example, a reading of a literary work is usually presented with a great deal of pinning down to the linguistic features. Sometimes, this can look like quite traditional stylistic analysis, but this observation is rather an indication that much traditional stylistic analysis was implicitly readerly and attentive to effects already. For example, an account of a love poem might focus on the linguistic features of addressivity, such as pronouns, naming patterns, informality of lexis, the way that the relationship between speaker and addressee assumes closeness and intimacy, and so on. Most of these are matters of deixis which can be identified precisely and their patterns sketched in a text, and this sort of analysis would not have looked out of place in a stylistics account from 30 years ago. However, the mental modelling of an alternate voice and persona, and the monitoring and maintenance of that viewpoint, and the sense of empathy or antipathy between reader and poetic voice, and the sense of closeness or distance – all these are matters of cognitively-based effects. A description of deixis that is located in the sorts of activities a reader is engaging in requires a recognition not only of the textual features that are present but also a simultaneous recognition of how those features are being used by the reader. The autonomous patterns in a literary text are part of its textuality, and the heteronomous quality of textuality as felt by a reader is its texture; the subtle distinction between the two captures the different emphases of stylistics proper and cognitive poetics (Stockwell 2009).

Cognitive poetics in the stylistic tradition makes its claims to validity on the basis of this sort of textual and textural evidence. There is an external claim to validity in the sense that the model or framework of analysis usually originates in the discipline of linguistics where it has been developed with some evidential or observed basis. The application of a tested language model to the language of a literary text thus has a degree of established validity. Secondly, there is an internal validity in that the framework has to be applied precisely. That is, the analysis itself gains in validity if it is an accurate parsing or pattern-identification of the text by the framework. At this level, it is possible simply to be wrong by misidentifying something. Lastly, there is a consequential validity in that the clear and accessible analysis is open for other readers to verify the account against their own perceptions. Taken altogether, these three aspects of validity allow a certain degree of introspection in cognitive poetics,
since the risk of idiosyncrasy or pure subjectivity is mitigated by a common currency of
terminology and framework, and by a generally observed tendency towards openness and
clarity in description.

Not all researchers working in cognitive poetics agree that this form of mitigated
introspection remains a valuable form of exploration. There is a very strong empirical
tradition in cognitive poetics which aims to establish a more laboratory-style methodology for
the discipline. For example, it would not be enough for a scholar to demonstrate the nature of
the interaction between textual features and cognitive processes for himself; he would also
have to replicate the effects in a sample of informant-readers under some level of controlled
situation. This methodology can be relatively qualitative, taking the form of focus groups,
questionnaires or self-reporting protocols that aim to gain a rich sense of readers’ responses;
such responses are then theorised and accounted for in terms that stand as a cognitively
plausible account. For example, there is a currently developing interest in the use of reading
groups as a form of exploratory data. Groups of mainly non-academic readers often gather in
people’s homes, local libraries, schools or cafés to discuss (usually) novels that they have
read. Many of these informal groups keep notes and post their reactions as blogs. The
researcher then has access to natural data that was not generated by the research, or is able to
participate in the group as an involved researcher. It is possible, in this emerging tradition, to
see a blended methodology that draws on discourse analysis and sociolinguistics (see

Alternatively, there is also a strong empirical tradition that is relatively quantitative. This
work identifies specific features or effects in a literary reading, and readers’ responses to
tightly controlled laboratory tests are collected, tabulated and rendered for their statistical
significance. Sometimes studies such as this use objectively measurable data such as facial
and gestural reactions, eye-tracking equipment, and other physically apparent behaviour, and
sometimes the results are in the form of marking-up a text, or self-report by informants (see

There is a very strong, even foundational tradition in cognitive poetics deriving from the
pioneering work of Reuven Tsur (1992) that takes this empirical approach to its most
materialist form. Inevitably, it is this collision of neuroanatomy and literary scholarship that
has been popularly presented in the media as the emergence of ‘literature and the brain’
studies. For a taste of the half-understood nature of this, and an equally ill-informed criticism
of it, see Byatt (2006) and Tallis (2008). Most researchers working in cognitive approaches to
literature insist on placing a conception of the mind between the brain and the book. It is
possible to record the brain activity of a reader in an MRI scanner, but the resulting images of
the reader’s brain will not tell you anything interesting about that literary work. Such findings
might make for interesting data about language processing, or about creative engagement in
general, but they are so far removed from the singularity of the particular text as to be
irrelevant as literary research. Such studies might be good brain science but they are not good
literary study.

The charge of reductionism is often laid against cognitive poetics. It is possible to identify, for
example, similar schematic patterns in reading all literature, or all novels, or all lyric poetry,
or all action scenes in a narrative, and so on. It is possible to notice a particular conceptual
metaphor (life as a balance, or emotions as containers, for example) in texts as diverse as
Shakespeare’s Macbeth and Paul Auster’s The New York Trilogy. A researcher might argue
that it is interesting to note similarities across great works of literature, but a critical view
might point out that the differences between Shakespeare, Jane Austen and James Joyce are lost in this general reduction to common patterns and principles. The accusation of reductionism, of course, applies also to structuralism, stylistics, canonisation by genre, the historical periodisation of literature, and any other approach that seeks to find patterns across and between literary texts. In fact, the charge of reductionism can be levelled at any human investigation of anything at all, since the purpose of theorising by definition is to move towards generalisable patterns out of the messiness of particular data (see Freeman 1996).

Nevertheless, there is an emerging tradition of evolutionary criticism, or neo-Darwinism, that has been associated (wrongly in my view) with cognitive approaches to literature (see Boyd 2009, Boyd et al 2010, Carroll 2011). This tradition turns the approach around: instead of saying, for example, that our evolutionarily evolved capacity for identifying other people helps us to understand what a reader does with characters in a Jane Austen novel, the evocritic interprets how we deal with these characters as evidence for how our brains evolved in pre-history. This is a sort of paleopoetics (Collins 2012) that is of course both highly speculative and ultimately untestable and unprovable, even if it is fascinating as a thought-experiment (Gottschall 2008, Dutton 2009, Swirski 2010).

When it is handled well, this sort of cognitive literary studies (to give it its most broad name) has a lot to say about literariness, about the connections between humans across cultures, and thus about our common human condition. It can allow the expression of a genuine passion for character and narrative that has often been missing in recent years from literary criticism (Vermeule 2010). At its worst, however, it can simply be the exercise of some fashionable but vacuous generalisations. There can be a tendency to treat educated liberal American responses as the default model for all human responses, so that actual cultural, ethnic and ideological differences are blurred away (see again Hogan 1996). There is a tendency to focus on readerly consensus and patterns that denote psychological commonalities (and this is refreshing against a literary critical landscape that privileges eccentric innovation and creative but peculiar newness), but there is often a corresponding loss of the idiosyncracies and subjectivities of readers. Lastly, there can be a poor understanding of the scientific method when arts and literature academics attempt to draw on methodologies from psychology and neuroscience, so that key issues of replicability, falsifiability, the nature of proof, and the nature of evidence are lost.

The over-riding danger is that cognitive literary studies become merely a critical theory in the hands of people who do not really understand cognitive science. I believe cognitive poetics is more of a method than a critical approach: I do not think it is a critical theory at all. Without going as far as to say that some of its tenets are demonstrably true in a natural science sense, I would at least say that the object of analysis is present and available in the world to be discovered. To this extent, it is the exploration of something real, not the poetic construction of its own object of investigation. Cognitive poetics is a form of experiential realism. It is realist because it treats a natural phenomenon in the world (literary reading and its effects), but it is experientialist because it recognises that our only access to that world is through our experiencing senses of perception and cognition. It seems to me that if cognitive literary studies is regarded as a critical theory, then it will rise and fall very quickly according to the whims of intellectual fashion, and a moment of real advance in human understanding will have been lost.

My antidote to all of these risks around the field is – perhaps unsurprisingly – to insist on the centrality of textuality and texture (see Gavins and Stockwell 2012). In the heteronomous
event of a reader reading a literary work, the textual patterning of the work is traceable and evident. The textual pattern is a measurable commonality across different readers, a point of stability that offers the key to further explanations of reception and effect. Systematic analytical attention to the stylistics of the text retains the singularity of the literary work at the centre of the study. Even if common patterns between this text and that text can be discovered, close cognitive stylistics will always emphasise the singular qualities that make this text different from that text. An emphasis on textual evidence also ensures that claims made about literature are open to agreement or falsifiability, are open to verification and checking for accuracy, and are fundamentally testable in order to advance the discipline.

The principles of cognitive poetics

Cognitive poetics draws most of its foundational principles from cognitive science in general. Some of these are general scientific principles and some are specific to the non-scientific domain of literary art. The latter require some adjustment in the way a scientific method is applied outside the natural sciences: for art, the object of investigation is heteronomous and dependent on consciousness.

The most important principle of cognitive science is the principle of continuity. This applies to several aspects of the field. Keeping continuity as a principle entails remembering that categories are not distinct labels but blend into each other and can reconfigure themselves in different situations. For example, the distinction between mind and body which has been implicit in Western philosophy for most of its history and explicitly since Descartes is refuted by cognitive science. It is clear that the human mind functions not only because of brain processing but also because that brain is carried in a body configured in a characteristically human way, living in a social culture usually understood as a common human condition. Contrary binaries that follow from this Cartesian dualism are also refuted: nature and artifice; reason and emotion; meaning and feeling; science and art – it is increasingly apparent that these distinctions drawn by the European Enlightenment do not match up with the ways that humans seem to think. The mind is an embodied mind (Lakoff and Johnson 1999), subject to experiences, memories and interactions that determine our consciousness.

Secondly, there is a sense of continuity in our perception of evolution. The human language capacity has not been around long enough (up to 100,000 years in a form we would recognise as full language ability), and the human literary capacity for experiencing writing has not been around long enough (only 11,000 years) for much specific evolutionary adaptation to have taken place. Cognitive language capacity is almost certainly built on top of existing structures that allow us to see, hear, distinguish things, understand similarities and differences, use tools and avoid walking into objects. Certainly with writing, it is probably safe to find continuities between, for example, the way that figure and ground in the visual field and in the virtual field of a fictional landscape work (Stockwell 2009). World-building in novels and poems is highly likely to be the same as maintaining a mental representation of remembered places. Characters in fiction are people in every respect except the trivial sense that they are not real. There is a theoretical linguistic debate over whether there is such a thing as a ‘language module’ in the brain, but we can with some certainty say that there is not a ‘literature module’ which treats fictional worlds and dramatic lyrics as special and unique. Instead, our encounters with literature are extrapolations of our encounters with life in general.
A continuity is assumed between language and thought. Since our only access to thought is through language, and any experience of self-consciousness can only be a form of representation, then understanding how language works is at the heart of understanding thought. In cognitive linguistic terms, this line of argument has led to a central focus being placed on conceptual metaphor, schematisation, and idealised cognitive models that stand as our personal building blocks of experience (Lakoff 1987). The importance of metaphor in cognitive linguistics was established from an early stage (Lakoff and Johnson 1980), and it continues to be a central concern (Fauconnier and Turner 2002). It is of course obvious that the cognitive linguistic emphasis on metaphor has an immediate application in literary works, where metaphors of various types can be observed from the micro-level (lexical blends, similes, copula metaphors, grammatical metaphors) right up to the text and discourse levels (allegory, symbolism, emblematic patterns and other parabolic (Turner 1996, 2006) features of satire, irony, social observation and moral didacticism.

Key features that have been of central concern in cognitive linguistics include the gestalt distinction between figure and ground, the understanding of categorisation as always being defined by prototypicality effects, and the importance of rich schematic framing of experiential context. The first of these is apparent in the well-documented and easily observable effects of visual, auditory and tactile figure/ground differentiation which can be reasonably safely extended into an understanding of how the fully encountered world is a landscape of shifting attention and attentiveness. The distinction between figure and ground has been developed into a full cognitive grammar (Langacker 2008), which has then been applied to literary texts (Hamilton 2003, Stockwell 2009, 2010). A passage of prose fiction can be analysed for the energy that seems to drive its narrative dynamic across clauses and indeed across the whole discourse (see also Talmy 2000a, 2000b), and this developing cognitive discourse grammar has also been applied in order to understand the workings of lyrical energy in poetry (Yuan 2013).

Prototypicality has also been an important and longstanding element in cognitive poetics (Lakoff 1987). The notion that categorisation is best understood as a cline from best-example to worst-example (rather than a binary opposition) requires a re-evaluation of most of the categories of philosophy and analysis. With specific reference to literature, issues of genre and mode are best understood as being products of prototypicality effects (Stockwell 2002), such that different types of literature, literary movements and periods can be understood in relation to each other. Even the way that figure/ground-like distinctions are applied in literary analysis has been understood in terms of a blending of semantic domains (see Fauconnier and Turner’s 2002 conceptual integration theory, Dancygier 2006), or a loosening of the binary distinctions in phonetics between voiced and unvoiced, consonant and vowel, and so on (Evans and Green 2006: 34, Stockwell 2009), or as a sense that the difference between real people, fictional characters, and caricatures is one of prototypicality rather than essential ontological difference (see below for a further discussion of literary character).

Perhaps the oldest part of cognitive poetics is its interest in worlds and schemas – deriving originally from Liebniz (1710) and Kant (1781) but receiving its modern form through the adaptation of possible worlds semantics and schema theory in computer programming in the 20th century. The notion that an encounter with a literary work encourages a vivid mental representation in a reader is in itself clear, and there have been numerous models within cognitive poetics aimed at understanding the detail of the process. The most successful of these has been text world theory (Werth 1999, Gavins 2007), which suggests that a principle of ‘text-drivenness’ determines which aspects of memory and knowledge are deployed during
any particular reading of a literary text. Text world theory has been used to provide an explanation of literary meaning and information, but has also been used more recently to offer a coherent account of feeling: empathy and identification on the one hand, and ethical relationships between reader, author, narrator and characters on the other. These two domains of aesthetics and ethics are briefly addressed below.

**From cognitive poetics to aesthetics and ethics**

Much cognitive poetics over the past two decades has been concerned with a systematic account of meaning, and though this part of the project is by no means complete, the field has had a great deal of success. However, many researchers have also been interested in the emotional, artistic and felt effects of literary reading, and the principles of cognitive poetics have also offered a method of addressing these issues more precisely and with greater empirical psychological and linguistic grounding than has been the case so far. After all, natural readers take up a novel, short story or poem often not simply for its informational content but for its engaging aesthetic effect; natural readers talk about literature in terms of feelings, force and impact, rather than as puzzles of meaning to be solved (even where such a resolution is a major part of the narrative dynamic, as in crime fiction or spy thrillers).

For example, in my own work I have explored the ways in which a cognitive poetic account can illuminate the difference between empathy (a close readerly identification and feeling with a fictional character) and sympathy (a more detached or distant but nevertheless affectionate feeling for a fictional character). In particular, text world theory has proven useful in calibrating the emotional distance between reader and character across world-boundaries. For example, in Stockwell (2011) I examined two 16th century lamentations by the Polish poet Jan Kochanowski: these were useful because they were both written at around the same time, by the same writer, in the same literary form, and about the same topic – the death of the poet’s young daughter. These circumstances offer an unusually convenient literary control, allowing me to investigate why most readers felt that one poem was much more moving, involving and empathetic than the other. The conclusions suggest that a text world analysis offers a useful means of capturing what are actually very subtle emotional effects, and a way of describing what is often very difficult to articulate in any sort of principled or systematic way.

In general, it seems to me that cognitive poetics has the potential now to capture many of the extremely delicate, rarefied and subtle feelings that literary works evoke. These aesthetic effects are very often right on the verge of subjective conscious awareness, or even possibly subliminal but experienced in a vague, ineffable sense that feels like a mood, disposition, tenuous and fleeting sensation or mild perception. Those literary critical terms of atmosphere, tone and air around a literary work – which have been discussed only in highly impressionistic and unsatisfactory ways for many years – seem at last to be opening up to rigorous exploration.

The inevitable integration of meaning and feeling has been a consequence of taking the cognitive turn seriously (Strawson 2010). Meaningfulness is experienced, and that experience has an emotionally engaged and involving aspect that is difficult to separate out. Similarly, meanings and feelings almost always involve an interaction with other people, their views, beliefs and expressions, and so there is also an inseparable ethical dimension.
The reconnection of meaning, feeling and moral relationships allows cognitive poetics to encompass an integrated account of informativity, aesthetics and ethics as articulated by Aristotle (c.350 BC) in his three rhetorical Appeals of logos, pathos and ethos (in the Rhetoric I.ii). Modern cognitive science suggests that these three aspects of discourse are in fact intimately bound up with each other, and so current cognitive approaches to literature have a particular interest in developing a rigorous account of readerly ethics. For example, both Phelan (2005, 2012) and Vermeule (2010) treat character more in terms that natural readers would recognise, rather than in the more theoretically complex but disembodied way that character has been treated (or ignored) within critical literary theory. Phelan, in particular, sets out an approach to character that is defined by three components: mimetic, thematic and synthetic. A character is mimetic in that he, she or it is a possible person. A character is thematic in that he, she or it has an artistic, compositional significance. And a character is synthetic since he or she or it is a product of artifice.

However, applying a prototypicality scaling to the notion of self and other, reality and fiction, subjective consciousness and the apparent presence of other minds, we might characterise our sense of ‘people’ as follows: the best example of a person is myself; other people are less good examples of people, but are still better examples of people than objects, animals or abstractions (Stockwell 2012). It is not true to say that trees, or horses, or cleanliness are not people; instead it is better to say that they are simply very poor examples of people. At least, this would be true in normal circumstances, but in particular literary settings, trees in Tolkien, horses in Swift, and Cleanliness in an English medieval allegory can be personified – where ‘personification’ means an elevation to the point of impersonation. So, in my own mind, I am a better example of a person than you are; people in my family are better examples of persons than people I have never met; and fictional characters are better or worse examples of persons depending on how richly fleshed-out they have been made by the author and me working together.

With a cognitively-informed account of character and person along these lines, it is clear that the field potentially has things as yet unresearched to say about a socio-ideological positioning that we might more briefly describe as ethics. The ethics of reading applies not only to relationships between readers and characters, but also to narrators, authors and intervening editors, of course, and so cognitive poetics is bringing us also to a re-evaluation of the nature of authorial intention, production and culture. In short, we are able – this time with system, principle and evidence – to address all of the concerns and interests of literary criticism. The ultimate project draws on an adapted scientific method and aspires towards a reasonable account of consciousness in literary reading: an artful science of literature.

References


