



Universiteit
Leiden
The Netherlands

Affordances and limitations of algorithmic criticism

Verhaar, P.A.F.

Citation

Verhaar, P. A. F. (2016, September 27). *Affordances and limitations of algorithmic criticism*. Retrieved from <https://hdl.handle.net/1887/43241>

Version: Not Applicable (or Unknown)

License: [Licence agreement concerning inclusion of doctoral thesis in the Institutional Repository of the University of Leiden](#)

Downloaded from: <https://hdl.handle.net/1887/43241>

Note: To cite this publication please use the final published version (if applicable).

Cover Page



Universiteit Leiden



The handle <http://hdl.handle.net/1887/43241> holds various files of this Leiden University dissertation

Author: Verhaar, Peter

Title: Affordances and limitations of algorithmic criticism

Issue Date: 2016-09-27

Chapter 8

Data analysis

8.1. Introduction

Many of Louis MacNeice's poems express a fascination with the notion that changes in perspective can produce entirely new experiences. In "Under the Mountain", for instance, the poet captures the idea that the view from a mountain can miraculously purge a landscape of its everyday banalities. At ground level, "[t]he breakers are cold scum and the wrack / Sizzles with stinking life", but the mountain view reveals the poetry of the landscape and changes the bay into "a goose-quill that feathers ... unfeathers itself" (ll. 2-3). Similarly, in the early poem "Morning Sun", MacNeice concentrates on the transformative powers of sunlight. While daylight "blazons / The red butcher's and strolls of fish on marble slabs" (ll. 14-15), this "hanging meat / And tiers of fish are colourless and merely dead" at dusk (ll. 19-20). The poem "Mutations" proposes additionally that such unexpected vistas and startling perspectives are essential for the intellectual development of human beings. Shifts in perspective have the effect that established views may "crack at times" and that "new / Patterns from new disorders open like a rose" (ll. 8-9).

The urge to apply computing in the context of literary research is often driven by the analogous conviction that changes in perspective can foster new kinds of insights. In literary informatics, the changed vista generally involves a transition from a detailed examination of a limited number of texts to the possibility to focus on collections in their entirety. Texts in which concepts and literary techniques are originally presented linearly are transformed into a collection of discrete data, allowing for structural or synchronic analyses of full corpora.⁵¹⁴ When data are generated out of primary sources, and when the resultant annotations are stored separately from the original documents, it may be argued that these structured annotations collectively form a new and more encompassing resource. Within the data set that is created, the original division into books and volumes becomes less

⁵¹⁴ Writing in 1978, Susan Wittig conjectured that progress in the field of computer-based textual analysis was hampered by the fact that texts were habitually viewed as linear structures. This view, which, according to Wittig, was inherited from New Criticism, undervalued the non-linear or network-like patterns that can be produced by the phonological, syntactic or semantic units within literary texts. See Susan Wittig, "The Computer and the Concept of Text", in: *Computers and the Humanities*, 11 (1978), p. 212. Although the view on whether or not the original text should principally be viewed as a linear entity may differ, the results of algorithms for the creation of structured annotations about texts are typically discrete and context-insensitive.

consequential. In a sense, such large databases implement the logic that was envisaged for Vannevar Bush's memex. Bush proposed that knowledge which is traditionally organised according to physical boundaries of the printed work should be reorganised on the basis of the principle of "associative indexing, the basic idea of which is a provision whereby any item may be caused at will to select immediately and automatically another". In the memex, individual pages could also be "gathered together to form a new book".⁵¹⁵ The discrete data sets that are generated out of linear texts can likewise serve as proxies in which fragments from one particular text can be connected directly to fragments from other texts, purely on the basis of semantic or linguistic similarities. Gooding argues, more strongly, that digitisation projects decontextualize and deconstruct texts, to such an extent that there no longer are any books. The original texts are replaced by a vast corpus of words and metadata, which can all be divorced from the original context and which can collectively be searched uniformly.⁵¹⁶

The benefits of digital tools often seem proportional to the volume of the text collection that is analysed. Tasks which require few data as input are generally performed most effectively by human beings, especially when such work demands interpretation. The analysis of corpora that span hundreds or thousands of works is strongly complicated, however, by the fallibility of human memory and the inconsistency of human concentration levels. Methodical comparisons of thousands of texts can generally be performed exclusively when scholars make use of digital research methods. In his monograph *Macroanalysis*, Matthew Jockers clarifies the value of computational analyses by drawing an analogy with the field of economics. Whereas microeconomics concentrates on individual companies or on individual families, macroeconomics aims to explain the economic behaviour of countries or of continents in their entirety.⁵¹⁷ Literary informatics comparably aims to uncover the larger trends in text collections via rigorous manipulation of large-scale aggregations of digital data. Other scholars have proposed similar metaphors to clarify the rationale of the shift to a larger scale. John Burrows compares large text corpora to handwoven rugs in which the "principal point of interest is neither a single stitch, a single thread, nor even a single color but the overall effect".⁵¹⁸ In the first pamphlet of the literary lab, Allison et al. use architecture as an image, suggesting that literary texts "like buildings, possess distinctive features at every possible scale of analysis".⁵¹⁹

⁵¹⁵ Vannevar Bush, "As We May Think", in: *The Atlantic*, (1945), n.pag.

⁵¹⁶ P. Gooding, M. Terras & C. Warwick, "The Myth of the New: Mass Digitization, Distant Reading, and the Future of the Book", in: *Literary and Linguistic Computing*, 28:4 (13 August 2013), p. 428.

⁵¹⁷ Matthew Jockers, *Macroanalysis : Digital Methods and Literary History*.

⁵¹⁸ John Burrows, "Textual Analysis", p. 324.

⁵¹⁹ Sarah Allison et al., *Quantitative Formalism: An Experiment*, p. 8.

In this thesis' case study, data have been produced about a number of literary phenomena, and these structured data allow for a wide range of statistical analyses. This chapter discusses a number of ways in which the data can be processed.

8.2. Data analysis

As a rule, computer-based literary research is based on the central assumption that texts can be compared on a quantitative basis. While this reduction to numbers evidently implies a full neglect of the numerous unfathomable and unquantifiable ways in which poetry can produce effects, ensuring the statistical soundness of numerical comparisons poses challenges as well. The software that was developed for the purpose of this study has considered MacNeice's verse line by line. In agreement with Martin Mueller's observation that text analysis tools essentially produce lists,⁵²⁰ the direct output of the application consisted of a simple index of all individual occurrences of the literary devices under investigation. To shift away from a focus on separate occurrences, and to be able to extend the scope to larger units, such as poems or volumes in their entirety, it is necessary to aggregate these individual values.

In most cases, instances of literary phenomena can be counted in a variety of ways. In the case of alliteration, for instance, metrics can be produced by counting the different repeated sounds that occur within a line. Alternatively, metrics may also reflect the number of words that alliterate, or the number of verse lines within a poem that actually contain alliteration. Applying these three distinct methods to the line "This wind from the west be backed by waves" (l. 94), in "Country Weekend", for instance, would result in values two, six and one respectively. The number of rhyming sounds in a poem can similarly be counted in different ways. Stanza totals may either reflect the number of repeated sounds within a stanza or the total number of lines that rhyme. The second stanza of the poem "Passage Steamer", for instance, consists of 7 lines which end in "require", "desire", "sun", "none", "mast", "past" and "begun", making the rhyming scheme AABCCB. In this situation, there are obviously seven lines that rhyme, but only three different rhyming sounds. In this situation, the choice of the summation method depends on the objective of the analysis. Counts of the lines that rhyme give an impression of the overall regularity of the stanza. Counts of rhyming sounds, by contrast, indicate either the monotony or the versatility of the lines that rhyme.

Counts are rarely neutral, as they invariably demand decisions on what and how to count. Such discussions of the various methods for producing counts are not trivial, as decisions on how to count can directly have a strong impact on the outcomes of subsequent statistical processing. As a result of the relative novelty of this type of research, few best practices have emerged. Transparency on how

⁵²⁰ Martin Mueller, "Digital Shakespeare, or towards a Literary Informatics", p. 291.

numbers have been produced, therefore, is of crucial importance in literary informatics. In the case of alliteration, it was observed that there are important differences in the intensity of occurrences. Lines such as “Further who failed last Friday to feel grieved” in “Easter Returns” (l. 11), or “The flotsam of private property pekinese and polyanthus” in “An Eclogue for Christmas” (l. 100) clearly alliterate heavily. These examples differ in a qualitative manner from a line such as “Round the Corner is sooner or later the sea” (l. 2) in “Round the Corner”, as this line only contains one repetition of an /s/ sound. In this study, it was decided to give expression to the intensity of occurrences of alliteration by counting the number of words that alliterate. The same decision has been taken for occurrences of assonance and consonance. With respect to rhyming sounds, it was decided to count the number of lines that rhyme, rather than the number of rhyming words, as analyses of rhyme are mostly aimed at determining the regularity of stanzas.

An additional complexity is produced by the fact that the poems in the corpus obviously differ in length. When different poems need to be compared directly, it does not seem legitimate to compare texts on the basis of absolute counts. In a long poem such as “Valediction”, which contains 105 lines and 856 words in total, the author obviously has much more opportunities to use alliterative language than in “Aubade”, which only contains 53 words. It has been decided, for this reason, to correct all absolute numbers, by dividing the total number of occurrences either by the total number of tokens or by the total number of lines. Alliteration is obviously a device which is applied at the level of tokens. In theory, an author can introduce alliterations with each new word that is used, and the total number of alliterative sounds can never be higher than the total number of tokens. For this reason, alliteration, assonance and consonance have been normalised using the token counts. Devices such as perfect rhyme, slant rhyme and semi-rhyme, however, operate at the level of the verse line, and the absolute counts of all occurrences or rhymes have been normalised on the basis of the total number of lines. On the basis of these principles, normalised values have been produced for all the 311 poems in the corpus. The long poem “An Eclogue for Christmas” contains 1444 tokens, and the absolute counts for alliteration, for instance, amounts to 253. The normalised value is $(253 / 1444 =) 0.175$. The shorter poem “Invocation”, which has 146 words, has 52 alliterative words. The relatively high intensity of alliteration in this poem is expressed in the normalised value 0.356. During initial tests, an additional measure proved to be necessary to ensure the accuracy of the normalisation method for counts of rhymes. A number of poems contain stanzas which consist of one line only. One example is MacNeice’s “Autobiography”, in which the haunting refrain “Come Back Early or Never Come”, is not contained within a longer stanza. When stanzas contain one line only, it is evidently impossible for these to contain rhyme. The number of single-line stanzas have therefore been deducted from the total number of lines before the normalisation.

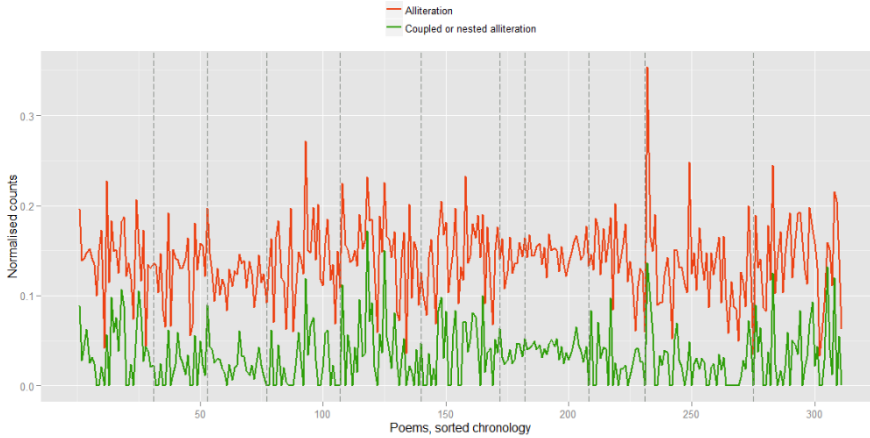


Figure 8.1. Normalised counts of occurrences of alliteration

The values that have been calculated can be used, firstly, to trace particular historical developments at a macro-level.⁵²¹ Figure 8.1 displays the distribution of the counts for alliteration across MacNeice's entire oeuvre. In the diagram, the poems have been sorted according to their location in the *Collected Poems*, edited by Peter MacDonald, and published in 2007. This arrangement roughly corresponds to a chronological order. The sections which are distinguished using the dashed vertical lines correspond to the 11 volumes in which these poems have appeared. As can be seen from the peaks and the troughs in the diagram, the values that have been calculated for alliteration fluctuate heavily during most of MacNeice's career. The usage is relatively stable, nevertheless, in *Autumn Journal*, *Ten Burnt Offerings* and *Autumn Sequel*.⁵²² As was also explained in Chapter 6, MacNeice frequently composed verse lines containing notable combinations of alliterative words. Many lines contain alliterations nested within another alliteration (e.g. "As gay trams run on tracks and cows give milk" in line 16 of "An April Manifesto" or "Greek words sprout out in tin on sallow walls" in line 324 of "Eclogue from Iceland") or structures in which two sounds in the first half of the line are repeated in the second half (e.g. "The ancient smiles of men cut out with scissors and kept" in line 17 of "Perseus" or "Columns of ads the quickest roads to riches" in line 25 of "Christmas Shopping"). The green line in figure 8.1 represents all occurrences of such specific forms of alliteration. This analysis indicates that MacNeice used such nested or alternating structures in most of his poems, and,

⁵²¹ The visualisations that are discussed in this chapter were created in R or in Processing. The code that was used to create these visualisations can all be found at <<https://github.com/peterverhaar/Phd>>.

⁵²² In figure 8.1, the poems in *Autumn Journal*, *Ten Burnt Offerings* and *Autumn Sequel* have been assigned the numbers 53 -77, 172 -182 and 183-208 respectively.

additionally, that the ratio between regular alliteration and such specific forms of alliteration is more or less stable.

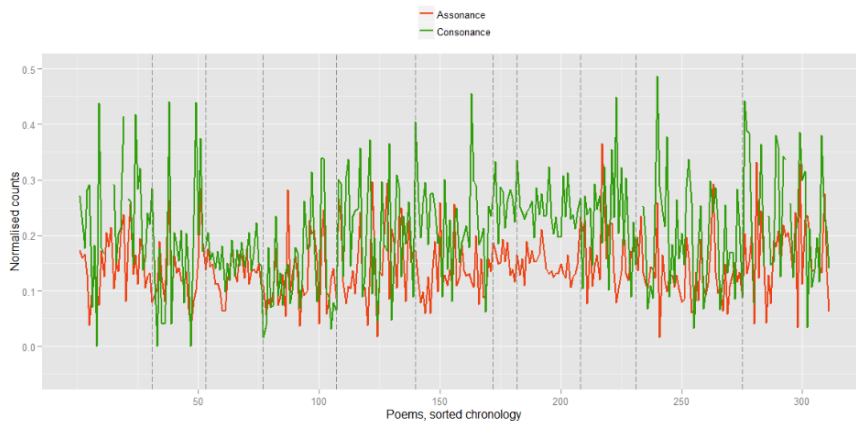


Figure 8.2. Normalised counts of occurrences of assonance and consonance

Figure 8.2 shows that MacNeice likewise made a relatively solid use of consonance and assonance. In general, the counts for consonance are higher than those for assonance. As is the case for the use of alliteration, the poems in *Autumn Journal*, *Ten Burnt Offerings* and *Autumn Sequel* display less variability with respect to the occurrences of assonance and consonance. Figure 8.2 also reveals, interestingly, that there are three poems which are fully without any occurrences of consonance. In this study, consonance has been implemented strictly as “the repetition of the sound of a final consonant or consonant cluster in stresses, unrhymed syllables near enough to be heard together”.⁵²³ Although some authors view alliteration as a specific form of consonance in which sounds are repeated at the beginning of stressed syllables, this study has distinguished alliteration and consonance more sharply. The three poems which were found to be without consonance are “The Sunlight on the Garden”, “Cuckoo” and “Sand in the Air”. Consonance, like assonance, can be used to suggest connections between words, and, fittingly, the poems which are devoid of consonance all describe a loneliness or an isolation. “Cuckoo” references Symeon the Stylite who was solitary on his pillar and who formed the “Personification / Of distance” (ll. 7-8). “Sand in the Air” and “The Sunlight on the Garden” both describe a feeling of intense desolation after being abandoned by another person. The persona in the former poem bemoans his apprehension that in a “shrivelled world / There is only I” (ll. 21-22). The fact that the three poems that were mentioned lack consonance can easily be

⁵²³ “Alliteration”, in *The Princeton Encyclopedia of Poetry and Poetics*, p. 299.

overlooked in studies based on traditional methods. Human scholars tend to concentrate on the textual phenomena which are present, and the basic fact that certain other phenomena are absent is often difficult to notice.

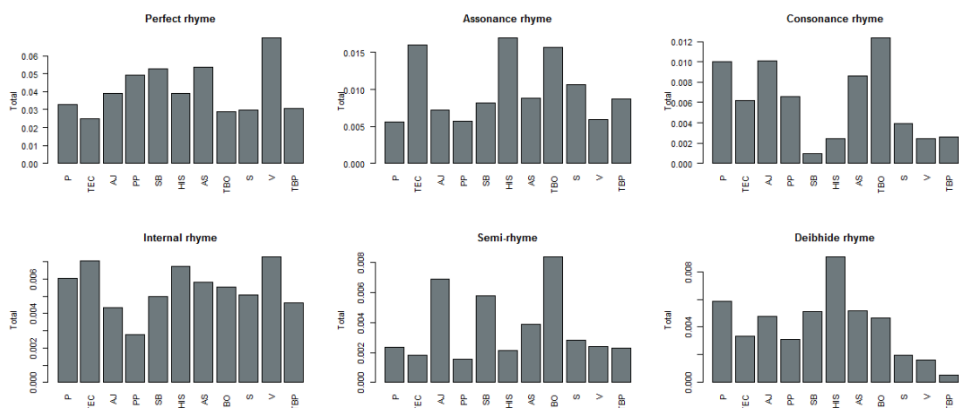


Figure 8.3. Normalised counts for perfect rhyme, assonance rhyme, consonance rhyme, internal rhyme, semi-rhyme and deibhide rhyme, aggregated at the level of volumes

To allow for a more focused comparison of the use of literary techniques throughout MacNeice's career, the normalised metrics can also be aggregated at the level of volumes.⁵²⁴ Figure 8.3 displays the main differences in the use of the various forms of rhyme. The diagram was created, firstly, by counting all the occurrences in the various volumes, and, secondly, by dividing these counts by the total number of tokens in each volume. As can be seen, perfect rhyme has been used most frequently in MacNeice's penultimate volume *Visitations*, and it has been used most sparsely in the second volume *The Earth Compels*. For some volumes, the frequencies of perfect rhyme are inversely proportional to the frequencies of slant rhyme. *Visitations* has low scores for both assonance rhyme and consonance rhyme, while *The Earth Compels* has one of the highest values for assonance rhyme. The diagram also reveals that the volumes *Springboard* and *Holes in the Sky* appear to make use of a very specific form of slant rhyme. In these volumes, there are many occurrences of assonance rhyme, and relatively few occurrences of consonance rhyme. Other interesting findings are that the use of semi-rhyme is most frequent in *Autumn Journal* and *Ten Burnt Offerings*, and that the use of internal rhyme roughly remains at the same level throughout the

⁵²⁴ *Autumn Journal* and *Autumn Sequel* are evidently not volumes. They are lengthy poems consisting of separate cantos. For ease of reference, however, this text will refer to these collections as volumes as well.

poet's career, with the exceptions of *Autumn Journal* and *Plant and Phantom*. In general, the frequencies of internal rhyme, semi-rhyme and deibhíde rhyme are much lower than those of perfect rhyme and slant rhyme. The use of deibhíde rhyme is relatively stable in MacNeice's first eight volumes, with a very sudden upsurge, however, in *Holes in the Sky*. The usage of this Celtic rhyming device diminishes quite abruptly in the last three volumes.

In this case study, data have also been collected about the imagery that occurs in MacNeice's poems. Figure 8.4 clarifies a number of historical developments in the use of imagery. The diagram was produced by dividing the absolute number of references in each volume by the total number of tokens in each volume. The sizes of the circles reflect the number of references to the various images.⁵²⁵ On the basis of the normalised counts of all references, it can be observed that the images "Religion", "Plants", "Transportation", "Light" and "Water" have been used most profusely within the corpus. The counts provide support for Brown's observation that references to the sea and to trains and other modes of transportation are frequent in MacNeice's verse. The volume entitled *Plant and Phantom* also contains the highest number of references to plants. Religious imagery is used consistently throughout the entire career, and most frequently in *Springboard*, *Holes in the Sky* and *Visitations*. In the early volumes *Poems* and *The Earth Compels*, there are mainly reference to Christian religion, and the passages in which they are used often stress a conflict between religious ideals and the observed practices of a secular society. In "An Eclogue for Christmas", people decorate their houses with "tinsel and frills" to "announce that Christ is born among the barbarous hills" (ll. 8-9), and people principally worship "the cult of every technical excellence" (l. 153). "Belfast" portrays a beggar woman "[t]o whose prayer for oblivion answers no Madonna" (l. 20), and "Birmingham" describes the failed "endeavour to find God" (l. 15). The volume *Plant and Phantom* refers more frequently to non-Christian religions and to mythology. "Stylite", for instance, describes a "white Greek god" (l. 21), "Plant and Phantom" mentions the "pawky Fates" (l. 9), and "Novellettes" refers to an "Aztec pyramid of sacrifice" (l. 2017).

Whereas MacNeice wrote in the *Poetry Book Society Bulletin* that his volume *Solstices* "contains practically no allusions to either Graeco-Roman or Christian legend",⁵²⁶ this study did identify many references to Christian religion in this volume. "Bad Dream", for instance, contains references to a bible and a "crucifix on the wall" (l. 22) and "Jericho" mentions "the fires of Pentecost" (l. 24). "Il Piccolo Rifiuto" describes the irritability that can result from the sense of being lost in translation in another country. The persona noticed that, during moments of

⁵²⁵ During various trials and errors, I discovered that more common visualisation techniques, such as line diagrams or bar charts could not effectively be used to represent the developments in the use of imagery. The visualisation in figure 8.4 was programmed using the Processing environment.

⁵²⁶ Louis MacNeice, Peter McDonald (ed.), *Collected Poems* (London: Faber & Faber 2007), p. 795.

miscommunication, “God / began to limp” (ll. 12-13). “The Blasphemies” and “The Messiah” probably contain the most obvious references to Christian religion. The former poem describes the persona’s troubled fidelity to religious faith. As a young boy, the persona is afraid to commit the “sin against the Holy” (l. 1), while in later life, he “turned to parody / Prayers, hymns, the Apostles’ Creed” (ll. 14-15). “Windowscape”, furthermore, is a poem which focuses on a lacklustre suburb, which reinforces its central sense of doom and fatalism by stressing the obsolescence of religious practices. Although the “[w]indow-cleaner and postman call just once a year”, there is “never a priest” (ll. 8-9).

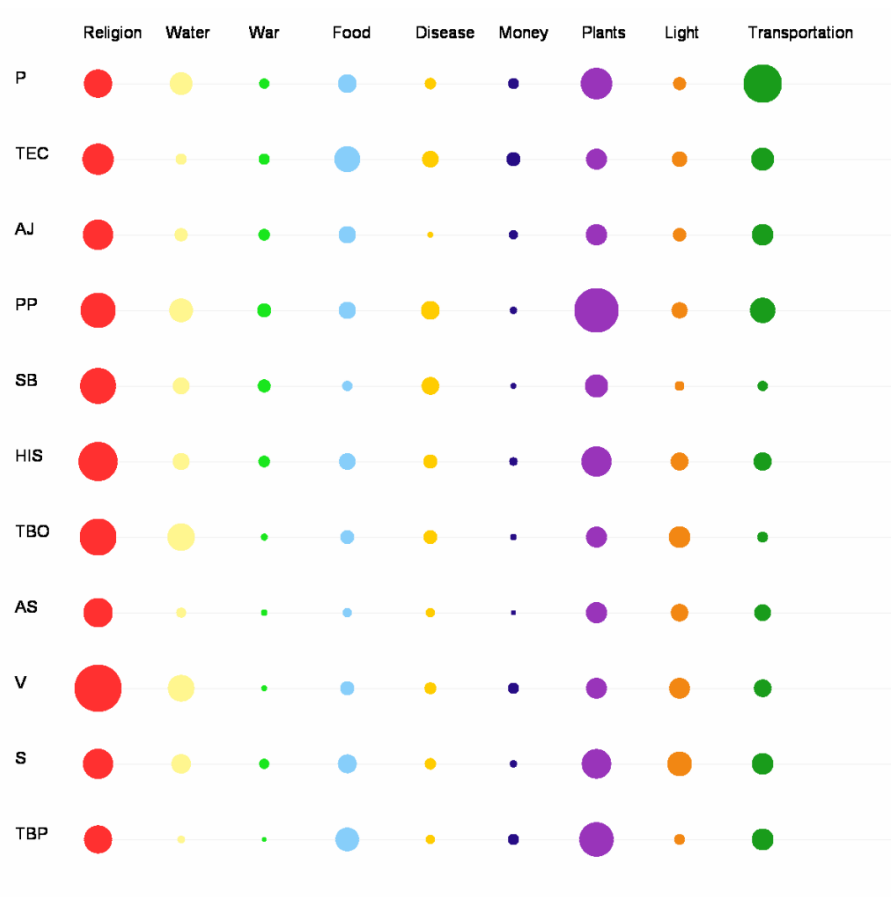


Figure 8.4. Use of imagery

One notable finding is also that war imagery remains relatively stable throughout the entire oeuvre. The war imagery is not significantly higher in the volumes

written during the Second World War. In *Poems* and *The Earth Compels*, war imagery is often employed to evoke a general sense of doom, or to underscore the menacing aspect of forces that represent change. In “Spring Voices”, the coming of spring is presented as a preparation for a war. The new season is “massing forces”, and the main character is afraid “of the barrage of sun that shouts cheerily” (l. 2). In “June Thunder”, the sudden storm is presented as “the sword of the mad archangel” (l. 23). Within the poem “Brother Fire”, which is arguably the poem in which MacNeice addresses the atrocities of war most poignantly, no explicit war images were found.

Throughout MacNeice’s corpus, copious references were also found to consumer articles and to objects from popular culture. Images such as “Food”, “Drinks”, “Media”, “Sports” and “Money” all have comparatively high frequencies. References to money and to wealth often occur, expectedly, in passages which express a criticism of a capitalist and a consumerist society. The poem “An April Manifesto”, for instance, demands that “April must replenish / Our bank-account of vanity” (l. 10). Similarly, in “Christmas Shopping”, the shop windows “marshal their troops for assault on the purse” (l. 13). Interestingly, however, words connoting financial opulence are also used frequently to evoke a sense of happiness or of vitality. In “Snow”, for instance, the persona finds that the room “was suddenly rich” (l. 1). When the first shepherd in “Eclogue by a Five Barred Gate” recounts his blissful dream, he stresses that the light he saw was “delicate as the chink of coins” (l. 103). Additionally, MacNeice stresses in “Train to Dublin” that, by celebrating the transient phenomena which are enumerated towards the end of the poem, people can “find that they are rich and breathing gold” (l. 55). In the depiction of scenes from his marital life in *Autumn Journal*, moments of felicity are often associated with affluence. The persona recalls that “We slept in linen, we cooked with wine / We paid in cash” (“Autumn Journal VIII”, l. 38) and that “till life did us part I loved her with paper money” (“Autumn Journal I”, l. 67). MacNeice’s rejection of consumerism is apparent from poems such as “Christmas Shopping” and “In Lieu”, but the imagery that is used throughout his verse also prompts a reconsideration of the poet’s position towards the pursuit of wealth. MacNeice’s oeuvre, furthermore, contains many images of a gastronomic nature. In “Littoral”, for instance, the shore is compared to “[d]amson whipped with cream” (l.2). In “The National Gallery” the persona sees a “pink temple of icing-sugar” in a picture (l.12). The fragrant trees which were remembered from childhood in “When We Were Children” were “breakfast for the gluttonous eye” (l. 10). The poem “Constant” depicts the remnants of the various civilizations in Istanbul via a complaint that there are “[t]oo many curds on the meat” (l. 1). This cursory survey of imagery illustrates that MacNeice often manages to convey profound concepts effectively via objects taken from quotidian life.

Most critics agree that MacNeice's oeuvre can be divided into three broad phases.⁵²⁷ The poetry written during the second phase in the early 1950s is mostly considered to be of a more languorous and a less inspired nature. The data set that was developed in this study can be used to investigate whether or not the changes in the critical acclaim also correlate with specific formal properties of the poetry. Figures 8.1 and 8.2 both suggest that specific clusters of poems are distinctive within the context of MacNeice's oeuvre not because of a particularly high or low value for a specific metric, but, rather, because of a lack of variation. As a lower variation might also imply a more predictable and a less creative use of devices, the hypothesis may be formulated that there is a correlation between the perceived quality of the poems and the ability of the poet to vary the application of literary techniques.

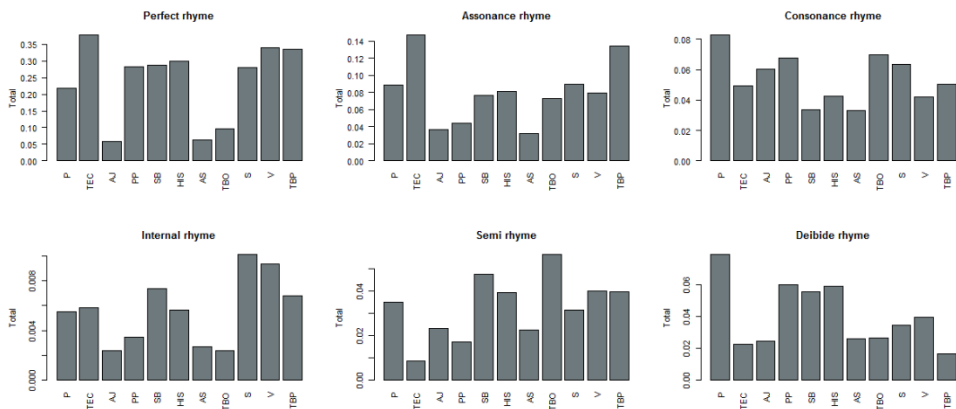


Figure 8.5. Standard deviations of volume-level counts of perfect rhyme, assonance rhyme, consonance rhyme, internal rhyme, semi-rhyme and deibhide rhyme

To explore the hypothesis more precisely, standard deviations have been calculated for a number of variables in all eleven volumes.⁵²⁸ The results of these

⁵²⁷ Alan Gilles explains that “[t]he contours of Louis MacNeice’s career are rarely contested: from the high point of his nineteen thirties work, reaching a crescendo with *Autumn Journal* (1939), he drifted into a slump, reaching a nadir with two collections from the early nineteen fifties, *Ten Burnt Offerings* (1952) and *Autumn Sequel* (1954), before reviving to develop a startling new style at the end of the decade”. See Alan Gillis, ““Any Dark Saying”: Louis MacNeice in the Nineteen Fifties”, p. 105.

⁵²⁸ Standard deviations are calculated by taking the square root of the variance of all the values. The variance, in turn, is produced by calculating the average of the squared differences between all values and their mean. The standard deviation gives an indication of the variation within a collection of data values. When its value is close to zero, this implies that there is little variation. High values, conversely, indicate that values are more diverse.

calculations are shown in figure 8.5. Interestingly, it can be observed that *Ten Burnt Offerings* and *Autumn Sequel* largely have the lowest values of all volumes. While some other volumes combine low volumes for one variable with high values for other variables, the poetry of the early 1950s scores lowest on almost all metrics. An exploration of the variations in the occurrences of literary devices cannot fully account for the differences in critical acclaim, however. *Autumn Journal* is frequently regarded as one of MacNeice's best works, but the standard deviations that were calculated for this collection are largely on the same level as the poetry of the early 1950s. The stasis in the usage of literary devices can partly be explained through the fact that *Autumn Journal*, *Autumn Sequel* and *Ten Burnt Offerings* all contain long poems with relatively consistent rhyme schemes. *Autumn Sequel*, for instance, consistently uses the terza rima rhyme form, in which the first and the third line of each triplet rhymes. For this reason, it must be concluded that the critical acclaim does not correlate completely with the variability in the use of literary devices.

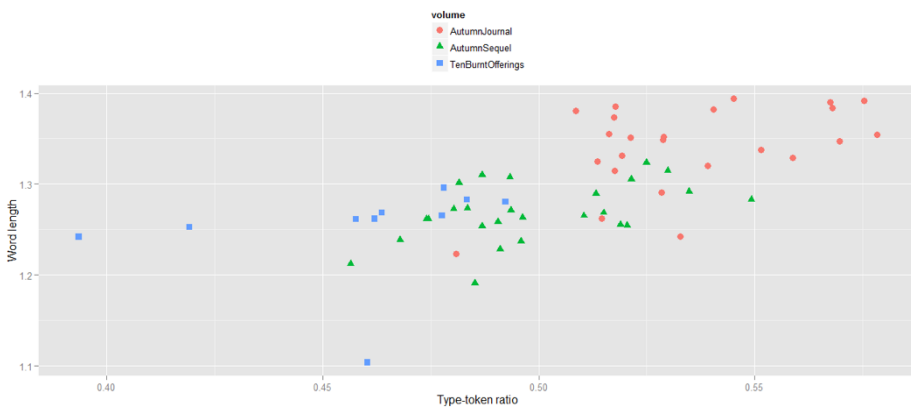


Figure 8.6. Type-token ratios and average number of syllables per word for all poems in *Autumn Journal*, *Autumn Sequel* and *Ten Burnt Offerings*

Prompted by the cursory impression that the language that is used in *Autumn Sequel* and in *Ten Bunt Offerings* is generally more rambling and more stagnant than that of *Autumn Journal*, I have created a scatter plot which shows both the type-token ratios and the average word lengths for all the poems in these three volumes. The word lengths were calculated by dividing the total number of syllables in each poem by the total number of tokens. Figure 8.6 indicates that the values that were calculated for *Autumn Journal* clearly differ from those of the two collections that were published in the early 1950s. On the whole, the poems in *Ten Burnt Offerings* and in *Autumn Sequel* have lower type-token ratios, which implies that there is also a lower diversity in the vocabulary. Additionally, the poems in the

latter two collections poems also contain words with fewer syllables. *Autumn Sequel*, in fact, contains many monosyllabic lines, such as “The wits of Bath for all their sense of form” (“Canto XXI”, l. 105), “Put back no clock; clocks were made for men” (“Canto I”, l. 16) or “We are bound to live and give and make and act” (“Canto XVI”, l. 52). On the basis of these findings, it may be hypothesised that that this more repetitive verbiage, and the many reiterations of monosyllabic words, strongly contribute to the overall sense of aridity in the latter two collections.

The analyses that have been performed to this point largely concentrated on the differences and the similarities between the various collections of poems. As was also alluded to in the introduction to this chapter, however, computational methods also stimulate scholars to study the text in a corpus as nodes in an interconnected network, and to abate the original boundaries formed by the physical books and volumes which originally contained these texts. Writing about theories of literature, Wellek and Warren distinguish approaches which view literature as “a series of works arranged in a chronological order and as integral parts of the historical process”, on the one hand, and approaches which view literature as “a simultaneous order”, on the other.⁵²⁹ In this latter approach, texts are juxtaposed on the basis of their stylistic features and regardless of when they were written. Computational methods support both approaches to the study of literary works. Using the data that were generated, poems can be grouped on the basis of a correspondence in formal properties, and the individual poems that are set apart in this manner can subsequently be examined in a more targeted manner.

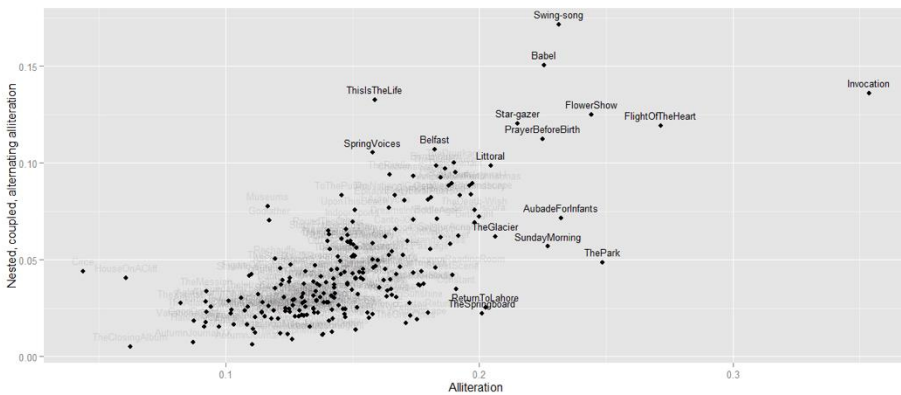


Figure 8.7. Normalised counts of alliteration in general and normalised counts of specific forms of alliteration

⁵²⁹ Rene Wellek & Austen Warren, *Theory of Literature*, p. 39.

Figure 8.7 shows the normalised counts of all occurrences of alliteration for each poem, together with the counts of the structured types of alliteration, in which two pairs of matching sounds are combined in specific ways. Notably, some of the poems which have high scores for alliteration also have similar themes. “Invocation” and “Flower Show”, for instance, both describe a withdrawal from a physical reality, followed by an escape into a dream-like or imaginative world. The heavy alliterations in “Invocation” (e.g. “Fetch me far a moon in a tree / Fetch me far a phrase of the wind”, ll. 15-16) and in “Flower Show” (e.g. “Fanged or whaleboned wattled or balding brimstone or cold / As trout or seaweed these blooms ogling or baneful all”, ll. 8-9) have an incantatory effect which helps to evoke the otherworldly nature of the environment which is depicted. Alliteration is used in a comparable manner in “Littoral”. In this poem, the persona is swept away by the many associations that are provoked by the image of the sea. Two other poems which are highlighted in figure 8.7 are “Babel” and “Flight of the Heart”. Next to the fact that these poems both feature a tower, they are both about an impulse to move away from social and political realities. The former poem uses the myth of Babel to underscore the social conflicts that can exist between the denizens of a modern city, and between “[p]atriots, dreamers, die-hards, theoreticians” (l. 13). “Flight of the Heart”, which describes an urge to fully retreat into the cellar of a tower, may be read as a reflection on the relation between the self and the external world. Terence Brown stresses that the separation between the self and its physical environment is a “foregone conclusion” in “Flight of the Heart”. In “Babel”, however, the questions “Can’t we ever, my love, speak in the same language?” and “Have we no aims in common?”, which are repeated insistently in each stanza also implies a hope that the distances can ultimately be bridged.

The historical overview of the occurrences of literary devices in figure 8.3 suggests, among other things, that semi-rhyme has not been used consistently throughout MacNeice’s career. It may be instructive, for this reason, to concentrate more closely on the poems that contain examples of this particular form of rhyme. Figure 8.8 displays the poems which, when the corpus is sorted according to the scores for semi-rhyme, belong to the upper five percent. Semi-rhyme evokes the impression of an order which is incomplete, or of a uniformity which emerges illogically from chaos. This device is often used in poems which depict an eerie atmosphere, or to evoke a confusing or a chaotic reality. “The Pale Panther” has one of the highest scores for this particular form of rhyme. In this poem, MacNeice connects “lamented” to “bent” and “cure” to “surely”. MacKinnon describes the poem as one of MacNeice’s nightmares, in which “the mood of bleak despair is not balanced by any of his old sardonic optimism”.⁵³⁰ The semi-rhymes in “Bottleneck” and “Spring Voices” have similar effects. The former poem argues that social and

⁵³⁰ William T. McKinnon, “MacNeice’s Pale Panther: An Exercise in Dream Logic”, in: *Essays in Criticism*, XXIII:4 (1973), p. 388.

political questions are often too complex or too confounding to be able to take a firm definitive stance. The main character's idealistic but confused politics aim to be "combined / Into a working whole but cannot jostle through / The permanent bottleneck of his highmindedness" (18-20). In "Spring Voices", the uneasy association of words such as "air" and "warily" (ll. 1-3) effectively reinforce the central opposition between the lures of the sunlight and the new spring on the one hand, and the nihilism and the apathy of the "householder" which is depicted on the other. Like "Spring Voices", "Nostalgia" describes a struggle between two impulses. The urge to focus on actuality and on facts can be countered by a pressing "longing / For what was never home" (ll. 3-4). This conflict is mirrored, to some extent, between the uncomfortable conjunction of the words "lull" and "vulnerable" and "slow" and "lonely".

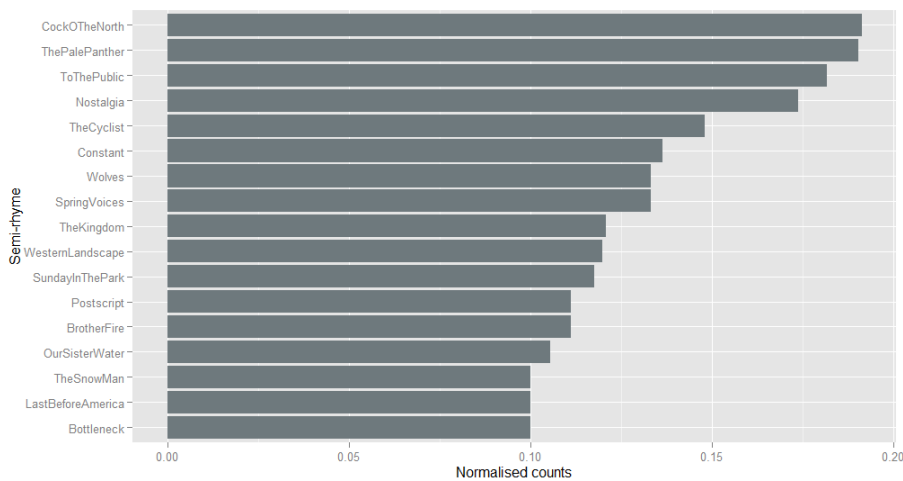


Figure 8.8. Poems with the highest normalised counts for semi-rhyme

A similar form of analysis can be performed to study regularities in the usage of deibhide rhyme. Figure 8.9 lists the individual poems which have the highest values for this specific device. The deibhide rhyme juxtaposes pairs of words which sound similar but which nevertheless do not agree fully. Reviewing the poems that are culled in figure 8.9, it may be concluded that such sonic agreements between words with dissimilar stress patterns have frequently have been used in poems which are concerned with social differences and with interpersonal conflicts. "Conversation" and "A Contact" form two clear examples. In the former poem, MacNeice proposes similarities between the words "way" and "yesterday" and "straight" and "interpolate". The poem concentrates on the uneasy moments during which the customary superficiality of social interactions is interrupted by inadvertent outbursts of genuine emotions. "A Contact", which rhymes "windows"

with “goes”, muses on the differences between a train that is chosen and the trains that pass by, and uses this imagery to evoke a sense of isolation and separation.

The notion that MacNeice frequently used deibhide rhyme to emphasise a strained relation between the self and the other is also exemplified clearly by some of the other poems which are enumerated in figure 8.9. “The Individualist Speaks”, for instance, contains two examples of deibhide rhyme: “As chestnut candles turn to conkers so we / Knock our brains together extravagantly” (ll. 5-6) and “Crawling down like lava or termites / Nothing seduces nothing affrights” (ll. 13-14). The poem as a whole imparts MacNeice’s mistrust of established creeds and of existing philosophical systems. Ideologies such as socialism are depicted as exhibits in a fair “pitched among the feathery clover” (l. 1). The persona, however, ultimately chooses to “escape, with my dog, on the far side of the Fair” (l. 16). The use of deibhide rhyme appropriately underscores the poem’s tension between communality and individuality. Deibhide rhyme is also used in the first and third stanzas of “Birmingham”. In the description of the slums which houses the factory workers, the “vista thins like a diagram / There unvisited are Vulcan’s forges who doesn’t care a tinker’s damn”. In the third stanza, the “shopgirls’ faces relax / Diaphanous as green glass empty as old almanacs”. Rhymes of stressed and unstressed syllables evoke a sense of inequality, and “Birmingham” productively uses this device to accentuate social distances. Along similar lines, “Bar-room matins” describes an inability or an unwillingness to empathise with the victims of World War II. It connects the words “sky” and “alibi” (ll. 22-23) and “sea” and “commodity” (ll. 26-27).

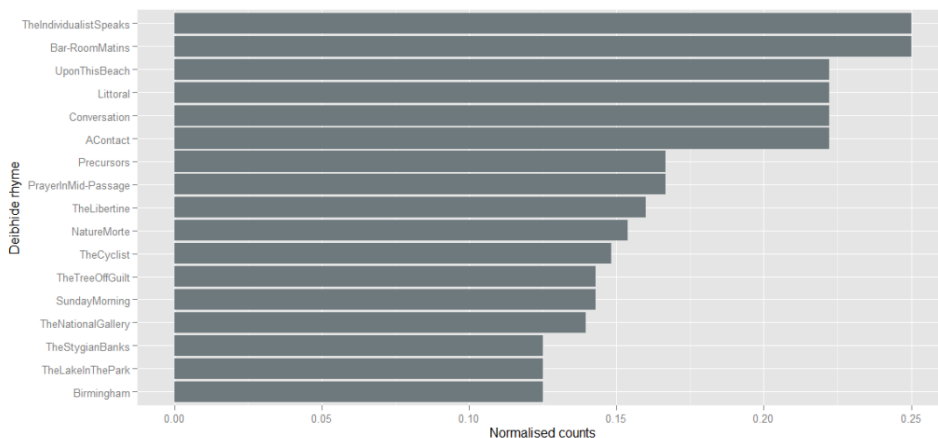


Figure 8.9. Poems with the highest normalised counts for deibhide rhyme

The list that is shown in figure 8.9 was produced by sorting the poems on the basis of their normalised counts. Figure 8.9 only lists 17 out of the 116 poems which contain instances of deibhide rhyme. During a closer inspection of the remaining

poems, several other striking instances were found. It is interesting to observe, for instance, that this device, which was derived from the Celtic verse tradition, is also used in “The Hebrides” and in “Valediction”, which both describe parts of the British Isles with a strong Celtic heritage.⁵³¹ “Valediction” describes the city of Belfast as “devout and profane and hard / Built on reclaimed mud, hammers playing in the shipyard” (ll. 15-16) and warns all tourists in Ireland not to “pay for the trick beauty of a prism / In drug-dull fatalism” (ll. 65). The two poems express dissimilar messages about the regions they portray, however. “The Hebrides” idealises a traditional and rural way of living, and also hints at some of the negative implications of modernity and consumerism. “Valediction” does not idealise Ireland’s past, but emphasises the asphyxiating effects of Ireland’s obsession with its history and with its religion. Notable instances of deibhide rhyme can also be found in “Les Sylphides” from “Novelettes”, which rhymes “ballet” with “grey”, and in “The Libertine” which connects “alone” to “telephone”. The latter two poems provide additional support for the claim that deibhide rhyme is often used in passages which concentrate on interpersonal distances. “Les Sylphides” depicts a failed marriage, and “The Libertine” explores the bewilderment of a man who, after a life full of promiscuity, only desires to be left alone. In the context of algorithmic criticism, the objective of data analysis is generally to expose aspects of literary texts which can enkindle interpretation, and, in view of this aim, algorithms ought to be able to detect remarkable instances of literary phenomena. Given a particular data set, however, it can often be challenging to develop an effective method for the discovery of significant fragments. Since the ability to advance interpretation seems difficult to quantify, an approach which is based solely on the number of occurrences does not necessarily form a suitable heuristic for ensuing interpretative readings.

The comprehensive data set that was developed in this study can productively be actuated to expand the results of previous critical studies. In a discussion of *Leaving Barra*, Neil Corcoran notes that the poem consists almost exclusively of feminine line endings. Corcoran argues that the lowering intonation in these lines, in which a stressed syllable is followed by one or more unstressed syllables has a melancholy connotation, evoking “courage in the face of loss, something falling poignantly away”.⁵³² Using the data set that was developed in this study, it is possible to investigate if such line endings have similar effects in other poems. Via

⁵³¹ Next to deibhide rhyme, these two poems also contain other examples of devices which, according to Skelton, were inherited from Celtic verse. “The Hebrides” contains both regular aicill rhyme (“Are glad to have their land though mainly stones / The honoured bones which still can hoist a body”, ll. 52-53) and a specific form of aicill rhyme in which the final consonants of a line are repeated in the line that follows (“And all the neighbours celebrate their wedding / With drink and pipes ...”, ll. 84-85). A similar example can be found in lines 51 and 52 of “Valediction”: “Of green marble or black bog-oak run up to Clare / Climb the cliff in the postcard visit Galway city”.

⁵³² Neil Corcoran, “The Same Again? Repetition and Refrain in Louis MacNeice”, p. 215.

the phonetic transcriptions that were made available, poems with similar properties can be identified relatively swiftly. For each poem, the lines which end in unstressed syllables were counted, and, after this, these counts were divided by the total number of lines. Figure 8.10 displays the 20 poems with the highest percentages of feminine line endings. The yellow lines indicate feminine line endings, and the blue lines indicate masculine line ending. Most of the poems in this group of texts indeed focus on a loss. “June Thunder” concentrates on the poet’s divorce from his wife, and the stream in the final stanza of “The Rest House” is read by Brown as “some allegorical river of life, winding from birth to death”.⁵³³ “New Jerusalem”, furthermore, depicts the city of London, which changed from a place that could excite and inspire into a location full of “compulsive, illusory distractions born out of a society more firmly tethered to a rampant consumer culture”.⁵³⁴

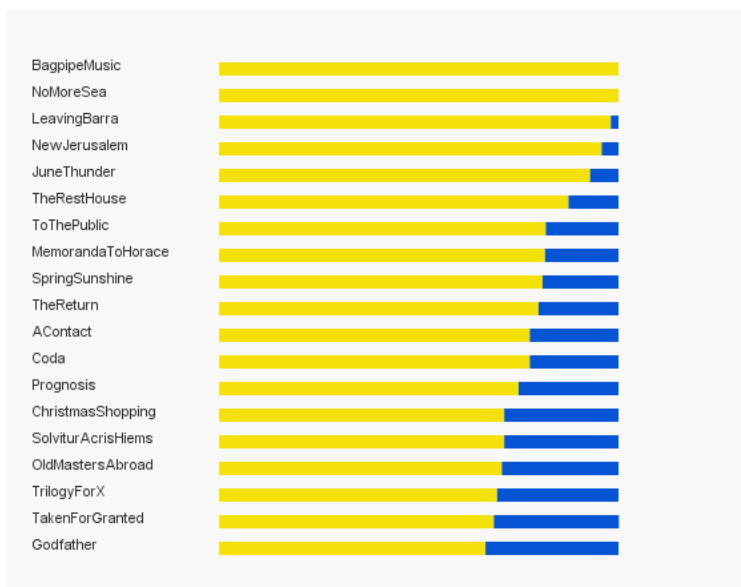


Figure 8.10. Proportions of feminine and masculine line endings

Interestingly, the three poems in the corpus whose final words are almost exclusively feminine all focus on islands. “Bagpipe Music” focuses on the decay of a harmonious rural culture on the Western Scottish islands, and on the transition to

⁵³³ Terence Brown, *Louis MacNeice: Sceptical Vision*, p. 141.

⁵³⁴ Simon Workman, ““To Be Tired of This Is to Be Tired of Life”: Louis MacNeice’s London”, in: *Irish Writing London: Volume 2*, Tom Herron, London: Bloomsbury Academic, p. 139.

a consumerist and violent society. Brown notes that the poem has “a nihilistic cruelty in its pointless violence”.⁵³⁵ “Leaving Barra”, as discussed by Corcoran, describes the poet’s continuing need for religious beliefs or for elucidatory theories. The departure from the island of Barra symbolises the concurrent inability to commit to such creeds. The thematic concerns of “No More Sea” are strongly related to those of “Leaving Barra”. In the latter poem, the sea represents a “vast nothingness”.⁵³⁶ Human beings can shield themselves against such nihilism either by devising individual guidelines, or by joining the accepted rules and dogmas of a larger group. “No More Sea” represents the clash between individuality and communality as the difference between islands and the mainland, and advocates the value of islands. In all these three poems, the islands represent a moral or spiritual superiority. The poems also describe the annulment of these values, either by modernity or by physical leavetaking. The feminine rhymes, which, as Corcoran stresses, evoke the notion of loss, poignantly mirror such processes.

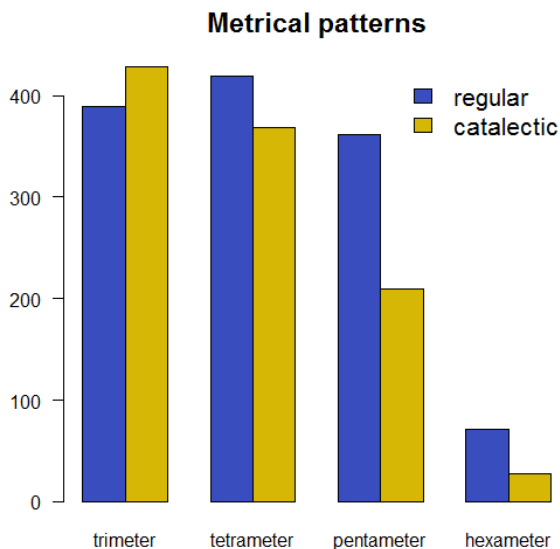


Figure 8.11. Absolute counts of regular and catalectic trimeters, tetrameters, pentameters and hexameters

As was noted above, having data about all the texts in a corpus enables scholars to investigate whether or not characteristics observed in individual poems are also representative of the full corpus. Traditional close reading tends to focus on a

⁵³⁵ Terence Brown, *Louis MacNeice: Sceptical Vision*, p. 152.

⁵³⁶ *Ibid.*, p. 119.

limited number of texts, but by making use of computation, scholars can investigate the profusion of phenomena in a more circumferential manner. Digital methods may be used, for instance, to verify Robin Skelton's claim that many of MacNeice's metrical patterns are irregular. Skelton argues that the preference for an odd rather than an even number of syllables is a reflection of the poet's Celtic background. He uses the catalectic patterns in "The Sunlight on the Garden" to illustrate his argument. Allan Gilles notes, in a similar vein, that the poem "All Over Again" contains many skewed lines because of the addition of an extra stress.⁵³⁷ To investigate the claim that MacNeice has a preference for irregular metrical feet, both the regular and the catalectic varieties were counted of all the verse lines whose metrical patterns could be classified automatically. Separate counts were also produced for the various line lengths. As can be seen, the catalectic pattern is used more frequently than the regular pattern in the case of trimeters only. In all other line lengths, the regular meters surpass the catalectic patterns. The diagram does clarify, however, that, although the catalectic lines do not dominate in absolute numbers, these incomplete patterns are still used highly insistently. To be able to interpret the results more thoroughly, nevertheless, data on the use of catalectic verse lines by other poets would clearly be valuable, as these could allow for a comparison of the use of incomplete metrical patterns among English and Irish authors.

In this case study, a number of experiments were also conducted to explore the relevance of sentiment analysis. Within the *Multi-Perspective Question Answering* (MPQA) tool, the *Harvard General Inquirer* package and the sentiment analysis tool that was developed by Bing Liu, lexicons have been developed with words that have positive and negative connotations. These lexicons have been made available separately, but, for the purpose of this study, these have been merged. On the basis of these combined lexicons, the tokens in this study's central corpus have been tagged. Subsequently, for each poem, the number of words with a positive and a negative connotation have been counted. The counts were normalised by dividing these by the total number of tokens. Next, the values for the negative words were deducted from the values for the positive words. Totals that were higher than zero were understood to indicate a chiefly buoyant register of speech, and negative values indicate texts that are overtly negative in tone. This relatively simplistic approach clearly has shortcomings. Words can express differing degrees of negativity, but, in this study, all positive and negative words were counted equally. Secondly, this analysis was based on counts individual words in a context-independent manner, and it failed to take into account the effect of negations. The simple counts of words from the lexicons can help, nevertheless, to produce a rough approximation of the connotations of the vocabulary.

⁵³⁷ Alan Gillis, "'Any Dark Saying': Louis MacNeice in the Nineteen Fifties", p. 118.

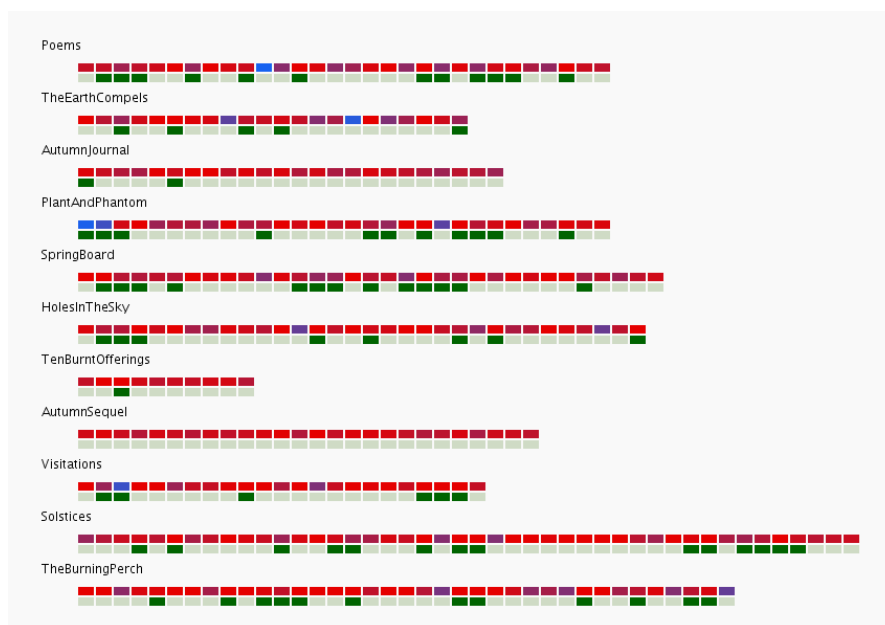


Figure 8.12. Results of sentiment analysis performed on all poems in the corpus

The results of the analysis are shown in figure 8.12. In this diagram, the blocks represent individual poems, and the overall sentiment of each poem is indicated by colours in a gradient scale. A predominantly negative vocabulary is shown using a red colour, and poems with a positive tone are rendered blue. The diagram also displays information on the themes of the poems, which had been assigned manually. If a poem was found to concentrate on decidedly negative themes, such as doom, nihilism and alienation, this was represented by a green block, shown directly below the bar that indicates the result of sentiment analysis. While, in most cases, the sombre themes are also described using words with negative connotations, as expected, there are also a number poems in which the sinister or morose themes are combined with mainly positive terms. Examples of poems in which a positive vocabulary coalesces with negative themes include “Wolves”, “Prognosis” and “Stylite”. The poem “Wolves” concentrates on the futility of critical reflection and evokes a fear of “the wolves of water”, which, as in many other poems, represents a nothingness and a nihilism. The text contains positive words such as “laughter” and “want” as it also proposes an antidote to nihilism. One solution is to unite with fellow human beings, and to imagine that dangers do not exist. The fear of disorder and of disintegration is also stressed effectively by the use of semi-rhyme. “Prognosis” is set during the end of a winter, and the persona reflects on what the new spring will bring. The poem contrasts a number of

favourable and disagreeable options. Since the poem ends with a negative outcome (“Or will his name be Death / and his message easy”), the poem as a whole evokes a sinister sense of doom, connected to the imminence of the Second World War. The persona counteracts this fear by a belief in an alternative, more pleasurable reality. “Stylite” similarly presents opposites. It contrasts the central tenets of Christian religion, which demands austerity and a denial of earthly pleasures, with Greek philosophy which fully embraces the physical world. The positive words are mostly used in the description of the Greek god, which sits opposite the hermit on the stele. The poem as a whole describes a solipsism, and an isolation of the self from the world. A general pattern that may be observed is that poems with negative themes may also contain imaginative evocations of optimistic alternatives which counteract the negativity.

Figures 8.8 and 8.9 above list the poems with the highest normalised counts for specific forms of rhyme. Such analyses, which concentrate on literary devices in isolation, can often help to clarify the effect of these literary techniques within specific texts. As poems frequently contain salient combinations of different forms of rhyme, it can also be instructive to explore the values for these various forms of rhyme simultaneously. Figure 8.13 lists the texts which contain the high values for perfect rhyme, in combination with the counts for assonance rhyme and consonance rhyme.⁵³⁸ The corpus contains 16 poems in which all the lines rhyme perfectly, and these have been ignored.⁵³⁹ When a high percentage of one type of rhyme co-occurs with a relatively low percentage of another type of rhyme, this may warrant a more detailed close reading, as it is frequently the case that the lines that deviate in notable ways from a pattern which is otherwise regular are highlighted for a particular purpose. A clear example is “Autobiography”, which has a score of 0.85 for perfect rhyme. All couplets contain perfect rhyme, with the only exception of the last couplet: “I got up; the chilly sun / Saw me walk alone”. It may be argued that, while the stanzas that rhyme perfectly evoke the innocence and the simplicity of childhood, the breach of this pattern at the conclusion of the poem emblematises the psychological damage with which the poet entered adulthood.

Instances of perfect rhyme and slant rhyme have likewise been combined effectively in many of the other poems which have been identified in figure 8.13. In “The Grey Ones”, the lines that rhyme perfectly evoke a traditional and increasingly outmoded form of living. Peter MacDonald argues that the state of greyness functions “as the drab background to modernity”.⁵⁴⁰ The evocation of modern life in

⁵³⁸ The 20 poems with the highest values for perfect rhyme did not contain any semi-rhymes.

⁵³⁹ The poems in which all the lines rhyme perfectly are “Chess”, “Greyness is All”, “Birthright”, “The Revenant”, “Beni Hasan”, “Whit Monday”, “The Ear”, “Figure of Eight”, “Mutations”, “Passage Steamer”, “Place of a Skull”, “The Heated Minutes”, “Bar-Room Matins”, “Bluebells”, “The Mixer” and “To Hedli”.

⁵⁴⁰ Peter MacDonald, “Louis MacNeice: The Burning Perch”, in: Neil Roberts (ed.), *Companion to Twentieth-Century Poetry*, Hoboken: Wiley-Blackwell 2003, p. 496.

the lines “Sprawled against the Gates of Doom / Whence all kebabs and cockstands come” (ll. 33-34) discernibly intrude into the surrounding lines. In the short poem “To Posterity”, the first line (“When books have all seized up like the books in graveyards”) is isolated from the rest of the poem because it is not part of the perfect rhyming scheme. This exclusion of the lines that describe traditional books stress the central fear in the poem that communication via written texts will be replaced by non-verbal media. In “Off the Peg”, line 8 (“Chameleons can adapt to whatever sunlight leaks”) is the only line which does not rhyme perfectly with any of the other lines. The poem as a whole describes the sensation that hackneyed phrases and worn-out habits can occasionally feel new and very relevant. The line which breaks with the regular pattern is also the point at which the poem begins to discuss the revitalisation of clichés.

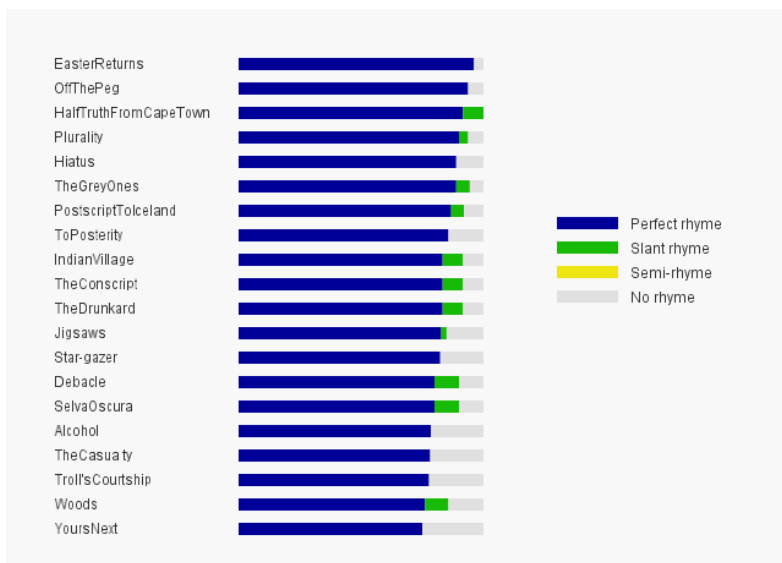


Figure 8.13. Percentages of lines that contain perfect rhyme, slant rhyme and semi-rhyme. The list was sorted using the values for perfect rhyme

The list which is shown in figure 8.13 is headed by “Easter Returns”. The poem focuses on the increasingly profane nature of religious festivities. The last word of each stanza rhymes with the first line of the stanza that follows. The last stanza forms the only exception to this regular structure. The schism in the formal structure of the poem suggests that, although traditional religious practices can no longer be followed, man can still find new forms of religious inspiration. Through such changes, man may experience that “the myth returns” (l. 22). The function of the rhymes and the non-rhymes in “Easter Returns” can be clarified further by linking this poem to “Place of a Skull” and to “Whit Monday”, which exclusively

have lines which rhyme perfectly. These poems, like “Easter Returns”, focus on a loss of religion. “Place of a Skull” recounts the banal and irreverent reaction of the Roman soldiers to the death of Christ. The fact that the coat that was found on the dead body did not fit the soldiers epitomises the shift to a more nihilistic and a more secularised worldview. “Whit Monday” emphasises the idea that religious traditions and creeds have lost their relevance and that they fail to banish anxieties caused by modern life. These three poems on the disintegration of Christian faith effectively use regular perfect rhymes to underline the spiritual emptiness of religious practices.

A similar diagram can be produced to identify poems in which high values for slant rhyme have been combined with other forms of rhyme. One poem in which occurrences of slant rhyme alternate regularly with lines that contain perfect rhyme is “The Habits”. The poem has a very intricate structure, consisting of five stanzas, of five lines each. Two of these lines end in words that rhyme perfectly (“best” and “dressed”). Each stanza also has one line that ends in the word “habits”. The second, third and fourth stanzas contain two additional lines whose final words are loosely connected to the word “habits” via assonance rhyme (“carried”, “hypodermic”, “affidavit”, “sjambok”). The alternation between predicative perfect rhymes and the more uneasy assonance rhymes suggests that personal habits and social conventions can both be comfortable and vexatious. The regular pattern is discontinued in the final stanza. The absence of rhyming words in the poem’s last lines undergirds the notion that, in later life, these enforced habits “[o]utstayed their welcome” (l. 22).

Slant rhyme and perfect rhyme have also been combined intriguingly in “Dogs in the Park”, which centres around an opposition between the dog owner’s desire to control and to tame the dogs on the one hand and the savage and freedom-seeking dogs on the other. These stanzas have two lines which rhyme perfectly and two lines which contain slant rhyme or no rhyme. This structure reflects the tension between the urge to impose order on the one hand, and the desire to break away from this control on the other. Another poem which contains conspicuous combinations of slant rhyme and perfect rhyme is “The Rest House”. The poem, according to Terence Brown, evokes a scene which is “nightmarishly alive” and in which “objects have an unpleasant, spontaneous life of their own”.⁵⁴¹ The notion that the room has normal objects which nevertheless behave unexpectedly is bolstered by the form of the poem. “The Rest House” consists of two stanzas, and they both contain one instance of perfect rhyme (“veranda” and “Uganda” and “much” and “such”). The strangeness of the imaginative world is mimicked by the words which are coupled loosely via slant rhyme: “window” and “river” in the first stanza and “splintered”, “filtered” “children” and “extinguished” in the second stanza.

⁵⁴¹ Terence Brown, *Louis MacNeice: Sceptical Vision*, p. 141.

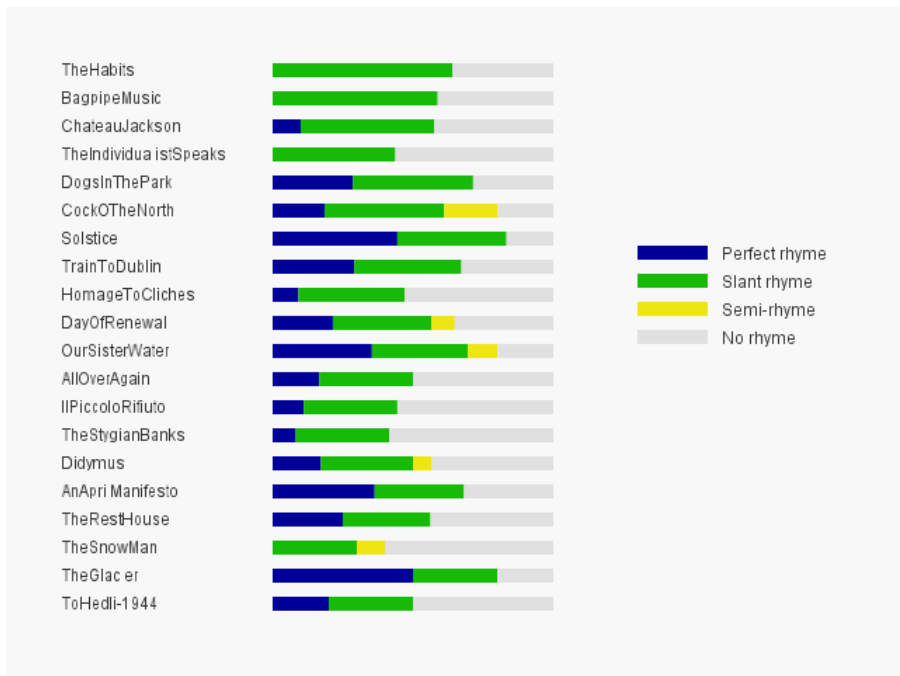


Figure 8.14. Percentages of lines that contain perfect rhyme, slant rhyme and semi-rhyme. The list was sorted using the values for slant rhyme

The various data values that have been produced about the formal aspects of MacNeice's poetry can be used productively to assemble texts which make use of the same literary techniques, and within such clusters of poems, scholars may investigate whether or not these poems also convey related ideas. Such considerations of the differences and similarities between texts can often quicken an interpretative reading of these texts. In the analyses that have been discussed above, poems have only been clustered on the basis of one or two variables, but, as this study has produced values about a wide range of different textual aspects, the similarities between different texts can also be investigated in a more encompassing manner. In this study, I have developed a method for the calculation of similarities between poems which was based on a concurrent analysis of 15 variables. For each poem, a vector was produced listing the values which were calculated for perfect rhyme, internal rhyme, alliteration, consonance, assonance, consonance rhyme, assonance rhyme, semi rhyme, deibhde rhyme, aicill rhyme, consonantal repetition, paronymy, iambic rhythms, trochaic rhythms and anaphora. The values that have been generated in this study could not all be compared directly, however, as a result of the fact that the different values have also been normalised in different ways. Values in the data set reflected either the number of

counts per item, or the number of counts per verse line. To address these complications arising from these different units of measurement, all values have been rescaled through the calculation of z-scores.⁵⁴² An important advantage of working with z-scores is that they are essentially dimensionless. They are calculated by dividing all distances from the mean by the standard deviation. Both of these values are without a unit of measurement.

A first attempt at identifying related poems was based on a calculation of the cosine similarity for each possible pair of poems within the corpus, using the values for the 15 variables that were mentioned. Cosine similarity expresses the degree of similarity between two separate vectors as a value in between zero and one. Using this method, it was found that 118 pairs of poems displayed a cosine similarity of 0.95 or higher. On closer inspection of these results, it was found, however, that these cosine similarities were often high for poems which share a value of 0 for devices such as paronymy, semi-rhyme, and anaphora, which have only been used in a limited number of poems. In this study, the aim of identifying formal similarities was ultimately to explore whether or not the formal aspects of the poems correspond to what the poem intends to convey. Given this objective, it was more fruitful to concentrate on texts in which these literary techniques have actually been used, and to disregard poems without any occurrences. For this reason, an alternative method has been developed. The analytic algorithm that was used to produce figure 8.15 firstly calculates, for each pair of poems, the difference between all 15 values. Values were considered to be similar if the absolute difference was found to be less than 0.05, and, importantly, if the values for these variables were higher than 0. The network in figure 8.15 shows the poems in which at least three variables have such matching values. In these calculations, I have decided to leave the long compositions *Autumn Journal* and *Autumn Sequel* out of consideration.

Figure 8.15 connects poems which are formally similar. “Mayfly” and “Day of Renewal”, for instance, both contain internal rhyme: “One only day of May alive beneath the sun” (“Mayfly”, l. 4) and “Blood cholera flies blank eyes becoming forty” (“Day of Renewal”, l. 32). They have similar percentage of alliterating words, and they both contain aicil rhyme: “Flowers in the sun beside the jewelled water / Daughter of the South call the sunbeams home” (“Mayfly”, ll. 6-7) and “Claiming a different birthplace a wild nest / Further more truly west on a bare height” (“Day of Renewal”, ll. 13-14). There is also a thematic connection between the two poems. “Mayfly” muses on the transience of valuable experiences, using the brevity of the lifespan of mayflies as a poignant metaphor. The short lives of the mayflies, which are nevertheless stretched out “taut and thin and ringing clear”, inspire an urge to

⁵⁴² Z-scores indicate the distance from the mean of the data-set, expressed in standard deviations. As such, they offer a good indication of the position of a particular value within the entire distribution. See Roger Bilisoly, *Practical Text Mining with Perl*, p. 193.

savour precious moments by regarding these as moments which exist outside of time. At the same time, the poem admits that such permanence and timelessness are impossible, emphasising that such experiences in which time is made “elastic” can only exist as the result of a creative transformation. Like “Mayfly”, “Day of Renewal” reflects on the passing of time. The poem describes a number of past birthdays and notes that the human mind can transform events and locations into something permanent and static, while these continue to evolve in reality. A place which is recollected “in itself has changed but in our mind / Does not become but is” (ll. 27-28). “Day of Renewal” recommends a response to the transience of life which differs from the advice which is given in “Mayfly”. While the latter poem urges us to counteract the “pathetic fallacy of the passing hours” by using our imagination, “Day of Renewal” argues that evolution and moments of renewal need to be accepted, as these are essential to being alive. It stresses that “[d]eath is” and that “life becomes” (l. 19).



Figure 8.15. Similarities between poems, based on a concurrent analysis of 15 variables

Although systematic explorations of the formal similarities between texts cannot expose the thematic or the conceptual links between poems in themselves, such comparisons of linguistic properties or of literary techniques can often lead to

juxtapositions of texts which may not have been related otherwise. Reading poems against the background of other poems can often further their interpretation. Figure 8.15 interestingly lists many poems which, next to their formal similarities, also share certain thematic concerns. “Cradle Song for Eleonor” and “Bar-room Matins”, for example, both concentrate on escapism. The speaker in the first poem urges the listener to look away from the “pity of it all” (l. 22). “Bar-room Matins” is written against the backdrop of the Second World-War and likewise describes a conscious but alcohol-induced effort to ignore the “doom tumbling from the sky” (l. 22). While the algorithm that was developed to establish similarities between poems did not consider their vocabulary, it coincidentally connected two poems whose titles share an adjective. “Slow Starter” and “Slow Movement” both contain trochaic lines, they have the same percentage of modal verbs and they have the same value for consonance. “Slow Starter” centrally describes a conflict between two conceptions of time. The first view posits that there is always time in abundance, and assumes that good things will come to those who are patient. This trite view is challenged by a second view which stresses that time is scarce, and that it is important take decisive actions. “Slow Movement”, from *Holes in the Sky*, likewise describes two different ways of experiencing time. It depicts a persona who sits opposite an unknown woman who is asleep. The persona appears to image that he features in her dreams. The dream sequence appears to exist outside of time, with “yesterday cancelled out” (l. 3). The end of the dream, which is described in the final lines of the poem, also marks a return to the fast-paced time of external reality, “[a]ccelerando con forza” (l. 23).

Striking commonalities were also found between “Homages to Clichés” and “Woods”. “Homages to Clichés” initially presents the social conventions that make interpersonal communication predictable as comfortable and appealing, but, as the poem progresses, the clichéd interactions increasingly begin to amplify the distances between the two persons who converse. The stale and banal conventions which are enacted also represent an inability to reflect critically and originally, and the poem eventually expresses a longing for a radically nonconformist world in which “[n]ext year is this year, sometime is next time, never is sometime” (l. 79). “Woods” contains a very similar progression from glorification to debasement. While the poem originally presents the English woods as kingdoms “free from time and sky” (l. 10), the penultimate stanza also accentuates the insipid and artificial nature of the tidy woods which were “assured / Of their place by men” (l. 25). “Homages to Clichés” and “Woods” both focus on controlled and predictable phenomena which can be satisfying to some extent, but which ultimately fail to stir meaningful or profound emotions.

Analyses of the similarities between poems can help to expose conspicuous aspects of MacNeice’s poetic language. The methods that were developed for this study did not primarily concentrate on the semantic contents of MacNeice’s poetry,

but on devices that are based on sounds. Edna Longley stresses, nevertheless, that MacNeice had a “continual drive to reconcile the demands of form and content”.⁵⁴³ Perhaps to an even greater extent than some of his contemporaries, MacNeice was deeply interested in the complex relations between sound and meaning. As has been shown in this chapter, a systematic analysis of the sonic patterns in the various verse lines can productively fortify an interpretative reading of MacNeice’s poetry.

8.3. Conclusion

8.3.1. Discussion of the case study

The case study that was conducted for this study primarily aimed to contribute to a general understanding of the concrete challenges that can emerge during processes of data creation and data analysis. Next to this main objective, however, it has been shown additionally that the various methods that are associated with literary informatics can effectively be applied to generate relevant new insights about the poetry of Louis MacNeice. In “Postscript to Iceland” MacNeice writes that his urge to find those texts that can truly inspire is frustrated by the feeling of being inundated by the many works that are available:

Rows of books around me stand,
Fence me round on either hand;
Through that forest of dead words
I would hunt the living birds

The stanza may simultaneously be read as an admission of the difficulty of finding a unique poetic voice amidst the literature produced by previous generations. This study aimed to chase “the living birds” in MacNeice’s own literary oeuvre by making use of algorithms. Computers can detect patterns within texts fragments with an acute exactitude and with an implacable consistency, and, utilising this capacity, this case study was able to disclose a number of aspects of MacNeice’s poetic language which had not been discussed yet in earlier studies. I have shown, for instance, that MacNeice often used alliteration in poems with an escapist surge, or in poems which ruminate on the distance between the self and the external world. The computer’s ability to detect patterns also helped to disclose the fact that MacNeice frequently made use of specific combinations of alliterative words, in which two pairs of sounds are repeated, or in which one pair of sounds is nested within a second pair of sounds. The methods that were discussed in this thesis also facilitated a systematic examination of all occurrences of semi-rhyme and deiphide

⁵⁴³ Edna Longley, “Louis MacNeice: Aspects of His Aesthetic Theory and Practice”, p. 151.

rhyme. The type of sonic agreements these devices are based on are often difficult to notice for human scholars. These forms of rhyme frequently occur in poems which concentrate on stark oppositions, or on the petrification or the inertia that results from an unruly plurality. Deibhíde rhyme is often used, more specifically, to describe social distances. This study has also demonstrated that the poems which lack consonance all describe a solitude, and that poems about the loss of religion often make a heavy use of perfect rhyme. None of the critical studies that have been investigated for this study, moreover, have commented on the importance of money, war and food imagery.

Computational methods can be applied effectively to investigate whether or not the qualities which earlier scholars have observed within singular text are also present in other texts. In this study, such critical expansions have often resulted in new observations. Robin Skelton has discussed the nature and the function of Celtic poetic devices such as deibhíde rhyme and aicill rhyme, illustrating his explanation via a close reading of the two poems “Aubade” and “Order to View”. Using computation, it was possible to study the profusion of these specific literary devices within MacNeice’s entire literary career. In combination with a more comprehensive assessment of the poet’s use of catalectic metrical patterns, such methodical examinations of deibhíde rhyme and aicill rhyme can lead to a sharper delineation of the Celtic influences in MacNeice’s work, and, ultimately, to an improved understanding of the poet’s relation to Ireland and to the Anglo-Irish tradition. Along similar lines, Neil Corcoran’s observation that the feminine line endings in “Leaving Barra” evoke a melancholy sense of loss formed the incentive for a wider exploration of all the poems in which the majority of lines have feminine endings. Interestingly, it was found that the poems whose line endings are almost exclusively feminine all depict islands. Computation can also be used to supply new forms of support for observations made by previous critics. Critics generally agree, for instance, that *Autumn Sequel* and *Ten Burnt Offerings* are distinctive within MacNeice’s oeuvre because of its concentration on philosophical topics, and because of the absence of engaging imagery. The case study has indicated, furthermore, that the poems from the early 1950s have a lower type-token ratio, and that they generally contain words with a lower number of syllables.

The case study that was conducted for this thesis did not aim to falsify the results of earlier studies. As was explained, most of the algorithms were inspired by observations which had been made in earlier studies, and the aim was often to extrapolate from claims which were made about individual texts. As the emphasis was primarily on occurrences and frequencies of literary devices, this study did not produce interpretations which were radically different from previous critical studies of MacNeice’s work. Algorithmic criticism, of the form that was illustrated by the case study, primarily aims to invigorate interpretative readings which, in line with the traditional close reading method, are based directly on the formal characteristics of the text. Computational methods can effectively identify texts with distinctive formal or structural properties, and such information about their

uncommon or their unique properties can often be used productively in critical assessments of texts. Digital methods cannot of themselves add interpretation, and for his reason, this study did not result in strong refutations of earlier interpretations. The toolset that is associated with algorithmic criticism cannot independently establish, for instance, that a poem which has conventionally been viewed as good literature is in fact of a poor literary quality. Statistical methods can occasionally lead to reassessments of texts, nevertheless. Many of the analyses that were conducted for the case study highlighted the statistical particularity of the poems "Place of a Skull", "Memoranda to Horace" and "Spring Sunshine". These poems have not been discussed at all in Edna Longley's extensive monograph on MacNeice's verse. Evidently, an unusually high or an exceptionally low frequency of literary devices does not necessarily imply that these devices have also been used in an interesting or in befitting manner. As qualitative methods allow for a more embrative and a more dispassionate form of research, they can prompt scholars to ask the question whether or not particular poems have been neglected deservedly.

8.3.2. Discussion of data analysis

In Chapter 4, it was argued that literary informatics mainly buttresses three basic scholarly primitives: annotation, comparison and discussion. Chapter 6 has discussed the main difficulties which can arise during the creation of annotations. This chapter has focused in more detail on the many different ways in which digital methods may support comparison and discovery. Using the practical work that was undertaken for this study as inspiration, two different types of comparison can be distinguished. Analytic methods can firstly compare the stylistic qualities of groups of texts in their entirety. In the case study, I have applied various analytic methods to explore the differences and similarities between 11 volumes of poetry, focusing more specifically on the differences between the volumes which were written during the early 1950s, and the volumes which were before and after this period. Comparative analyses may focus, secondly, on the differences and similarities between singular texts. When the results of analytic methods contain references to individual poem titles, these poems can subsequently be analysed more closely, using the information about the distinctive qualities of these texts as reading guides. In Chapter 2, literary texts have been described as linguistic compositions in which particular recondite concepts have been translated into formal features. If it is shown that a poem makes an extensive use of a specific literary device, such an observation may often be explained by relating some of the connotations of this literary device to the central thematic concerns of the text which contains this device. The statistical artefacts which are constructed, importantly, do not contain interpretation in themselves. The resources that result from algorithmic processing function as means rather than as ends, as these mainly form the materials with which scholars can fortify their interpretation.

Scholars who aim to compare texts can commonly choose from a broad range of statistical techniques, and it can often be taxing to select an appropriate method. The differences between two distinct classes of texts may be examined using supervised learning techniques, of which Student's t-test, logistic regression and Naïve Bayes all form concrete examples. When scholars aim to subdivide a corpus into smaller clusters, they can make use of k-means clustering, calculations of Euclidean distances, PCA or nearest neighbour analyses. These different methods are all based on different algorithms, and they consequently produce different results. Such differences can be subtle in some cases, but also quite dramatic in other cases. Even when scholars have decided to make use of one particular technique, they frequently have the possibility to influence the results by varying some of their initial parameters. In the case of classification, the results of the analyses can often be manipulated directly by varying the sizes of the training sets and the test sets. In this study, it was found, for example, that the nature of the network diagram displaying formal similarities between poems can change dramatically along with the variables which are considered.

Within the context of algorithmic criticism, there are few guidelines which can help scholars to make a reasoned selection. Statisticians can generally characterise the varying level of accuracy that can be achieved by these analytic algorithms, by making use of measures such as the precision and recall ratio, and the residual sum of squares. In studies which aim to assign an author to an unassigned text, or in studies which aim to date a particular text, the accuracy of a method can be verified through tests based on texts whose authors are known. This concept of validity seems appropriate in studies which operate under the scientific assumption that questions can be answered conclusively through a single answer. In the context of algorithmic criticism, however, such strict criteria for the evaluation of the adequacy of analytic methods can seldom be applied. In this specific discipline, the results of statistical procedures can be considered relevant when they ignite new ideas about the text that are studied, or when they convincingly stimulate hermeneutic interactions. A method which performs poorly according to statistical accuracy measures can still be valuable for literary criticism if it helps scholars to discover striking formal or thematic similarities between two texts. Such new ideas can be fostered when statistical methods manage to expose striking differences between different sets of texts, or when statistical methods can dispose formal similarities between texts which, according to the scholar, also share particular thematic concerns. During the case study, I experienced that many of the methods that were applied failed to inspire such new ideas. The statistical techniques often have to be used in an aleatory or experimental fashion, and it is often difficult to know beforehand whether these analyses can yield interpretable results. The choice of a particular method often reflects an individual scholar's preference.

Next to the fact that this multitude of statistical methods can complicate an effective comparison of texts, the discovery of relevant texts can likewise be challenging because of the absence of a single reliable heuristic method. In the case

study, I tried to find interesting examples of the use of semi-rhyme or deibhíde rhyme by sorting all the poems according to the frequency counts of these devices. The poems in which devices have been used most profusely do not necessarily contain the most interesting instances. The observation that a particular fragment has literary significance demands a qualitative assessment, and it continues to be challenging to detect such significant cases solely on the basis of quantitative methods. In *Modern Poetry*, Louis MacNeice wrote that “the rules or ‘laws’ of poetry are only tentative devices” and that “[t]here is no Sinaitic recipe for poetry, for the individual poem is the norm”.⁵⁴⁴ MacNeice’s explanation implies that, because of the intrinsic uniqueness of each poem, it is exceedingly difficult to develop steadfast rules for the identification of the qualities that can make a poem extraordinary. Because of its fortuitous and unpredictable nature, the discovery of all relevant texts is difficult, in traditional and in computer-based research alike. It can be expected, nevertheless, that the methods of algorithmic criticism, which can recognise occurrences of many literary devices, and which, despite specific challenges, can systematically compare the texts in a given collection, can ultimately help scholars to make such serendipitous discoveries more quickly.

In the scholarly writing about the marriage between literary studies and informatics, one crucial and fundamental question reoccurs insistently: do computational methods essentially enable scholars to perform the same set of tasks, albeit with more speed or with more precision, or can these various technologies genuinely effectuate fundamental transformations concerning the knowledge that can be produced? The case study that was conducted for this thesis has shown that digital techniques, because of their unparalleled capacity to methodically detect all occurrences or literary devices, can radically expedite and ameliorate parts of the traditional methodology of literary criticism. They can vastly magnify the range of textual aspects that can be observed within literary works. It is not immediately clear, nevertheless, that such heightened forms of perception can veritably qualify as a fundamental epistemological change. Martin Mueller has argued convincingly that epistemological innovation is often a gradual and incremental process, and that many piecemeal changes, occurring resolutely over the course of a longer period, can ultimately amount collectively to a potent “change in kind”.⁵⁴⁵ This study has assumed analogously that, although the methodological renewal that can be instigated by computation does not instantaneously result in a profoundly different form of knowledge, many cumulative methodological emendations can eventually result in a palpable transformation of the nature of literary research. Whereas the results of computer-based research rarely produce draconian disciplinary transformations directly, the capacity of digital methods to recognise patterns in the usage of literary devices can be of clear benefit to the field of literary

⁵⁴⁴ Louis MacNeice, *Modern Poetry: A Personal Essay*, p. 33.

⁵⁴⁵ Martin Mueller, “Digital Shakespeare, or towards a Literary Informatics”, p. 286.

criticism. Scholars with a genuine interest in literature ought to eagerly embrace every additional instrument which can shed new light on sinuous and multifaceted literary texts.