Stylochronometry: Stylistic Development, Sequence of Composition, and Relative Dating

Constantina Stamou
Private Researcher

Correspondence:
Constantina Stamou,
Ricability, 30 Angel Gate,
City Rd, London,
EC1V 2PT, UK.
E-mail: cstamou@gmail.com

Abstract

This article examines representative successful and unsuccessful applications of stylochronometric approaches of the last sixty years in a thematic fashion, aiming to present in a concise manner, although not exhaustive, modern approaches. Differences concerned with adopted methodologies, stylistic markers, and text size render any comparisons among the studies difficult. Nevertheless, common problems may be traced, whereas groups of different stylistic marker types of potential use for applications concerned with stylistic change in time are identifiable.

1 Introduction

In 1998, Forsyth coined the word 'stylochronometry' to define 'stylometric studies dealing with chronology attribution problems' (R. S. Forsyth, 2004, personal communication). The first recorded attempt at stylochronometry relates to Campell's (1867) work on the Platonic corpus, after which German researchers took over during the end of the 19th and the beginning of the 20th century. Summaries of their work exist in Brandwood (1990). Generally, those studies are characterized by subjective selection of stylistic markers of often low numbers and analyzed by descriptive statistics leading to generalizations of doubtful accuracy. Paired with insufficient external evidence to support any assumptions, the true chronology of Plato's undated dialogues remains a problematic case and probably will continue to remain.

However, recent, modern, and systematic research on Plato and other cases has not been hindered by such obstacles. This is due to the most fascinating aspect of stylochronometry which is the very same assumption it has been founded upon: the hypothesis that certain aspects of an author's writing style evolve rectilinearly over the course of an author's life time, hence, with appropriate methods and stylistic markers, such changes ought to be detectable (Frischer, 1991). This assumption has resulted in a number of successful studies which have primarily concentrated on the more straightforward subject of detection of stylistic development, taking for granted that general stylistic development not only exists, but is additionally detectable. Such studies operate as text classification for a limited number of age/year categories principally for the purposes of description. Notably, however, the path has not been as promising for the remaining stylochronometric branches, detection of chronological order and relative dating. Although these branches have actually taken the main assumption a step further—that is, if development does exist and is detectable based on internal textual evidence, then it ought to be traceable in a manner resulting into a correct chronological order of the ordinal kind for the documents under investigation, or even the acquisition of relative dates based on date of composition by treating date as an interval variable—their lack of universal credibility stems from the ambitious studies that have been

undertaken. That is, so far the main problems of interest have been works for which sufficient external evidence does not exist, such as Plato's dialogues, but also Shakespeare's plays, Horace's *Ars Poetica*, and Euripides' tragedies. Consequently, any conclusions drawn have been impossible to verify.

Furthermore, it appears that a number of developmental studies share similar characteristics by having been performed on known authors, either by external evidence or by literary criticism, for their identifiable stylistic trends beforehand (Jaynes, 1980; Opas, 1996; Can and Patton, 2004); therefore, it has been simple, first, to detect developmental patterns, and second to fall back on literary criticism to reinforce observations, dispute literary criticism (Jaynes, 1980), and link any findings to external evidence. However, literary criticism is based on subjective opinion and interpretation. Moreover, such studies have given birth to the implied generalization that some aspect of stylistic development ought to be universal and may be traceable in the wider authorial population, just as stylistic stability is considered universal and detectable in certain aspects. However, the tailored-to-each-author manner in which stylochronometric studies have been conducted until recently does not permit speculation on the existing universal applicable and inapplicable side of the rectilinearity hypothesis, which is reinforced by the limited number of studies reporting results highlighting stability instead of development (Jaynes, 1980; Laffal, 1995; Whissell, 1999; Smith and Kelly, 2002).

To answer questions of the kind 'is it feasible to identify and efficiently use universal stylochronometers?' and 'is there application for the rectilinearity hypothesis in the stylistics domain?', it is believed to be important to primarily investigate past studies in all three stylochronometric branches (nominal, ordinal, and interval) to obtain an overview of the stylochronometric domain. Such an investigation ought to shed light on common problems and provide insight for improved strategies.

Consequently, the purpose of this article is to offer a broad perspective of the field by the exposition of cases on which a breadth of methodologies, approaches, stylistic markers, and text sizes have

been adopted. The aim is to provide an overall account of the situations researchers face and to determine the nature of any common problems that have arisen by describing the most notable modern representative studies characterized mainly by the use of computer-based methods which have in recent years promoted the acquisition of vast amounts of machine-readable literary material, the straightforward and time-efficient application of both univariate and multivariate statistical methods, and the development of 'new concepts of textual analysis such as vocabulary distributions, word frequencies, etc.' (Temple, 1996). The second aim is to identify common variable types that have been employed across the selected stylochronometric studies to compile a list of chronometers of potentially universal use.

2 Stylistic Development over Time

Whissell (1996) notes that 'development comes into question when works of a single author or in a single category are spread across a considerable span of time, during which they might have been expected to change'. Although the studies described adopt a loose interpretation of the phrase 'considerable span of time' [ranging from the Beatles's (Whissell, 1996) eight-year career to Yeats's fifty-eight year poetic career (Jaynes, 1980)], they do successfully report stylistic development in the authorial works investigated.

The first study of interest is attributed to Jaynes (1980) who concentrated on the syntactical aspects of Yeats's poetry due to literary criticism focusing on radical systematic and unsystematic stylistic and semantic differences between Yeats's early and late work. Jaynes computer-analyzed twenty-five typical Yeatsean poems published in two year intervals—ranging from 87 to 319 words totalling 4,656 words—from eleven poetry volumes representing Yeats's poetic career from 1889 to 1939. The results were analyzed using *t*-tests, Pearson correlation coefficients between word-class and function raw scores to detect linear relationships, Spearman correlation coefficients between word-class and function ratio categories to detect nonlinear

relationships and linear trends in the raw scores, as well as time-cluster analysis. Fifteen variables[1] were selected and used with time-cluster analysis, but the obtained results, although significant changes were detected in nouns as heads, prepositions/nouns ratio, main verbs as heads, auxiliary verb heads, and prepositional phrases words, produced nonlinear patterns from which it was difficult to generalize across the entire Yeatsean sample. Consequently, Jaynes concluded that Yeats's syntactic style remained stable during his poetic career, which was contrary to critical opinion or Yeats's personal remarks.

Turning to a different set of variables derived from diction,[2] and again drawing upon critical observations on thematic shifts, Jaynes concluded that, based on t-tests, Yeats's poetry is characterized by 'an early Celtic period composed of words of moderate length and proportion of content words, followed by a more personal, informal period of longer words but fewer content words, and finally a homogeneous period of much shorter words, and with a higher percentage of content words'.

In 1996, Opas produced a multivariate and multi-dimensional analysis of Beckett's prose texts with the assistance of Biber's (1988) model.[3] Of primary interest was to detect Beckett's stylistic development over time. To test her assumptions, Opas selected eleven Beckett plays in 1,000 word blocks comprising around 37,000 words in total, approximately one work per decade, and covering the period between 1938 and 1982. Then, Opas used Biber's software to obtain standardized frequency counts of the variables required by the model and submitted them to factor analysis in order to examine the behaviour of the mean scores of the texts along the model's dimensions.

By comparing Beckett's works to the standardized genre-specific norms provided by Biber, Opas was able to identify the extent to which Beckett's texts are typical narratives rather than highly individualistic. In the first two model dimensions (involved versus informational production and narrative versus non-narrative concerns) the Beckett texts have general fiction characteristics, however, on the remaining three dimensions examined (explicit versus situation-dependent reference,

overt expression of persuasion, and abstract versus nonabstract information) the scores suggest that often some of Beckett's texts appear more explicit, more situation-dependent, and more abstract compared to readers' expectations of narrative fiction since the scores are situated beyond the fictional dimension range provided by the Biber model.[4] Consequently, Opas concluded that although Beckett's writing style evolved during his lifetime from the more typical narrative productions to his individualistic type of expression, his change has not been uniform. Each work appears to be stylistically innovative, which is confirmed by lack of consistent trends in any of the dimensions tested, and also by literary criticism in accordance with Opas's results.

Whissell (1996) assessed the extent to which Paul McCartney and John Lennon differed in their lyrics during the years 1962–70 by using emotional stylometry[5] with content analysis of the emotionality[6] of words in terms of pleasantness and activation. Based on the assumption that pleasantness, activation, and cheerfulness would have decreased over the Beatles' career in four life stages, Whissell automatically scored 155 Beatles songs with the Dictionary of Affect (Whissell et al., 1986) to obtain percentage usage of the emotional and stylometric categories comparable for the two composers. The aim was to both identify whether such a transition existed and to locate the stage at which the change occurred. Chi-square assessed differences between lyricists and career stages, the Kruskal–Wallis test of medians compared the median of word lengths across the four stages, and the Mantel-Haezel test for linear association examined the possible existence of linearity between variables across the different life stages under the assumption that any linearity identified evolved gradually, be that positive or negative.

Results indicated that during the two middle musical stages words were relatively longer than those in the remaining stages; that pleasantness, activation, and cheerfulness diminished in strength, while nastiness came in force along with softness; that sadness followed a rise-then-fall-then-rise path; and that words like 'the, a, an', and whoops, repetition, and punctuation became more common, whereas first and second pronouns, negatives, 'and', 'love',

and 'girl' were found to diminish. Whissell's results agree with critics' observations on the Beatles's lyrics according to whom Lennon is 'the less pleasant and the sadder lyricist', whereas the lyrics written by both song-writers are most noticeably negatively affected by Lennon's mood in the first and second stages of their career by being 'less pleasant, less active, and less cheerful over time' (Whissell, 1996). Finally, observation of variables' trends' percentage rates for the different stages revealed that the greatest change for pleasant words and emotionality reduction occurred between the first two stages by noting rarer instances, whereas softness and nastiness rose between stages three and four.

The most recent developmental stylochronometric study is attributed to Can and Patton (2004) whose work concentrated on detecting changes in the writing style of two Turkish authors, Çetin Altan and Yaşar Kemal. Can and Patton examined type and token length, and the frequency of most common words using ANOVA, linear regression, logistic regression, principal components analysis, and stepwise discriminant analysis. They also used sixteen fixed-size blocks of 2,500 words each for each old and new period reaching a total of 80,000 words of journalistic columns by Altan published during 1960–69 and 2000, and similar size blocks by Kemal obtained from his 1971 and 1998 novels in their entirety. Their aim was to detect a marked positive change between the sampled early and late works as expressed in terms of the stylistic markers.

ANOVA showed that for each author the average token and type length were significantly longer for the new works. Linear regression on Altan's data for token lengths and text age was successful, although not highly impressive ($R^2 = 0.2415$) suggesting that other factors must also be involved. Logistic regression on the usage rates of types and tokens, with binary dependent variable values 'new' and 'old' and frequencies of tokens or types of a particular length as the independent variable, showed that, for both tokens and types, word length of three to eight was predominately occurring in the old works, whereas word length of nine or greater was characteristic of the new works with significant results for words greater than ten characters. For Kemal, word length between three and six for tokens, five or less for

types, and sixteen for both types and tokens was detected as a characteristic of the old work, whereas types of length between three and twelve and tokens of size two, ten, and twelve were abundant in the new work. It was concluded that such change may be attributed to the authors' progressive mastery of language hence the use of longer words which the agglutinative nature of Turkish allows.

Comparison of the frequencies of the fourteen most common words for Altan with logistic regression yielded five words as significant out of which two (*bu, gibi*) were more frequent in the old texts and the remaining three (*çok, da, de*) in the new texts. From the fifteen most common words used for Kemal, two (*dedi, gibi*) were found significantly more often in the old blocks and two (*daha, kadar*) in the new. Principal components analysis with all the variables in was conducted to provide graphical illustrations for the separation between the works. Finally, stepwise discriminant analysis with cross-validation chose the most efficient chronometers. The different discriminant analyses run for Altan produced the best token length selected ranging between four and eighteen characters, the best type length between three and fifteen characters, and the best common words among the selected fourteen which were *da, de, çok, bu, için,* and *sonra*. Average success classification rate was 98.96%. Kemal's discriminant analyses produced type length frequencies ranging from two to eighteen, character-length of four for tokens, and the words *dedi, kadar, da, sonra, ne,* and *daha* among the most common fifteen. Kemal's average success classification rate was 84.38%. The difference in the two classification rates was thought due to the greater time distance between Altan's works, hence allowing for more opportunity for his writing style to evolve.

3 Order of Composition

Studies on developmental work have presented evidence in favour of detectable and identifiable developmental stylistic trends. Relying on such findings then, order of composition ought to additionally be measurable. Admittedly, though, many more parameters need to be taken into account in such cases, and different versions of developmental

order are in theory plausible based on the accuracy of the available data and results. Theoretically, literary criticism and historical evidence can reinforce or refute claims of chronological ordering. However, at least in particular cases such as the Classics, and Shakespeare in which historical evidence does not suffice, it has still not been possible to reach definite conclusions.

3.1 Plato

Known as the author of the *Republic* and the *Apology*, Plato left a stylistic and thematically variable work which was presumably composed over a long period of time allowing for the development of his thought and philosophical theory. Problems with authenticity relate to *Epistle* 7 and other dialogues due to comparisons and disagreement on what the content of his philosophical theory should be, however, the problem with the ordering of the Platonic corpus has been active since antiquity. Scholars relying on biographical and historical material have been able to identify three stylistic periods (early, middle, late) in Plato's writings to which dialogues have been categorized accordingly. Unfortunately, however, the order of the dialogues within each stylistic period has not been possible to ascertain, especially the late dialogues which show unexpected variability in terms of style (Ledger, 1989). In particular, although scholars generally accept the *Laws* as the last Platonic work, preceded by the *Timaeus*, the *Critias*, and the *Republic*, 'before them in reverse order of composition, according to the majority view, came the *Philebus*, *Phaedo*, *Symposium*, *Politicus*, *Sophist*, *Parmenides*, and *Theaetetus*', while there is no agreement concerning the remaining works (Brandwood, 1990). Furthermore, as Brandwood (1990) mentions, 'the greater consensus which exists today about the chronological sequence itself considerably different from that just mentioned can rightly be claimed as the achievement of the stylistic method'. Speculations center around Plato's ability later in life to use a plethora of styles or to editorial and secretarial influences on the late dialogues. Theslef (1967) attributes any possible failure to obtain an accurate order for Plato's dialogues to lack of a linear development in Plato's corpus in the first place since

his diction may well had been altered deliberately from dialogue to dialogue to suit Plato's purposes. Therefore, the order that would have to be detected may as well obey a different type of curve.

Cox and Brandwood (1959) investigated the distribution of clausulae in the entirety of Plato's late dialogues by measuring only the last five syllables of each sentence in the texts and categorizing the syllables either as long or short, based on the assumption that Plato's late style changed rectilinearly. The frequency distributions of the resulting thirty-two qualitatively different sentence-ending types revealed a 'marked difference' between the percentages for the *Republic* and for the *Laws*. The aim of the study was to obtain a decreasing order of composition for the remaining plays with affinity to the *Republic*, based on the assumption that Plato's writing style changed linearly with time. Linear discriminant analysis was the method used to score each different type of sentence ending which were subsequently assessed for significance. A negative mean score was assigned to the *Republic* and a positive mean score to the *Laws*. It was anticipated that the remaining plays' scores would provide a suggestive ordering, which was eventually the following: *Republic*, *Timaeus*, *Critias*, *Politicus*, *Philebus*, and *Laws*. *Critias* was the only play of ambiguous dating since based on the statistical evidence it could have also been placed anywhere before *Timaeus* or anywhere before *Politicus*. The researchers acknowledged that such an ordering would not be in complete accordance with views on Platonic chronology at the time, although there was a minority group sharing similar positions.

Wishart and Leach (1970) extended Cox and Brandwood's work by investigating the behaviour of percentages of five-syllable sequences as in 1–5, 2–6, 3–7, etc. up to *N*-4 throughout the text instead of at the end of a sentence, using cluster analysis, principal components analysis, and multidimensional scaling. The data in thirty-three samples ranging between 1880 and 3667 numbers of syllable groups which had been derived from the ten Platonic books *Timaeus*, *Sophistes*, *Philebus*, *Critias*, *Laws*, *Seventh Epistle*, *Republic*, *Phaedrus*, *Politicus*, and *Symposium*, were first considered individually and then concatenated into the original ten books to

obtain accurate representations. All the books, except *Phaedrus*, *Symposium*, and the *Seventh Epistle*, represent the late group used by Cox and Brandwood (1959). Moreover, only continuous prose was used, unlike Cox and Brandwood. Euclidean distance measured resemblance between samples. In reality, the researchers concentrated on computing the sum of the squared error for each pair of samples based on the selected variables and used the values to form clusters by adding cases to a group-sample in succession whose Euclidean distance from the group average was minimum.

It was noted that the obtained chronological order for the late group of writings was the same for the one produced by Ward's method, Cox and Brandwood, and multidimensional scaling, since for Plato to maintain his theory of rhythm and use of clausulae, experimentation, development, and adaptation of rhythmic variation was expected to be reflected throughout his prose. The suggested seriation was *Phaedrus-Republic-Symposium*, *Timaeus*, *Sophistes*, *Critias* (*Seventh Epistle*, *Politicus*), *Philebus*, and *Laws*, with the *Seventh Epistle* and *Politicus* being the only dialogues deviating from the ordering, and the *Republic* and the *Symposium* were noted to occupy the same position although not in any specific order. The *Phaedrus* data did not produce any evidence of a homogeneous late book, but rather placed it 'controversially' (Robinson, 1992) directly opposite to the *Laws* which is universally considered to be the latest (Wishart and Leach, 1970). However, Keyser (1992) expressed doubts about the obtained results since considerable variation remained unaccounted for which placed great demands on the data for production of the principal components and multidimensional scaling plots. Moreover, Brandwood (1990) has also criticized their work by noticing that the five samples used from *Phaedrus* were taken from speech parts, thus biasing the results.

Ledger (1989) stylometrically analyzed, with the help of the Greek alphabet,[7] carefully-selected 1,000 word blocks (total 702) from all of Plato's works to obtain the order of composition of Plato's dialogues. The methods employed were canonical discriminant analysis and stepwise discriminant analysis. Ledger constructed four matrices which

he populated with the dialogues based on cross-referencing from within the dialogues themselves. The categories consisted of overlapping sets of dialogues, ranging from having one dialogue in each early and late category, to having thirteen in the early one including the *Republic* and four to the late one including the *Laws*. Stepwise discriminant analysis was used in order to select a subset of maximum six variables able to discriminate between the groups of early and late dialogues and that subset was subsequently used with canonical discriminant analysis to obtain a unique combination of the original variables that would maximally correlate with group membership and order each work according to canonical variate's value. In particular, canonical discriminant analysis was first used with the first ten variables, then, with the best nine and so on until the best three are used. After, the procedure was run with all thirty-seven variables in the model with each variable set (ALETS, BLETS, CLETS) separately. Mahalanobis distance was employed to examine distance between different dialogues.

Ledger agrees with current scholarship in the existence of two periods of early and late dialogues. Within those periods he places the *Seventh Epistle* and *Epinomis* with affinity to the *Laws*, and *Hippias Major* and *Alcibiades* 1 with affinity to the *Republic*, whereas the order obtained for the later works was *Sophist-Politicus*, *Laws-Epinomis*, and *Timaeus-Critias*. Furthermore, in this order, *Philebus* is controversially placed very early, and *Timaeus* appears very late. The placing of the last pair of dialogues, however, is in disagreement with that of Cox and Brandwood (1959) and Wishart and Leach (1970) who place it at the beginning of the late period. Problems related to Ledger's method relate to disputed authenticity of some of the plays.

The last stylochronometric attempt on Plato is attributed to Temple (1996) who examined thirty-eight linguistic, thirty-two clausular, seventeen reply, hiatus, and thirty-seven orthographic variables from past research in twenty-five platonic dialogues with the help of principal components analysis and MST (Minimum Spanning Tree)[8] as an attempt to rework the 'old' stylometry. Temple first aimed at dimensionality reduction with principal components analysis on the general linguistic,

clausular, and reply data, while subsequently he used a total of six variables selected from the hiatus and orthographic[9] variable types, as well as the first and second principal components for the linguistic and clausular data obtained from his first experiment in a final principal components analysis in order to work with reduced noise.

Separate principal components analyses on the linguistic and clausular variables indicated similarities in the dispersion of points, that is, early and middle dialogues appear to be easily distinguishable from the late dialogues, although the late dialogues were characterized by greater variability. They also seemed to be associated with an increase in frequency of linguistic variables, superlatives, rhetorical questions, the words 'orthos/orthotata', reduction of the words 'pane je' and clausular patterns. A similar trend was observed for the reply variables.

In the second phase, the ten selected variables produced relatively similar results to those obtained from the analysis on the linguistic variables. The early and middle dialogues revealed a clear separation from the late ones which were found a lot more dispersed compared with the tight gathering of both early and middle dialogues, with the *Parmenides* and the *Theaetetus* being placed to the middle period, whereas the *Sophist*, *Timaeus*, and *Critias* to the beginning of the late one. The discrimination was based on a decrease in hiatus, changes in orthography and the first clausular and linguistic principal components from the first analysis, and was justified either as a nonmonotonical stylistic change in Plato's writing or influences from secretaries or editors who affected the dialogues. Furthermore, the *Meno*, contrary to popular theory who places it later than the *Gorgias* as a transitional work, was in this case been placed early; also, the *Euthyphro* was placed relatively late although generally it is believed to be early; finally, the first book of the Republic group was found clearly to precede the remaining ones. On the whole, however, subdivision of early and late dialogues was observed to agree with current views.

3.2 Other studies

In 1982, by attempting to obtain the order of composition of Marlowe's plays, Ule used nine[10] variable types on the assumption that works composed sequentially entertain a particular affinity related to the variables measured. Affinity was measured with Euclidean distance[11] between pairs of distribution curves, whereas RVO was used for vocabulary overlap. To obtain possible sequences of composition, Ule computationally generated all possible sequences of the seven plays. Ule observed that all the obtained sequences agreed in placing the *Tamburlaine* first, whereas in three sequences, based on word length, connectives, and RVO, the second part of the play accurately followed the first. However, as Ule (1982) notes, such a procedure will not be effective in more populated corpora due to restricted required machine time and resources available in one's lifetime.

Burrows (1987), using most common word frequencies without distinction in grammatical or lexical function, developed a method which identifies similarities and differences across a range of texts and text types by obtaining frequency profiles and correlating them with each other in descending order of frequency in order to map them using Eigen analysis. By comparison of several different parameters, Burrows identified authorial and chronological determinants. His method was tested on Jane Austen's published narrative using thirty most common words.

Initially, similar frequency patterns were obtained across different types of narratives (pure, character, dialogue) in the Austen texts. When eighteen words were tested after removal of those ones which changed form in the narratives, Pearson correlation coefficients indicated consistency in the frequency profiles. Subsequently, Eigen analysis graphs revealed among other things that each narrative corresponded almost perfectly to its assigned date of composition. The results remained of similar quality when all thirty words were used, and again when major homographic forms had been clearly defined and incorporated in the hierarchy of the frequency profiles.

Burrows also converted pure narrative into 2,000 word segments and obtained word frequencies in order to compare them with each other or with the corresponding segments from a different type of narrative, partly concentrating on the demonstration of differences between Austen's early and latest

narratives. Burrows noted the ability of the method to successfully distinguish between an author's different writing stages, simultaneously mentioning that the correct chronological sequence was disrupted only when Austen's last novels and the fragmented manuscript 'Sanditon', completed by an anonymous author after Austen's death, entered the data set. Moreover, it was observed that although the original 'Sanditon' appeared close to the early parts of *Persuasion*, the imitation part overlapped only slightly which could not be nevertheless 'mistaken for the original'.

Arguably, Horace's *Ars Poetica* has been deemed his hardest poem to date due to variation in the proposed assigned chronology and to lack of appropriate historical information to accurately date the remaining of his output as a reference point. Consequently, only relative chronology, although well established based on external evidence, is available. So far, various attempts at the task had relied either on placing the poem according to Horace's estimated developmental pattern of poetic ideas, or according to the life span of the historical figures mentioned in it, or on stylometric material of diction and meter (Duckworth, 1965). In 1991, after refuting Duckworth's results and showing that no connection exists between meter and chronology, Frischer (1991) set out to date the poem aiming at linking it to *Epistles* I instead of *Epistles* II1 as had been traditionally the case up to that date.

Accepting the relative chronology of Horace's canon, and after controlling for genre and meter by separating the poems in lyrical and hexameters, Frischer performed three experiments. In his first experiment, poems were treated as nominal variables to establish the assumption that poems found in clusters (early, early-middle, middle, middle-late, and late) did so because they had been composed during the same time period. Using the chi-square test to investigate the chronological association between four selected function words[12] and the poems[13], Frischer developed models by combining *Ars Poetica* with any other poem according to genre and time restrictions and by assuming closeness of composition based on time periods. The best model incorporated *Satires* II, *Epistles* I, and *Epistles* II.2, whereas models which combined *Ars Poetica* with

Epistles II.2 were also satisfactory.[14] Because *Epistles* II.2 was treated as a late hexameter poem, this placed the *Ars Poetica* in the middle period (Frischer, 1991).

In his second experiment, Frischer investigated 'macrochronometers' to detect universal stylistic development and strengthen the results of the first experiment. Mean word-length was found steady, whereas mean sentence-length in words as per strong stop fluctuated randomly; therefore, they were omitted from the analysis. Unique strings as a percentage of all strings showed that the variables poem and string-type (unique versus nonunique) are not independent, while graphical representation revealed a positive linear pattern. However, no reduced model performed better than the original full model, which suggested that there is 'no basis at this point in the analysis for associating the *Ars Poetica* with one of the datable hexameter poems' (Frischer, 1991).

In the final experiment, the hexameter and lyric poems were separately treated as interval variables by being assigned their relative year of composition in order to assign a precise date to *Ars Poetica*. Using only the ratio of unique strings, an R^2 derived from linear regression and equal to 99.1% was obtained for the lyrical poems, although the number of cases in the sample was only three. In the hexameter poems, R^2 was 69.6% for a total number of five cases, and the obtained graph revealed a more scattered picture. The hexameters' formula placed *Ars Poetica* around 20/21 B.C. Then, Frischer used the four function words with the *Ars Poetica* only in order to obtain most and least probable dates by running the regression method seventeen times with a variable *Ars-Poetica* date from 24 to 8 B.C., controlling simultaneously for the remaining poems. Frischer wanted to obtain the most and least plausible years of composition for the poem based on R^2, to circumvent the small sample sizes, and to 'increase the plausibility, if not probability, of [the] results by including the *Ars Poetica* in the analysis' (Frischer, 1991). The results showed a superior R^2 from the three function words *sed*, *per*, and *nec*, compared to the one from the unique strings, while *ad* produces a similar value. The ratio of the best to the worst R^2, 'which quantifies the

advantage of the best model over the worse', reveals that 'sed and per are very reliable chronometers indeed, both in absolute and relative terms' and that the best suggestive date is between 24 and 20 B.C. (Frischer, 1991).

4 Relative Date

Studies concentrating purely on relative dating are not as simple or as straightforward to conduct, even more so compared to studies dealing solely with order of composition since much more precision is needed in the obtained results; therefore, merely correct chronological order in this case does not suffice.

4.1 Euripides

Only a handful of Euripides's tragedies, out of almost ninety plays, have been securely dated based on external evidence. Part of the remaining ones have been approximately dated according to internal evidence derived from meter. Eighteen of his plays are extant, while fragments of some of the plays, often substantial, survive.

In 1981, Devine and Stephens investigated occurrences of appositives at Porson's bridge[15] in the trimeter parts of Euripides' plays due to observations placing higher concentrations of the stylistic marker in the later plays. The aim was to assign an accurate chronology to the texts and introduce evidence of the strong correlation between the appositives at Porson's bridge and resolution instances on which most work up to that date had concentrated. The selected measurements were rates per thousand trimeters and standard deviation of appositives at Porson's bridge, and rates per thousand trimeters and tetrameters of resolution with their respective confidence intervals which were used to establish a rank order of the plays.

To provide dates to the undated plays, linear regression was used on the well-dated texts Medea, Hippolytus, Troades, Helen, Orestes, and Iphigeneia at Aulis, with the exception of the Alcestis and the Bacchae because of their systematic deviation from the true chronology during the ranking phase. It was noted that the rate of appositives at Porson's bridge

resulted in 91% variance accounted for, that the same percentage of correct prediction results with the rate of resolution in the trimeter sections, and that 97% variance accounted for can be achieved by the rate of resolution in the tetrameter sections. Following, the formulae were used to obtain predicted dates for the disputed plays. It was observed that Heracleidae was dated later than Hippolytus; Hecuba was dated late; Supplices was dated at 423 B.C. although such a date disagrees with that assigned to Andromache with which it ought to be in the same group; appositives dated the Electra at 414 B.C.; Helen was dated the same by all criteria; tetrameters are the only ones to correctly date the Phoenissae later than the Helen; and on the whole the resulting model dates were very close to the actual known chronology (Devine and Stephens, 1981).

A number of minor resolution criteria were further examined to acquire additional independent but not duplicate information regarding the dating since all resolution types appear to increase over time. Besides significant variation from the restricted number of minor resolution occurrences, the overall obtained ranking agreement was found strong and statistically significant. Moreover, when the minor resolution ranking was correlated with the ranking of the dated plays obtained by the regression analysis, a significant correlation coefficient was obtained. When the actual dating was examined, the Phoenissae was wrongly dated before the Helen, and the Bacchae before the Orestes, whereas the Alcestis was placed after the Ion and the Andromache just before the Helen (Devine and Stephens, 1981).

Finally, the rate of resolution and appositives at Porson's bridge were predicted by the dates of the plays in order to examine the growth of the chronometers. A satisfactory goodness-of-fit was obtained, although the placing of the Bacchae and the Alcestis was characterized as 'anomalous', the justification being difference in genre since Alcestis is a pro-satyr play, whereas the Bacchae was probably affected by stylistic aspects of its language.

Cropp and Fick (1985) concentrated on occurrences of resolution in iambic trimeters of the dialogue scenes in Euripides' tragedies and their relation to the tragedies' chronology. Incidences of

'low resolution-rates and a restricted variety of resolution-types in the early dated extant tragedies—*Alcestis*, *Medea, Hippolytus*—and of increasingly higher rates and an increasingly wider variety of types in the later ones—*Troades, Helen, Phoenissae, Orestes, Bacchae, Iphigeneia at Aulis*' had been observed, interpreted as reflecting 'on the whole, a development over time in this feature of his style, rather than some chance combination of special motivations relevant to each individual play' (Cropp and Fick, 1985). Drawing from such observations, and from the assumption that plays were expected to show variation in their relative rates of resolution in accordance to a developmental time-line, regression analysis was utilized on data both from the fragmented and the fully-dated tragedies aiming to apply the obtained model on the undated extant and undated fragmentary plays. Therefore, data associated with particular years formed subsets for which each resolution rate was considered the best estimate of the year-subset. Noticeably, the later the year, the higher the resolution rate tended to be, although such a straightforward increase was not observed until the year 431 to 406.[16] Relative likelihood intervals were also established for the variable of interest.

It was finally observed that, for the eleven undated plays, *Danae, Kretes*, and *Protesilaos* were associated with the early group, whereas *Temenidai* (with plausible date 422>), *Polyidos* (421>), *Antigone* (420>), *Oidipus* (419>), *Meleagros* (418>), and *Auge* (414>) were associated with the late group. *Phaethon* was located between 427 and 414, while *Alkmene* was associated with the 420–410 decade. In addition, conflict between the resolution evidence and external evidence resulted in doubts about the established dates of *Erechtheus* and *Antiope*. On the whole, it was concluded that 'absence of conflict between the metrical and the nonmetrical evidence which tends to reinforce confidence in the validity of the metrical evidence as a whole, and in correcting if only with negative effect in some cases many of the inferences made in the past from the evidence of resolutions' (Cropp and Fick, 1985).

4.2 Shakespeare

The problem of dating Shakespeare's plays stems from the fact that 'Shakespeare did not personally oversee the publication of a collected edition of his words, and no autograph manuscripts survive of words attributed to him in his lifetime. The contents and chronology of his canon—. . . therefore remain, and will for ever remain, a matter of dispute' (Wells *et al.*, 1987). In particular, disputes are currently centering around three major areas: the relative order of the early comedies, the link between literary composition in varying genres, and the actual date Shakespeare began composing his plays. His poetry is well established chronologically though in his early work, therefore, it is of interest to detect possible links between the vocabulary of the early plays with that of the poems (Wells *et al.*, 1987). Jackson (2002) summarizes Shakespeare's poetic development in terms of meter as being away from a succession of strictly iambic, decasyllabic, end-stopped lines towards greater freedom, variety, and flow. The use of enjambment increased, the sense being run on, with no syntactical break, from the end of one line to the beginning of the next. More alexandrines appeared, and the verse began to carry a larger proportion of extra unstressed syllables, particularly those constituting feminine or double endings to a line, the final iambic foot being followed by an unstressed syllable. As end-of-line pauses became less frequent, pauses within the line became more frequent and heavier, and were more variously placed. Pentameters shared by two or more speakers became increasingly common.

To determine and assess the predictive accuracy of a number of lexical, metrical, and structural variables, Brainerd (1980) examined correlations with date of composition of Shakespearean plays with the intention of employing the best discriminator variables on the plays of disputed dates. The central aim was to observe, or not, deviations from the disputed dates given to several Shakespearean plays (Evans *et al.*, 1974). The selected variables included average verse line length in words, percentage of split lines,[17] percentage of contractions involving pronominal forms, and relative frequencies per hundred of 120 most frequent—or words in connection to high-frequency—lemmata on which no distinction of grammatical categories was made.

Brainerd initially identified the most sensitive discriminators by visual inspection of scattergrams

and bivariate correlations with calendar year for the thirty-eight plays in his sample, and selected those of an absolute correlation of 0.400 and above. Twenty lemmata were accepted as suitable, then, each one was separately regressed on a transformation of calendar year. Then, Brainerd selected the ten[18] best discriminators out of the final twenty when tested on the nineteen accurately dated plays. The ten discriminators were combined into a multiple regression equation with cross-validation, but detected instability in the regression coefficients led to use of principal components analysis and a more stable predictor. In phase three, regression on principal components was used on the dubiously dated nineteen plays for which the *t*-test determined deviations of the predicted dates from the observed ones for the plays *Love's Labour Lost*, *All's Well that Ends Well*, *King John*, *Richard III*, *Henry VIII*, *Hamlet*, and *The Two Noble Kinsmen*. Brainerd attributed the deviations to mixed authorship for *The Two Noble Kinsmen*, and possible revisions made at later stages for the remaining plays.

Derks (1994) used the Regressive Imagery Dictionary (RID) (Martindale, 1990) on thirty-seven Shakespeare's plays excluding *Two Noble Kinsmen* and to different Shakespearean play types (histories, comedies, tragedies) to measure incongruous juxtaposition, primordial content, and conceptual content as functions of time.[19] Derks observed a very weak linear increase in incongruous juxtaposition which was reinforced when the variable was transformed into its quadratic value. Primordial content was expected to decrease with age reflecting preferences away from the primitive and arousing that it represents, however, although the trend was identified according to expectations, it was not significant. Conceptual content was expected to remain stable or even decline with age in a typical individual, but not in Shakespeare's case, in which a high positive correlation value was detected. When individual play types were examined, incongruous juxtaposition was observed to decline with date for histories, primordial content dropped for tragedies and histories, and conceptual content increased with tragedies and histories. Overall, a tendency to alter content with time was detectable in the Shakespearean plays.

Jackson (2002), using Oras's (1960) pause data, investigated Shakespeare's stylistic development. The data were based on inclusive counts of all pause patterns (which also included heavy pauses and line-split pauses) in all of the plays in the First Folio of 1623, as well on *Pericles* and *The Two Noble Kinsmen* for which Oras had provided different counts on the different authorial parts of the plays. A matrix of Pearson product moment correlation coefficient was produced for each play or part of play when correlated with each other. Bivariate correlations were interpreted as a measure of similarity between plays based on pause patterns and written during time frames of close proximity, hence representing Shakespeare's stylistic development. It was observed that plays which tended to be highly correlated (approximately $r = 0.99$) belong to roughly the same chronological date according to the dating provided by the Oxford Shakespeare's chronology and in Wells *et al.* (1987), whereas plays which were not in chronological proximity, reached correlation values approximately as low as 0.38 (the majority of the correlations were smaller than 0.70). Jackson used his results to stress that the correlations provided by his study 'provide independent testimony to the general rightness of the Oxford chronology and the extraordinary consistency of Shakespeare's metrical development', although he noted that a number of plays did not clearly support the Oxford chronology. However, it remained unclear whether the data provided by Oras were normalized for differences in size and whether they were appropriate for analysis with bivariate correlations.

4.3 Other studies

Craik and Kaferly (1987) investigated vowel and consonant usage in Sophocles' *Trachiniae* in order to confirm scholarly assumptions which place the play at an early date. This was achieved by the development of a software program which detects and describes in detail Greek phonological information based on a standard transliteration table. All Sophocles' seven extant plays (280,413 letters) were tested on vowel-to-consonant ratio (VCR), consonant group representation, and consonant group position as in initial, medial, or final, and consonant group alliteration in trimeter sections. A VCR in

favour of consonants was observed accompanied by a slight increase in vowel usage over the course of Sophocles' career. In addition, a trend for an increase in initial consonants in the trimeter sections was detected, suggesting that if the trend is genuine then the *Trachiniae* should belong to the early period.

Furthermore, principal components analysis was run in order to linearly combine the measurements and observe their behaviour for the seven plays on a two-dimensional graph. Accounting for 66.2% of the variance, principal components analysis revealed a strong association between dentals and sibilants depicting a correlation between the measurements and the second principal component. This suggested that the first principal component represented the 'maximum variation and the contrast among the remaining consonant groups' (Craik and Kaferly, 1987). Moreover, a striking separation was noticed along the principal components between *Philoctetes* and *Oedipus Cycle* which are generally attributed to the late period indicating that since the remaining plays do not show signs of such extreme variation, possible stylistic changes have taken place over time. For this purpose, the researchers suggested the order of the *Trachiniae, Antigone, Ajax* (late 440s), *Oedipus Tyrannus* (early 420s), *Electra* (410s), *Philoctetes* (409), and *Oedipus Cycle* (401), since the *Trachiniae* appears to be maximally separated from *Philoctetes* and *Oedipus Cycle*; based on the principal components graph, it was assumed that the play may even be considered earlier from *Antigone* and *Ajax*. However, more tests are required for definite conclusions (Craik and Kaferly, 1987).

In 1999, Forsyth produced a study which concentrated on Yeatsean chronology. Forsyth utilized a quasi-random Monte Carlo search algorithm to identify word substrings inclusive of punctuation marks in 142 Yeatsean poems of rather small length (1,000 character chunks approximately), which he thereafter assessed for distinctiveness in four separate experiments. In the first experiment, ten poems were assigned to their correct chronological period, which were not included in the initial training sample. The study begun with a random search for the identification of suitable short substrings of random length between one and eight in the training data set which were ranked based on

chi-square according to distinctiveness. The texts were divided into groups of 'young' and 'old' of relatively equal size using the year 1915 as cut-off date. Eleven 'young' and nine 'old' substrings were the final ones selected for convenience. When five more poems were added to each category, in nine out of ten of them, the count of the substrings was higher in the appropriate category (Forsyth, 1999).

In the second experiment, two poems revised approximately thirty years apart were examined for inherent differences; it was observed that, during revision, the number of old markers was increasing and replacing the new markers. In the third experiment, a linear regression formula was obtained by the development of a Yeatsean index[20] used as the independent variable against year of composition. With the addition of three more poems to balance out the middle period, a highly significant R^2 of 0.71 was obtained. In the final experiment two short prose extracts from two essays (450 words approximately each) composed forty-six years apart were examined based on the same substrings in order to assess the appropriateness of the method on a different genre. The extracts were correctly classified to their chronological group based on observation, with the exception of only two out of the twenty substrings indicating otherwise (Forsyth, 1999).

In the same year, Forsyth *et al.* (1999) investigated the authenticity of the *Consolatio* (1583) with the secondary objective of testing methods previously used only on English texts on the language of Latin. However, as part of their analysis,[21] in order to accurately place the *Consolatio* in its appropriate time frame and compare the language of the text with the most suitable controls from the correct chronological period, the research crossed into the field of stylochronometry.

Using literary works by eleven authors—among them Sigonio and Cicero themselves—amassed in seventy samples of 305,000 total which covered the classical and neo-Latin periods, the researchers employed Burrow's (1992)[22] approach for fifty most common words with no lemmatization excluding content ones. Moreover, word length in syllables and syllable transition frequencies were automatically obtained. Because syllable counts were unsuccessful in positively affecting the analysis when

Cicero was compared with the classical authors and Sigonio with the renaissance ones, it was decided to test the variable across periods. Stepwise discriminant analysis without cross-validation was conducted on all texts based on the dependent classical and neo-Latin groups in order to classify the *Consolatio* under one of those chronological categories. Three function words (*ac*, *vel*, *vero*) and four syllabic transitions (ST13, ST22, ST43, ST44) proved significant with 98.53% correct classification rate except for one misclassified case in the neo-Latin group, the most modern one of the classical texts (*Tacitus*) (Forsyth *et al.*, 1999).

Furthermore, stepwise multiple linear regression was employed to obtain the *Consolatio*'s date by using sixty-six of the securely dated texts, the century of composition as the dependent variable, the forty-six most common words, and the twenty-two syllabic variables. The final five (ST44, *ac*, *vel*, S6, ST13) selected with positive coefficients suggest that, for the lexical variables and the six-syllable words, they are characteristic of a more evolved Latin writing style, although syllabic transitions did not provide a clear interpretation, besides the possible explanation that it is a product of a universal move from the classical Latin to Neo-Latin. The final formula was tested on the *Consolatio* which was eventually placed in the later period. It would appear that syllabic information, especially syllabic transitions are useful for 'temporal discrimination' (Forsyth *et al.*, 1999).

To address the issue of inequality in text lengths when measures of lexical richness are being investigated within and across corpora, and to consider whether vocabulary richness remains constant within an author's work across his life time, Smith and Kelly (2002) proposed an experimental method which relied on averaging measures of lexical richness across different standardized nonoverlapping blocks of the main text(s) under investigation. The measures they explored were average values of Yule's constant K (Yule, 1994), Zipf's Z parameter (Orlov, 1983), average portion of hapax legomena, average repetition of common words, and the length of the classical text they were investigating in iambic meters. The underlying assumption was that the average measures do not depend on text length; therefore, works may be compared in relation to each other. The researchers tested the early and middle comedies of Aristophanes (eleven plays), the middle tragedies of Euripides (fourteen plays), and the old comedies of Terence.

Initially, ANOVA (or Kruskal–Wallis rank sum when appropriate) was used to test constancy, with $N = 300$, of the average number of unique words per chunk of N, the average number of hapax legomena per chunk of N, and the average K per chunk of N. Then, weighted linear regression was employed to detect any association between the measures and the date of the plays' first performance. Prediction intervals were also estimated to calculate performance dates of plays not tested in the training sample. The results for Aristophanes revealed that the average number of hapaxes and average K were significantly different over time with hapaxes decreasing in frequency, while average K was found to increase but not significantly. After consideration of each play in its entirety, it was discovered that the total number of words in iambics and constant K also increased significantly over time, although constant K showed sufficient variability. For Terence's comedies, only average vocabulary showed evidence of change although not statistically significant. However, its significant regression formula revealed a positive pattern. When the comedies were inspected in their entirety, no measure displayed significant changes. Finally, only the average number of hapaxes showed significant trends on Euripides' middle tragedies, and with the exclusion of the outlier *Supplices*, hapaxes were found to significantly increase with time. Finally, the researchers concluded that the method followed 'reveals the possibility of significant variance of these measures of vocabulary among works of a single author and warns against the notion of some absolute authorial stylistic character' (Smith and Kelly, 2002).

5 Content Analysis in Stylochronometry

A number of studies approaching stylochronometric problems using content analysis, thus expanding the tools available to the interested researcher, have additionally been identified in the literature.

Laffal (1997) used a general content dictionary of 168 concepts to obtain frequency profiles for a number of word categories on twenty-one collections of tales and poetry by Edgar Allan Poe totalling 191,165 words and covering the whole of his literary career by two-year intervals in order to trace a developmental time line in Poe's literary output. Obtaining z-score frequency profiles of the distribution of the word categories and incorporating them into a factor-analysis, allowed Laffal to associate each text with dominant categories on each positive and negative side of the extracted factors. The texts resulting in the highest loadings for each factor were considered as characteristic of the factor, while concepts associated with those texts were selected based on significant z-scores. The time-oriented factor was characterized by early concepts of place, power, person values, nature, and body, middle concepts like place, qualities, and ideation, and late ones of union/separation, ideation, and time. Death, which is the central theme in Poe's work, remained constant. The result was a clustering of early and late texts at opposite ends of the third and fourth factors, reinforcing the idea of a detectable developmental life line in Poe's work although not in an absolute manner since the position of 'The Raven' provided evidence of a 'continuation of earlier modes . . . among the later texts. Equally, one might expect that intimations of later ideas would occur in earlier work' (Laffal, 1997).

Subsequently, Laffal associated events in Poe's life with the detected shifts, such as Poe's brother's and wife's death (in 1831 and 1847, respectively). It seems that the death of Poe's brother caused a thematic shift from nature to union/separation concepts resulting in higher concentration of quality and utilitarian words, and fewer occurrences of words of power, ideation, and person values. The second thematic shift was observed around the time of Poe's wife's death. Although the idea of death was always of major concern to Poe, following another major experience of loss, death attains a more abstract value, seen as a ' ''mediator of union and separation'' in a more philosophic [and] kabbalistic overview of the creation and destiny of man and the universe' (Laffal, 1997).

Whissell (1999) examined Robert Frost's poetry to evaluate, among other things, imagery patterns of variability in time. Her aim was to compare results with related literary criticism which view Frost's later poetry as more complex than earlier productions. Measuring increase in complexity reflected in increase of rare words and reduction of most common ones, Whissell found 'no significant correlation between the order of appearance of Frost's poems and their linguistic complexity', whereas the periods of early, middle, late were overall indistinguishable from each other. However, based on examination of variance, Whissell recorded a significant increase in linguistic complexity within individual poems of the later period, which led to the conclusion that 'Frost's later poetry was occasionally simpler and occasionally more complex than his earlier work' (Whissell, 1999). Furthermore, when mental imagery was examined to assess the degree of concreteness/abstraction in the poems, significantly higher frequencies of abstraction were identified in the later poetry based on a Kruskal–Wallis nonparametric analysis of variance, although greater variability from the early poems to later ones was also observed. Such findings lead to the conclusion that Frost's poetry was characterized by the tendency towards abstract language and greater variability in terms of both abstraction and complexity as the years went by, rather than been driven by definitive patterns. Finally, emotional restraint was observed as stable.

Lastly, Pennebaker and Stone (2003) conducted a psychology-oriented study in which fourteen language dimensions[23] were examined in the works of four novelists, two playwrights, and four poets[24] covering over nine million words and spanning over 500 years to minimize any cohort effects in order to examine any relationships between aging and language use. Bivariate correlations and forward entry logistic regression analysis were utilized. The underlying assumption was that 'aging can affect word usage' (Pennebaker and Stone, 2003). Weak but significant bivariate correlations revealed similar patterns of correlations across all language dimensions for all authors except Dickens, possibly attributed to the limited amount of his available data. Furthermore, forward entry logistic regression

was used with the standardized language dimensions to obtain predicted logits to correlate with each author's age. Larger correlations were interpreted as showing 'greater usage of language in an age-typed manner' (Pennebaker and Stone, 2003). In particular, it was observed that fewer self-references and occurrences of more future tense verbs and cognitive complexity tend to more complex cognitive patterns with increasing age, whereas no patterns were detected with positive or negative emotions, word size, first person plural, social references, past and present verbs, and time-related words. Significant patterns were detected in Millay, Eliot, Baillie, Yeats, Wordsworth, and Graves who revealed strong significant simple within-subject correlations between their age and their respective aging coefficients. No change was detected for Alcott, Austen, Dickens, and Shakespeare.

6 Conclusion

The description of the presented stylochronometric studies has made evident the plethora of subjectively guided complex stylochronometric practices according to the aims of the research at hand in terms of text selection, text size, sample size, stylistic markers, and statistical and computational methods, not to mention frequent use of material of uncertain composition date utilized for the completion of each individual study. Stylistic development as expressed by the frequency distributions of particular stylistic characteristics, as well as sparse evidence of stylistic stability in markers thought initially to depict change, have been detected. In addition, inconsistency in rectilinear trends was also present; that is, not every marker in every author appeared to decline in relation to advancing age; increasing trends were noted, too. However, although it is not currently possible to compare different studies due to lack of a unified methodology and standardized approaches, a number of different types of stylistic markers were judged to be of interest based on their rate of appearance:

(1) Punctuation marks: Whissell (1996), Forsyth (1999), Jackson (2002).

(2) Character markers: Ledger (1989), Ule (1982), Forsyth (1999).
(3) Parts of speech: Jaynes (1980), Ule (1982), Opas (1996), Whissell (1996), Pennebaker and Stone (2003).
(4) Most common words including function words: Brainerd (1980), Ule (1982), Frischer (1991), Forsyth et al. (1999), Whissell (1999), Can and Patton (2004).
(5) Various word frequencies subjectively selected and more closely related to content analysis: Jaynes (1980), Derks (1994), Whissell (1996), Laffal (1997), Pennebaker and Stone (2003).
(6) Syllable markers: Cox and Brandwood (1959), Wishart and Leach (1970), Jaynes (1980), Forsyth et al. (1999).
(7) Hapax legomena and/or hapax dislegomena: Ule (1982), Smith and Kelly (2002).
(8) Vocabulary richness: Ule (1982), Smith and Kelly (2002).

Other variable types shown to be more meter- and genre-oriented (Brainerd, 1980; Devine and Stephens, 1981; Cropp and Fick, 1985; Temple, 1996) have been judged as rather specialized, therefore, they have not been included in the list. It is maintained that commonly used markers may be of valuable use under the assumption that, since they have been incorporated in more than one study, they are of possible universal potential. It is also believed that, at this stage, the devising or discovery of new types of markers is not strictly necessary since the markers already used by stylochronometrists have not been fully explored. Consequently, their further and detailed investigation is recommended in order to identify the types of markers that work best, ideally universally, the patterns they follow, the different linguistic levels that are being combined and affected by them, and the circumstances under which detection and optimum function occurs.

These markers may play a central role initially in experimental studies which have been designed to preferably incorporate either well-dated material or material for which there is enough external evidence to support detected patterns or lack of them, in addition to some form of standardization regarding text selection, text size, sample size, and statistical

and computational methods which will offer an opportunity for comparisons. Thus, it will be possible to further observe standardization's feasibility and usefulness in stylochronometry as an improved approach which will either distract researchers from reinventing the wheel or accept it as the only suitable solution.

However, the usefulness of such an extended approach, besides the opportunity for direct comparisons and an automatic flavour to chronological problems, lies in the fact that, even if no single universal stylochronometer or type of it is eventually obtained, the possibility still exists that stylistic markers or types would be able to categorize authors in clusters according to detected patterns assuming that some markers will continue to work in some cases and not in others. Therefore, instead of testing for one single universal stylochronometer or type, it may be possible to test three or four or more depending on the level of detailed investigated.

Furthermore, other variables could possibly be incorporated to account for dimensions not readily captured by stylistic variables only, such as psychological factors, historical or personal events, health issues, etc. since documents are neither produced in isolation nor are completely independent from their author, his/her immediate surroundings, the culture he/she lives in, or time era. In addition, it would be of interest to investigate any influence factors such as authorship and genre may have on any detected statistical trends, tested on a variety of suitable textual material beyond the range of genre types already investigated and not strictly belonging to famous authors to examine the magnitude of stylistic change that one may detect.

Stylochronometry has untapped potential and an abundance of questions and problems that are waiting to be re-defined and explored. It is hoped that its complicated nature will attract enough interest to rekindle the search for answers from a different perspective and in a wider and less isolated research scale than previously. It is expected that in due course its results, be that positive or negative, will contribute to the broader understanding of the creativity process related to the written word, and ultimately in the better understanding of the individual human mind as it evolves in the course of its life time.

Acknowledgements

References

Biber, D. (1988). *Variation Across Speech and Writing*. Cambridge: Cambridge University Press.

Brainerd, B. (1980). The chronology of Shakespeare's plays. *Computers and the Humanities*, **14**: 221–30.

Brandwood, L. (1990). *The Chronology of Plato's Dialogues*. Cambridge: Cambridge University Press.

Burrows, J. F. (1987). Word-patterns and story-shapes: the statistical analysis of literary style. *Literary and Linguistic Computing*, **2**(2): 61–70.

Burrows, J. F. (1992). Not unless you ask nicely: the interpretative nexus between analysis and information. *Literary and Linguistic Computing*, **7**(2): 91–109.

Campbell, L. (1867). *The Sophistes and Politicus of Plato*. Oxford: Clarendon Press.

Can, F. and Patton, J. M. (2004). Change of writing style with time. *Computers and the Humanities*, **38**: 61–82.

Cox, D. R. and Brandwood, L. (1959). On a discriminatory problem connected with the works of Plato. *Journal of the Royal Statistical Society B*, **21**: 195–200.

Craik, E. M. and Kaferly, D. H. A. (1987). The computer and Sophocles' *Trachiniae*. *Literary and Linguistic Computing*, **2**(2): 86–97.

Cropp, M. and Fick, G. (1985). Resolutions and chronology in euripides: the fragmentary tragedies. *Bics Supplement*, **43**.

Devine, A. M. and Stephens, L. D. (1981). A new aspect of the evolution of the trimeter in euripides. *Transactions of the American Philological Association*, **111**: 43–64.

Derks, P. L. (1994). Clockwork Shakespeare: the bard meets the regressive imagery dictionary. *Empirical Studies of the Arts*, **12**(2): 131–9.

Duckworth, G. E. (1965). Horace's hexameters and the date of the *Ars Poetica*. *TAPA*, **96**: 73–95.

Evans, B. G., Levin, H., Baker, H. *et al.* (1974). *The Riverside Shakespeare*. Boston: Houghton Mifflin.

Forsyth, R. S. (1999). Stylochronometry with substrings, or: a poet young and old. *Literary and Linguistic Computing*, 14(4): 467–77.

Forsyth, R. S., Holmes, D. I. and Tse, E. K. (1999). Cicero, Sigonio, and Burrows: Investigating the Authenticity of the *Consolatio*. *Literary and Linguistic Computing*, 14(3): 375–400.

Frischer, B. (1991). *Shifting Paradigms: New Approaches to Horace's Arts Poetica*. Atlanta GA Scholars Press.

Jackson, MacD. P. (2002). Pause patterns in Shakespeare's verse: canon and chronology. *Literary and Linguistic Computing*, 17(1): 37–46.

Jaynes, J. T. (1980). A search for trends in the poetic style of W.B. Yeats. *Association for Literary and Linguistic Computing Journal*, 1: 11–19.

Keyser, P. (1992). Stylometric method and the chronology of Plato's works (review article). *Bryn Mawr Classical Review*, 1: 58–74.

Laffal, J. (1995). A concept analysis of Jonathan Swift's 'A Tale of a Tub' and 'Gulliver's Travels'. *Computers and Humanities*, 29(5): 339–62.

Laffal, J. (1997). Union and separation in Edgar Allan Poe. *Literary and Linguistic Computing*, 12(1): 1–13.

Ledger, G. (1989). *Re-counting Plato: A Computer Analysis of Plato's Style*. Oxford: Clarendon Press.

Martindale, C. (1990). *The Clockwork Muse: The Predictability of Artistic Change*. New York: Basic Books.

Opas, L. L. (1996). A Multi-Dimensional Analysis of Style in Samuel Beckett's Prose Works. In Hocking, S. and Ide, N. (eds), *Research in Humanities Computing 4*. Oxford: Clarendon Press.

Oras, A. (1960). *Pause Patterns in Elizabethan and Jacobean Drama*. Gainesville, FL: University of Florida Press.

Orlov, J. K. (1983). Ein Model der Häufigkeitsstruktur des Vokabulars. *Studies on Zipf's Law*. Bochum: Brockmeyer.

Pennebaker, J. W., Francis, M. E. and Booth, R. J. (2001). *Linguistic Inquiry and Word Count (LIWC): LIWC2001*. Mahwah, NJ: Erlbaum Publishers.

Pennebaker, J. W. and Stone, L. D. (2003). Words of wisdom: language use over the lifespan. *Journal of Personality and Social Psychology*, 85(2): 291–301.

Robinson, T. M. (1992). Plato and the computer. *Ancient Philosophy*, 12: 375–82.

Smith, J. A. and Kelly, C. (2002). Stylistic constancy and change across literary corpora: using measures of lexical richness to date works. *Computers and the Humanities*, 36: 411–30.

Temple, J. T. (1996). A multivariate synthesis of published Platonic stylometric data. *Literary and Linguistic Computing*, 11(2): 67–75.

Thesleff, H. (1967). *Studies in the Styles of Plato*. Helsinki: Suomalaisen Kirjallisuuden Kirjapaino.

Ule, L. (1982). Recent progress in computer methods of authorship determination. *Association of Literary and Linguistic Computing Bulletin*, 10(3): 73–89.

Wells, S., Taylor, G., Jowett, J. and Montgomery, W. (1987). *William Shakespeare: A Textual Companion*. Oxford: Clarendon Press.

Whissell, C. (1996). Traditional and emotional stylometric analysis of the songs of Beatles Paul McCartney and John Lennon. *Computers and the Humanities*, 30: 257–65.

Whissell, C. (1999). Holding emotional and linguistic rulers up to the poetry of Robert Frost. *Psychological Reports*, 85: 751–8.

Whissell, C., Fournier, M., Pelland, R., Weir, D. and Makarec, K. (1986). A dictionary of affect in language: IV. Reliability, validity, applications. *Perceptual and Motor Skills*, 62: 875–88.

Wishart, D. and Leach, S. V. (1970). A multivariate analysis of Platonic prose rhythm. *Computer Studies in the Humanities and Verbal Behaviour*, 3: 90–9.

Yule, G. (1994). *The Statistical Study of Literary Vocabulary*. Cambridge: Cambridge University Press.

Notes

1 The markers Jaynes selected were frequencies of heads of noun phrases (nouns, pronouns), portion of head which are nouns, ratio of determiners to nouns, ratio of adjectives to nouns, ratio of preps to nouns, heads of verbal phrases (main verb followed by infinitive followed by participle), portion of heads which are main verbs, ratio of auxiliary verbs to verbal phrase heads, adverb proportion, conjunctions (coordinating and subordinating conjunctions, and relative and interrogative pronouns), portion of conjunctions which are coordinators, miscellaneous such as clause and phrase signal, 'not', and interjections, nominal to verbal generalizations, preposition phrase words, and predicate as part of the main clause-elements (predicate, subject, compliment) (Jaynes, 1980).

2 Jaynes's extra variables were first person pronouns, mean proportion of monosyllabic words, ratio of content to function words over the three life periods, and the general vocabulary based on three identifiable period-themes, namely Celtic-mythology-related words additionally involving love and romance, love and romance mixed with 'words of a much lower sort' such as 'dolt, sweat, drunkards, jolt', informality, and Classical Greek, Italian, and Modern Irish mythical figures, and finally a compromise of themes between the early and middle periods.

3 Biber (1988) identified a number of textual dimensions able to distinguish between speech and writing based on different situational contexts and verbal production means and purposes.

4 The texts exhibiting such characteristics, which set them apart from the earlier texts, had been produced after the 1960s when Beckett had begun experimenting with minimalism and had adopted the style for the remainder of his life. In addition, 'the spread of the scores along all the dimensions also suggests versatility in style not normally found in any other type of discourse' (Opas, 1996).

5 Emotional stylometry refers to a combination of traditional stylometric measures such as word length, word frequencies, use of first, second, and third person personal pronouns, punctuation, type/token ratio, song length, negatives, the definite and indefinite articles, the conjunction 'and', word forms of 'love', 'girl', 'dead', whoops, hollers, and nonsense syllables.

6 According to the theory of word emotionality, a word carries two different meanings, denotative or descriptive and connotative or emotional. The former is considered rather complex and hard to quantify, whereas the latter may be represented by three dimensions such as evaluation, potency, and activation—although in later years these dimensions have been reduced to two, evaluation and activation, since they are capable of explaining approximately 80% of the differences among words which have been rated according to binary categories such as good–bad, fast–slow, etc. (Whissell, 1996). In the present study, Whissell extended her examination of word emotionality to six emotional measures, namely pleasantness, activation, nastiness, cheerfulness, sadness, and softness (Whissell, 1996).

7 In particular, Ledger used thirty-seven variables which represented the proportion of words containing certain letters; ALETS corresponds to the proportion of words containing a given letter, BLETS those words which end in a given letter, and CLETS words that contain a letter in penultimate position.

8 Minimum spanning tree (MST) is an algorithmic aid which joins each point in multidimensional space to its nearest neighbor in order to clearly define relations (Temple, 1996).

9 The reply variables were excluded from the second analysis since reply data for all the dialogues do not exist.

10 The nine different variable types that Ule (1982) used were percentage of one, two, three, up to forty-letter words; length of sentences in terms of one, two, three, up to forty-words per sentence; relative frequencies per 100 of the forty most common words in Elizabethan plays; relative frequencies per 100 of forty prepositions found in Elizabethan English, forty function words; twenty-six letters of the alphabet and the ten Arabic numbers; relative frequencies per 100 of hapax legomena, hapax dislegomena up to words occurring forty times; the percentage frequency distribution of n times V_n; and the percentage relative vocabulary overlap (RVO).

11 Euclidean distance calculates the square root of the sum of squares and produces a sequence of composition according to which the distance between neighboring texts is minimum.

12 These were *ad*, *sed*, *per*, and *nec* (Frischer, 1991).

13 Two analyses were performed, one for *ad*, *sed*, and *per*, and one for *nec*. This is because the first three words had similar distributions of a 'high-low-high pattern' in all the chronological periods whereas *nec* had 'the mirror reversal of the first' (Frischer, 1991).

14 Although this was true only when the first three function words were examined.

15 Porson's bridge refers to 'the nature of the word boundaries permitted to stand between a long third anceps and the final cretic segment of the trimeter' (Devine and Stephens, 1981).

16 The period before 431 is characterized by the plays *Alcestis*, *Medea*, and *Hippolytus* (Cropp and Fick, 1985).

17 As Brainerd (1980) explains, split lines occur 'when a verse line begun by one character is finished by a second'.

18 The ten selected variables were relative frequency of split lines among lines of verse, 'because', the lemma *might*, 'most', 'which', and the lemma *who* which showed an increasing pattern, and the lemma *more*, the lemma *with*, the lemma *unto*, and 'why' which were depicted as declining.

19 Incongruous juxtaposition refers to the 'presence of distant metaphors' as measured by 'structured metaphors...creative combinations of words [and]... approach-avoidance' word pairs like sweet–sorrow,

senior–junior, etc. Primordial content refers to words that indicate primitive and arousing drives and sensations expressed by mental material. Finally, conceptual content refers to 'analysis, discrimination, and conscious awareness' (Derks, 1994).

20 The Yeatsean index is of the form $YYIX = (YY-OY)/(YY+OY)$. $YYIX$ stands for Yeatsean Index, YY for Young Yeats and OY for Old Yeats.

21 It had emerged that the language of the text is uncharacteristic of Cicero's writing style, with a high probability that it had been composed during the renaissance and not during the classical era. The analysis also pointed at Sigonio being the most likely author; however, the result was inconclusive, possibly due to Sigonio's superb skills as an imitator of Cicero.

22 Burrows's approach concentrates on the location of the most common words in the text under investigation (which could range from thirty to 100, for example, while usually content words are excluded) and the use of those words as stylistic markers based on rates of occurrence. Often these markers are analysed with the help of multivariate statistics (Forsyth et al., 1999).

23 The language dimensions examined were emotional processes (positive emotions, negative emotions), social and identity concerns (first person singular, first person plural, social references), time orientations (time-related words, past tense verbs, present tense verbs, future tense verbs), and cognitive complexity (words greater than six letters, total cognitive words, causation, insight, exclusive) (Pennebaker and Stone, 2003). The language dimensions were measured with the help of LIWC 2001 (Pennebaker et al., 2001).

24 The authors were Charles Dickens, Louisa May Alcott, Jane Austen, and George Eliot (novelists), William Shakespeare and Joanna Baillie (playwrights), and St Vincent Millay, William Butler Yeats, Robert Graves, and William Wordsworth (poets).