To what extent is the law of entropy applicable to literary language?

Literary language has been traditionally predominantly analysed from the perspectives of the arts and the humanities. However, since the second half of the 20th century, certain scholars have attempted to examine it using science instead. More specifically, the superconcept of entropy has been used to evaluate the value of the language used in literature. Though this claim is valid to a certain extent, in that writers will often tend to make their works more entropic to increase their ability to touch the public, which can be seen as a driver in the development of literary movements, the superconcept of entropy is not fully applicable to literature.

Superconcepts, as defined by Alan Wilson (2010, pp.40), are concepts applicable to many disciplines, though they often originate in only one. Entropy, one of the superconcepts he cites, originally stems from physics (Wilson, 2010, pp.43). Entropy in physics can broadly be defined in terms of statistics as a measure of the degree of disorder or randomness of energy particles (Razumovskaya, 2010, pp.263). According to the second law of thermodynamics (or law of entropy), in a closed system, entropy tends to increase (Tillman and Russell, 1965, pp.131) until the system has reached its most probable state, in which its disorder is maximised (Hockett, 1953, pp.92).

Entropy was first applied to language by Claude Shannon in what is known as information theory (Freund, 1980, pp.495). In language, unlike in physics, entropy measures the predictability of a text (Eco, 1989, pp.45). This is because, the more disorder in a passage’s organisation and structure, the trickier its content is to foresee (Holmes, 1985, pp.335). Hence, as entropy can broadly be seen as a measure of the degree of disorder, it can be applied metaphorically to language in that the more unpredictable a piece of writing, the more entropic (Velde, 1980, pp.302). In terms of information theory, the level of entropy of a given passage is equivalent to its level of information (Letzler, 2015, pp.28). Indeed, the level of information is calculated using a formally identical equation to Ludwig Boltzmann’s statistical equation for entropy in physics (Tillman and Russell, 1965, pp.127). Another important term to define in information theory is redundancy. Redundancy is essentially the opposite of entropy in that it is the parts of a texts which are determined by the structure and rules of language (Freund, 1980, pp.496). Therefore, redundancy makes a passage more predictable, and hence more organised (Eco, 1989, pp.51).
As the entropies of literary texts increase, so do their unpredictability. As a result, higher entropy (or information) has often been seen as a measure of unconventionality, and is thus seen as making works more valuable (Letzler, 2015, pp. 28). Norbert Wiener, for example, expressed this sentiment by pointing out that “clichés... are less illuminating than great poems” (Letzler, 2015, pp. 29). Similarly, Umberto Eco argued that “unorthodox uses of language” tend to result in poetry, whilst “conventional, probable” writing doesn’t (Letzler, 2015, pp. 29). For instance, when the Shannon entropies were measured for Shakespeare’s plays and Jin Yong’s works, it was found that, arguably, each of their most popular works, Macbeth and The Legend of the Condor Heroes, have the highest entropies (Chang, et al., 2016). This substantiates the argument that the more entropic, the greater the literary work. As a result, as innovation and originality (and thus entropy) are desirable in literary language, over time the entropy of literature in general would logically increase. Indeed, I. Kontoyiannis (1997) found that the King James Bible’s entropy was lower than those of Jane Austen’s novels’, which were in turn inferior to those of James Joyce’s novels. This shows a correlation between the level of entropy and time. Consequently, the law of entropy can be seen as being applicable to literary language.

Following the same line of argumentation, literary language in general may be viewed as more entropic than ordinary language (Thompson, 1977, pp. 228). This is supported by a study which measured the entropies of Greek and English literary works and Web-based news. It concluded that the language used in both English and Greek literature had a higher entropy than the more ordinary language used in the news (Papadimitriou, et al., 2010). From this, the creation and development of literature itself as cultures have advanced over long periods of time can also be regarded as exemplification of increases in entropy over time (and thus the law of entropy).

However, it is illogical for a passage with higher levels of entropy to necessarily be more valuable. If it were, the phrase “Constantinople fishing nasty pink” would have great worth (Eco, 1989, pp. 58-59), as it does not follow the rules of English syntax, and therefore has very low internal structure (redundancy) and high entropy. Yet, it is nonsensical. Moreover, this leads to the absurd conclusion that a string of random letters, which has the highest entropy, as it is completely unpredictable, is more valuable than any intelligible piece of writing (France and Hénaut, 1994, pp. 220). This is because information theory was developed for engineering purposes, in which information merely denotes the probability of the occurrence of certain messages (Thompson, 1977, pp. 228). Thus the
level of information (or entropy) in a message is not related in any way to its semantic content (Jones, 1973, pp.32). For instance, though, as seen earlier, novels written by James Joyce have greater entropy than the King James Bible, many find the latter to be more meaningful and valuable (Kontoyiannis, 1997). Therefore, a text cannot be praised solely because it has high levels of entropy (Letzler, 2015, pp.51). For texts to communicate any substantial meaning, they must possess a balance of order (redundancy) and chaos (entropy). This conclusion is supported by a study which analysed the entropies of *Moby Dick* and the Grimm Brothers’ Tales. It found that both bordered between periodicity and chaos (Ebeling and Pöschel, 1993). Texts with extremely high entropy, which are completely random, or extremely low entropy, such as the repetition of the letter “M” for 400 pages, usually have completely trivial meaning (Letzler, 2015, pp.51). Therefore, the law of entropy may not be so applicable to literary language, as it is doubtful that its entropy will continuously increase. This is the case as if a text’s entropy does not affect its value, authors may see little reason to make their writing more entropic. Furthermore, it would mean that eventually literature will be composed of nothing but totally random sequences of letters.

Nevertheless, though information, and therefore entropy, are not equivalent to meaning, and as a result, simply maximising entropy does not create more valuable literature, writers can still be viewed as attempting to increase entropy, not as an end, but rather as a means to an end. According to Eco (1989, pp.53), literary language is characterised by its capacity to create unusual meaning and emotions by expressing certain messages in more a entropic manner. In other words, an idea can be conveyed in many different ways, and the tendency in literature is to do so more entropically than if one were following precisely all of the rules of redundancy of a language. To illustrate this, he attempts to fully communicate all of the meaning in the first few lines of Petrarch’s poem “Chiare, fresche e dolci acque” (“Clear, fresh and sweet waters”) using all the rules of redundancy used in everyday language (thus writing much less entropically). (Eco, 1989, pp.53-54) In order to achieve this, it takes Eco almost a full page of text which is boring and forgettable, despite the fact that both Eco’s passage and Petrarch’s lines have the exact same meaning. What makes Petrarch’s poem so much more appealing is its originality of organisation and the deliberate improbability of its language compared to mundane language, and therefore its higher entropy, which informs us more powerfully about his love and the nature of love in general (Eco, 1989, pp.53-54).
The need for greater entropy in making a message more powerful can be explained by literature’s need for ambiguity in order to affect the reader (Razumovskaya, 2010, pp.262). Ambiguity is increased as writers express their ideas in new, relatively improbable and entropic ways. This is because, as pointed out earlier, entropy is fundamentally the opposite of redundancy, or the rules a language imposes to make messages more understandable (Freund, 1980, pp.496). Hence, an increase in entropy implies a decrease in redundancy, which results in the reduction of clarity of a text, and so makes it more ambiguous. According to R. Land, art without ambiguity is boring (Arnheim, 1974, pp.140). This is because ambiguity makes a piece more open to interpretation (Razumovskaya, 2010, pp.262), which gives it a greater chance of being able to intersect with some of each reader’s own personal emotional and cultural framework (France and Hénaut, 1994, pp.220), and therefore participate emotionally (Arnheim, 1974, pp.141). Hence, it tends to cause more appreciation, admiration, and identification with the artist on the part of the reader (France and Hénaut, 1994, pp.220). Literature’s impact on people is crucial to consider because it is the reader who determines how much value is placed upon different works (Land, 1973, pp.333). As Hénaut and France put it (1994, pp.221): “Without entropy there would be no possibility of exchange, without entropy there would be no art”. Therefore, authors are arguably inclined to write more entropically.

There are other factors which may make literature great apart from its form and ambiguity, such as the ideas conveyed in it. However, as Eco points out (1989, pp.55), unless there is novelty in what is being said, the novelty must lie in how it is said (the text’s structure, entropy, and ambiguity) in order for a piece of writing to be considered particularly valuable. As a result, as writers strive to express their ideas in new ways, they are likely to defy the given ‘order’ of literary language at the time. Effectively, they will introduce new entropic styles of writing, which, over time, may in turn become the norm, and thus the new ‘order’ that will eventually also be challenged. Subsequently, as time progresses, different literary styles and movements arise (Eco, 1989, pp.60-61). This process, driven by the need to write in an original, and thus relatively more entropic way compared to the standard of the time, demonstrates that entropy in literature is constantly being increased, just as the law of entropy predicts. Yet, simultaneously, the entropies of new styles decrease as they become more popular, and thus more common and probable. Consequently, although the relative entropy of literary language is always changing (being pushed up by innovation, and then back down due to increases in prevalence), the overall absolute entropy of literature can be seen as being relatively stable. Though it may have very gradually been increasing, as Eco suggests (1989, pp.61),
pointing out that contemporary poems are more ambiguous than classical ones, it is highly unlikely that it will continue to do so indefinitely, as literary language would eventually become nothing more than a string of random letters.

In conclusion, the law of entropy can be seen as applicable to literary language only to a certain extent because when observing literature within a relatively small timeframe, texts’ entropies and ambiguities tend to increase so that they may better impact readers. However, globally, the entropy of literary language seems to remain relatively stable because as certain unorthodox styles become popularised, they become more common and thus more predictable and less entropic. Furthermore, though it has been stated that the overall entropy of literature has been rising very slowly, it is unlikely that it will continue to do so after a certain point. This is because, the most entropic levels of language occurs when letters are completely randomly distributed. Were the law of entropy totally applicable to literary language, literature would eventually become incomprehensible and lose all value, at which point, one may wonder why have it at all.


