

Unit 5 Case studies

5.2 A Diachronic Study of *Melancholy* in a British Novel Corpus

Sally Kung, University of Birmingham

Introduction

In this case study, I compare the salient semantic features found in the contexts of the word *melancholy* in a diachronic corpus of fiction from 1722 to 1834. By building such a corpus, I aim to document and explain a change in the usage and meaning of *melancholy*. I intend to use corpus linguistics to provide empirical evidence of a semantic change of the emotion word *melancholy* from the pre-Romantic period to the Romantic period. I will use statistics such as frequency and the mutual information score to yield potential collocates and illustrate how *melancholy* has increased its popularity in usage among the novelists in the Romantic period. This study intends to show how a diachronic corpus study is a feasible way of explaining the semantic change of *melancholy*.

Data collection

As descriptive empiricists who value complexity and variation in language, corpus linguists observe patterns so as to put forth tentative hypotheses based on the principle that linguistic theory and description are best built on real and natural data. As Firth (1957: 199) says, '... it is unnecessary to assume any "facts" prior to statement [and] attested language text duly recorded is in the focus of attention for the linguist'. Nevertheless, this does not imply that corpus linguists need not use introspection; indeed, corpus linguists need insights and intuition to interpret natural data. The description of the naturalness of language resides on centrality and typicality. Empirical data support the establishment of what is central and typical in language. As Stubbs (1996: 28) concludes, corpus linguists maintain that 'Language should be studied in actual, attested, authentic instances of use, not as intuitive, invented, isolated sentences.'

Hunston (2002: 16) defines a diachronic corpus as 'A corpus of texts from different periods of time. It is used to trace the development of aspects of a language over time.' A diachronic novel corpus was built specifically for the purpose of this study to uncover any changes of the semantics relevant to *melancholy* in English novels over time. An obvious reason for building such a corpus is that no existing diachronic corpus exactly matches the objective of looking at the eighteenth century to early nineteenth century English discourse. Large amounts of electronic texts, mostly fiction, from this period are available in the electronic forms from large text repositories such as the Oxford Text Archive and Project Gutenberg. An instant advantage of using these novels as resources is that they do not involve the issue of copyright permissions.

In general, literary historians label the late eighteenth century and early nineteenth century the Romantic period. For example, Literature Online defines the literary period 1780-1837 as the Romantic period. Therefore, my corpus is divided into two sub-corpora, namely, the pre-Romantic period (1722-1779) and the Romantic period (1780-1834). The pre-Romantic sub-corpus consists of 11 novels (about 1.8 million tokens) and the Romantic sub-corpus consists of 21 novels (roughly 2.2 million tokens). The detailed composition of the corpus is shown in Appendix I. Each sub-corpus covers works authored and first published between the prescribed years. A date of first publication was accepted if the date of authorship was unclear and the work was not published posthumously. Only prose rather than poetry was included because poetry might contain different word order patterns. Short quotations from poetry were, however, common at the start of chapters at the time and these would not be removed as they were considered an integral part of the prose. Since it is a novel corpus, a large amount of direct speech, which sometimes reflects regional pronunciations, is present in the texts. It would be very difficult and disadvantageous to standardise this. Only British novels are included because my aim is to look at British society and discourse.

Data management

All the electronic texts were cleaned using Perl scripts to remove mark-ups. Next, WordSmith Tools was used as a concordancer to obtain concordances, retrieve collocates and generate word lists. The software can search all the occurrences of a specified word-form and display the node word in the centre of the screen within a pre-specified span for observation of patterns. An immediate question is how many words away from a node one must go before the collocate ceases to be affected by the node. In this paper, I follow the suggestion by Sinclair et al. (2004: xix) and use five words to the left of the node and four words to the right. In their experiment, they took one particular type as node and counted the collocates at the span position up to tenth order. They repeated the same procedure for all the types in the text. They estimated the expected value and the standard deviation of the number of collocates for any span position in a random text containing the same types. They used ten texts and for each text, they subtracted the expected value from the actual observed value at each position and divided by the standard deviation. They then found the mean value at each span position of all the ten texts and plotted the average node prediction graph.

The main statistical method I use in this study is the mutual information (MI) score. Church and Hanks (1990) propose the association ratio, which is similar to the mutual information $I(w_1, w_2)$. $I(w_1, w_2)$ compares the probability of observing w_1 and w_2 together (the joint probability) with the probabilities of observing w_1 and w_2 independently. It is defined as follows:

$$I(w_1, w_2) = \log_2 P(w_1, w_2) - \log_2 P(w_1)P(w_2)$$

In their application, word probabilities, $P(w_1)$ and $P(w_2)$ are estimated by $f(w_1)$ and $f(w_2)$, namely the number of observations of w_1 and w_2 in a corpus normalised by the size of the corpus, N . (They ignore pairs with small counts of less than five co-occurrences.) Joint probabilities, $P(w_1, w_2)$ are estimated by counting the number of times that w_2 follows w_1 in a window of w words, $f(w_1, w_2)$, and normalizing by N . The association ratio is not symmetric because $f(w_1, w_2)$ may not be the same as $f(w_2, w_1)$. This asymmetry of $f(w_1, w_2)$ gives order information and may shed further light on syntactic patterns. In addition, $f(w_1, w_2)$ may be smaller than $f(w_1)$ if w_1 and w_2 happen to appear several times in the window.

To compute the MI score using WordSmith, I first generated the two wordlists for the corpora of the first and the second period. I set the span to be five words to the left and four words to the right and the minimum number of occurrences of the collocates to be 2; WordSmith then calculated the MI score. Any interesting collocates with an MI score of 3.0 or above are discussed in the next section. Any MI score of below 3.0 is not regarded as significant.

Results and Analysis

Melancholy occurs 200 times among the 1.8 million pre-Romantic sub-corpus and 416 times among the 2.2 million Romantic sub-corpus. This means that, per 1 million tokens, *melancholy* roughly has 111 occurrences in the pre-Romantic period and 189 occurrences in the Romantic period. There are 167 significant collocates of *melancholy* in the earlier period and 321 significant collocates of *melancholy* in the later period. An immediate explanation is that more frequent usage of the emotion word attracts more collocates.

The first question that interested me was to find out what caused the emotion indicated by the word *melancholy*. From the 167 collocates, I found *grow*, *grew*, *became*, *cause*, *reason*. These seem relevant to my first query, although the instances with *cause* and *reason* found are not very relevant to the real causes of *melancholy*. After a closer examination of the concordances in context, I found that *melancholy* is typically caused by fear and financial loss in the pre-Romantic period. When the heroine in *Moll Flanders* (Daniel Defoe 1722) 'knew to be with a child by another man', she 'grew very melancholy' because she 'knew not what course to take.' Her melancholy was reflected in her physical illness - 'an ague' and was highly related to her 'apprehensions' that she should miscarry. We see that the subject was engaged

in thinking that something bad might happen and she feared as well as abhorred her thoughts.

Moll Flanders's husband *grew melancholy* because he lost a large sum of money. In his *melancholy*, he was 'disconsolate, and from thence lethargic, and died.' Her husband's *melancholy* gradually resulted in death. Another example of financial loss causing *melancholy* is that of Peregrine in *Peregrine Pickle* (Tobias Smollett 1751). He lost seven hundred pounds and came to Paris, and his friends engaged him 'in parties' so that he might not grow *melancholy* in reflecting upon his loss. A third example of becoming *melancholy* as a result of poverty is when Peregrine was unable to pay off his expense by 'his own industry' and suffered from ill health: 'his eye-sight failed, his appetite forsook him, his spirits decayed.'

In the Romantic sub-corpus, *nourish* is the most significant collocate of *melancholy*, with an MI score of 10. Looking at the two instances of *nourish* with *melancholy*, we can see that *melancholy* can be due to boredom and lack of acquaintances. Lady Jacintha in *Lovers and Friends* (Ann Hatton 1821) was 'dull and gloomy enough to nourish melancholy, and bring on madness' because she has 'no other society'. *Melancholy* is nourished by solitude of the environment and can cause delusions. In *The Monk* (Matthew Lewis 1796), Don Lorenzo's 'disposition to melancholy' was nourished by 'the calm of the hour and solitude' of the church. 'He thought of his union with Antonia,' his sister, and fell into *melancholy*, 'which accorded but too well with the religious gloom surrounding him.' We see that later Don Lorenzo had dreamt of his sister in the arms of a monster. Thus the causes of *melancholy* in the Romantic period are more likely due to the loss of people than to loss of money. Death would be the reason why the characters are *melancholy* in the Romantic period. The event itself is described as *melancholy*. There are 24 instances belonging to this category. In the pre-Romantic sub-corpus, I could find hardly any evidence of how to cope with *melancholy*. In contrast, *dissipate*, *dispel*, *divert* became significant collocates of *melancholy* in the Romantic period. For instance, after the sudden loss of her aunt Bianchi, her only relative, Ellena in *The Italian* (Ann Radcliffe 1797), endeavoured to 'dissipate melancholy recollections by employment'. These recollections were 'tender and mournful'.

The idea that *melancholy* can be dissipated by employment is not new. Indeed, Robert Burton's *Anatomy of Melancholy* ([1621] 1989), a treatise on the causes, symptoms and cure of melancholy, says 'for a melancholy young man urge it as a chief cause, why was he melancholy? Because idle.' The treatise itself was intended as treatment, as can be seen from the preface: 'I write of melancholy by being busy to avoid melancholy. There is no greater cause of melancholy than idleness, as better cure than business.'

Like Ellena in *The Italian* who lost her only relative, Antonia in *The Monk* also lost her beloved relative, her mother. She tried to read 'a volume of old Spanish ballads' to dispel her 'melancholy awe' when she recollected 'a thousand painful ideas' 'after her mother's death' although she did not quite succeed in dispelling her *melancholy*. Laurette in *The Orphan of the Rhine* (Eleanor Sleath 1798) was worried about her valuable friend, Madame Chamont, who had not returned home. She attempted to divert her *melancholy* by 'constant and unremitting employment'. 'The reward of effort' finally assumed 'the power of pleasing.' We see that people could work against the assault of *melancholy* by 'employment' and making effort.

Another measure to divert *melancholy* is to see nature. In *Lovers and Friends*, the Earl of Torrington had 'the late-obtained knowledge of Cecilia being his daughter' and his son Oscar could not remain in love with his own sister and said 'that she is my sister, is my grief – my misery; were she not your daughter, I would still seek and woo her for a bride', both of them become *melancholy*. The only method to divert their melancholy was to go to Lisbon.

In the novels of the Romantic period, it is described explicitly how the characters dealt with *melancholy*, either through labour or through travelling to have a change of environment. In the pre-Romantic period, *melancholy* resulted from financial loss, a deterioration of health, or pregnancy. In the Romantic period, it was a recollection of lost loved ones.

An important change in usage of *melancholy* from the pre-Romantic to the Romantic period is that there seems to be an increased correlation of *melancholy* with emotion words. In the pre-Romantic period, the collocates related to emotions are *gloomy*, *anguish*, *despair*, *uneasy*, *satisfaction*, whereas in the Romantic period, the collocates related to emotions have increased: *mood*, *deepest*, *pleasing*, *soothing*, *subdued*, *awe*, *gloomy*, *gloom*, *delightful*, *satisfaction*, *pleasure*, *solitude*, *deep*, *calm*, *anxious*, *happiness*. To support my claim that *melancholy* became increasingly related to emotion words in the Romantic period, I look at these emotion words in context.

The most significant collocate of *melancholy* in the pre-Romantic period is *pensive*. In *Moll Flanders* (1722), after her husband was told by Moll Flanders that 'I was not your lawful wife, and that our children were not legal children, so I must let you know now in calmness and in kindness, but with affliction enough, that I am your own sister, and you my own brother, and that we are both the children of our mother now alive,' he 'became pensive and melancholy' and 'the weight of it lay too heavy upon his thoughts', so he 'strangled himself'.

When one became *melancholy* in the eighteenth century, the period of Enlightenment in which rational thinking and reasoning was promoted, one became *pensive* as well. Mr. Glanville in *The Female Quixote* (Charlotte Lennox 1752) found Arabella, 'the pensive fair one, in a melancholy posture, her head reclined upon one of her fair hands; and though her eyes were fixed upon a book she held in the other, yet she did not seem to read, but rather to be wholly buried in contemplation.' From this instance, we see again that when one is *melancholy*, one contemplates in the pre-Romantic period. I did not find *pensive* among the collocates of *melancholy* in the Romantic period although there are two instances of *thoughtful*, a close synonym of *pensive*. In *The Last Days of Pompeii* (Edward Bulwer-Lytton 1834), a Greek character was described as having a particular gaze – 'His eyes, large and black as the deepest night, shone with no varying and uncertain lustre. A deep, thoughtful, and half-melancholy calm seemed unalterably fixed in their majestic and commanding gaze.' *Thoughtful* and *half-melancholy* are used to describe the calmness of the man's gaze. The usage of *half-melancholy* instead of *melancholy* hints a weaker link with thinking. Instead, *melancholy* more often associates with feelings in the late eighteenth and early nineteenth centuries. The French governess, Mademoiselle de Barras in *Evil Guest* (Joseph Sheridan Le Fanu 1830) 'had been called by every foul name which applied to the spy and the maligner; she could not bear it. Some one had evidently been endeavouring to procure her removal, and had but too effectually succeeded.' She was insulted and suspected by Mr. Marston and 'Mademoiselle was determined to go early the next morning; nothing should prevent or retard her departure; her resolution was taken. In this strain did mademoiselle run on, but in a subdued and melancholy tone, and weeping profusely.' She was 'a lost spirit, borne on the eddies of the storm; fearless and self-reliant, but with no star to guide her on her dark, malign, and forlorn way.' *Melancholy* in this instance describes the tone of the character with a *subdued* feeling which is reflected in *weeping profusely*. Crying is closely related to the emotion word *melancholy* only in the Romantic period. This can be supported by the evidence that *tears* is a significant collocate in the Romantic sub-corpus but not in the pre-Romantic sub-corpus. For instance, *melancholy* describes the countenance of a character, Matilda in *The Monk*, with a feeling of *sorrow*, who had *involuntary tears* because her inconstant lover Ambrosio had lost interest in her. 'While he waited for the opportunity of satisfying his unwarrantable lust, every day increased his coldness for Matilda', whose 'eyes filled with involuntary tears' and whose countenance had 'soft melancholy.' Although 'Ambrosio was not unmoved by her sorrow', Matilda tried 'in vain to regain his affection' and 'continued to treat her inconstant lover with her former fondness and affection'.

Melancholy collocates with *gloom* in both sub-corpora. I found in two instances, one in the pre-Romantic period and one in the Romantic period, that the subjects suffering *melancholy* had lost a friend. In *Love in Excess* (Eliza Haywood 1725), the Count D'Elmont was 'in a gloomy melancholy for the loss of his friend' Monsieur Frankville. In this case, we do not see any deep feelings attached to the *melancholy* due to the loss of a friend. In fact, the Count D'Elmont had a 'secret adoration' of his friend's daughter Melliora. The instance found in the

Romantic period is from *Frankenstein* (Mary Shelley 1818). Frankenstein was engaging in *melancholy* after the murder of his friend Clerval. Clerval's image was forever before him, 'ghastly and murdered'. As his health improved, he 'was absorbed by a gloomy and black melancholy that nothing could dissipate.' 'Soon, oh, very soon, will death extinguish these throbbings and relieve me from the mighty weight of anguish that bears me to the dust; and, in executing the award of justice, I shall also sink to rest.' The melancholic patient engages in thinking of dying. The idea that someone being *melancholy* will be suffering from thinking of death can be dated back to Galen ([ca. 165 C.E.] 1976), who says:

Although each melancholic patient acts quite differently than the others, all of them exhibit fear or despondency. They find fault with life and hate people; but not all want to die. For some the fear of death is of principal concern during melancholy. Others again will appear to you quite bizarre because they dread death and desire to die at the same time.

Melancholy in the Romantic period is described with greater specific details and more profound effects on the characters. *Melancholy* is being exhibited as a character being *unsociable* and being seen with compassion in *Things As They Are* (William Godwin 1794). Mr. Falklands was accused of murder and his personality changed as 'no two characters can be in certain respects more strongly contrasted, than the Mr. Falkland of a date prior and subsequent to these events'. His mood became a 'gloomy and unsociable melancholy' which was a complete opposite to 'the habits of his life' which was 'full of cheerfulness and tranquillity' and to his sanguine mind. 'From an object of envy he was changed into an object of compassion. Life, which hitherto no one had so exquisitely enjoyed, became a burden to him. No more self-complacency, no more rapture, no more self-approving and heart-transporting benevolence!'

The words *habitual* and *frequent* appear as more significant collocates of *melancholy* in the later period than in the earlier period. As Burton ([1621] 1989) states 'Melancholy the subject of our present Discourse, is either in Disposition, or Habite.' In *The Orphan of the Rhine*, St Gerard was in huge debt and asked his only sister Madame Laronne to 'request the loan of a sum sufficient to discharge the debt' but his sister refused his request. 'When the mingled emotions of indignation and anguish had, in some measure, subsided, he seemed to have lost all his energy of soul; nothing bestowed even a transitory pleasure, and he sunk into the most alarming melancholy!' 'His constitution, which in the early part of his life had suffered much from the severity of military discipline, now became visibly impaired; the disorder of his mind daily increased; melancholy became habitual to him, and so rapid was the progress it made in undermining his health.' He soon suffered 'a nervous fever' which 'in a short time terminated his existence.' Apparently, when *melancholy* becomes habitual, it can be fatal and cause death.

The last observation from the collocational profiles of *melancholy* is that words like *pleasure*, *pleasing*, *soothing*, *delightful*, *happiness* came up in the Romantic period, but not in the pre-Romantic period. The phrases *melancholy pleasure* and *pleasing melancholy* occur six and four times, respectively, in the later sub-corpus. Caleb Williams in *Things As They Are* explains, 'I have already said that one of the motives which induced me to the penning of this narrative was to console myself in my insupportable distress. I derive a melancholy pleasure from dwelling upon the circumstances which imperceptibly paved the way to my ruin.' It seems that pleasure and melancholy, the opposing feelings exists in the same space. In *Problems*, attributed to Aristotle ([4th century B.C.] 1957), melancholy is a type of character typical of brilliant and outstanding men.

Why is it that all men who have become outstanding in philosophy, statesmanship, poetry or the arts are melancholic, or some to such an extent that they are infected by the diseases arising from black bile, as the story of Heracles among the heroes tells?

The brilliance attributes of *melancholy* were revived in the literature of the Romantic period as a sensual pleasure and joy.

Conclusion

In this study, I have given evidence to support the contention that the meaning and usage of *melancholy* has changed by taking a diachronic corpus linguistic approach. I built a corpus consisting of two sub-corpora of English novels, one from 1722 to 1779 and one from 1780 to 1834, to represent novels from the pre-Romantic the Romantic periods. I combined the computational analysis with interpretations of the concordance lines to draw several conclusions. First, I found a higher relative frequency in the Romantic period sub-corpus, which subsequently led to more significant collocates. Second, the Romantic period provides more causes of and solutions to *melancholy*. Third, *melancholy* links to reasoning in the pre-Romantic period and links to emotions in the Romantic period. Fourth, it becomes a habit and can lead to death in the later period. More detailed descriptions of the behaviour of the characters are given in the novel literature of the Romantic period. Lastly, *melancholy* associates with pleasing mood in the Romantic period.

Appendix I

Table 1: The pre-Romantic sub-corpus: 1,791,582 tokens

year	author	title	tokens
1722	Defoe	<i>Moll Flanders</i>	136,495
1725	Haywood	<i>Love in Excess</i>	88,137
1741	Richardson	<i>Pamela, or Virtue Rewarded</i>	223,121
1749	Cleland	<i>Memoirs of a Woman of Pleasure</i>	85,786
1749	Fielding	<i>The History of Tom Jones</i>	347,650
1751	Smollett	<i>Peregrine Pickle</i>	346,959
1752	Lennox	<i>The Female Quixote</i>	147,126
1759	Johnson	<i>The History of Rasselas</i>	37,357
1760	Sterne	<i>Tristram Shandy</i>	187,456
1771	Mackenzie	<i>The Man of Feeling</i>	36,639
1778	Burney	<i>Evelina</i>	154,856

Table 2: The Romantic sub-corpus: 2,203,696 tokens

year	author	title	tokens
1793	Smith Godwin	<i>The Old Manor House</i>	219,363
1794	Lewis Radcliffe	<i>Things As They Are</i>	133,027
1796	Sleath	<i>The Monk</i>	136,869
1797	Edgeworth	<i>The Italian</i>	172,789
1798	Lewis	<i>The Orphan of the Rhine</i>	107,255
1800	Porter	<i>Castle Rackrent</i>	39,051
1808	Austen	<i>The Anaconda</i>	18,858
1810	Austen Peacock	<i>The Scottish Chiefs</i>	286,554
1811	Scott	<i>Sense and Sensibility</i>	119,832
1813	Peacock	<i>Pride and Prejudice</i>	123,446
1815	Shelley	<i>Headlong Hall</i>	27,664
1815	Anonymous	<i>Guy Mannering</i>	163,583
1818	Hatton	<i>Nightmare Abbey</i>	26,780
1818	Ainsworth Hogg	<i>Frankenstein</i>	75,273
1820		<i>The Bride of the Isles</i>	9,817
1821	Buckstone	<i>Lovers and Friends</i>	242,608
1822	Le Fanu	<i>The Spectre Bride</i>	3,801
1824	Bulwer-Lytton	<i>The Private Memoirs and Confessions of a Justified Sinner</i>	84,382
1826		<i>Luke the Labourer</i>	14,090
1830		<i>Evil Guest</i>	46,471
1834		<i>The Last Days of Pompeii</i>	152,183

References

- Aristotle ([4th century B.C.] 1957) book XXX of Aristotle: *Problems*, vol. 16, translated by W.S. Hett Cambridge, Mass.: Harvard University Press
- Burton, R. ([1621] 1989) *The Anatomy of Melancholy* Oxford: Oxford University Press
- Church, K. & Hanks, P. (1990) 'Word Association Norms, Mutual Information and Lexicography' in *Computational Linguistics* 16(1), 23-29
- Firth, J. (1957) 'Modes of Meaning' in Palmer, F.R. (ed.) *Papers in Linguistics 1934-1951* London: Oxford University Press, 190-215
- Galen ([ca. 165 C.E.] 1976) 'Function of Diseases of Brain and Spinal Cord' in Diseases of the Black Bile, *On the Affected Parts*, translated by Rudolph E. Siegel. Basel: S. Karger
- Hunston, S. (2002) *Corpora in Applied Linguistics* Cambridge: Cambridge University Press
- Literature Online <http://lion.chadwyck.co.uk/refineSelectList.do>
- Sinclair, J., Jones, S. & Daley, R. (2004) *English Collocation Studies: The OSTI Report* Continuum: London
- Stubbs, M. (1996) *Text and Corpus Analysis* Oxford: Blackwell