A Computational Approach to Poetic Structure, Rhythm and Rhyme

Rodolfo Delmonte
Department of Linguistic Studies and Comparative Cultures
Ca' Bembo 1075 – Università Ca' Foscari – 30123 Venezia
Email: delmont@unive.it

Abstract

English. In this paper we present *SPARSAR*, a system for the automatic analysis of English and Italian poetry. The system can work on any type of poem and produces a set of parameters that are then used to compare poems with one another, of the same author or of different authors. In this paper, we will concentrate on the second module, which is a rule-based system to represent and analyze poetic devices. Evaluation of the system on the basis of a manually created dataset - including poets from Shakespeare's time down to T.S.Eliot and Sylvia Plath - has shown its high precision and accuracy approximating 90%.

Italiano. *In questo lavoro presentiamo SPARSAR, un sistema per l'analisi automatica di poesia inglese e italiana. Il sistema è in grado di lavorare su qualunque poesia e produce un insieme di parametri che vengono poi usati per confrontare poesie e autori tra di loro. In questo lavoro ci concentreremo sul secondo modulo che consiste in un sistema a regole per rappresentare e analizzare i dispositivi e le tecniche poetiche.*

Introduction

In this paper we present *SPARSAR*[1], a system for the automatic analysis of English and Italian poetry. The system can work on any type of poem and produces a set of parameters that are then used to compare poems with one another, of the same author or of different authors. The output can be visualized as a set of coloured boxes of different length and width and allows a direct comparison between poems and poets. In addition, parameters produced can be used to evaluate best similar candidate poems by different authors by means of Pearson's correlation coefficient. The system uses a modified version of *VENSES*, a semantically oriented NLP pipeline (Delmonte et al., 2005). It is accompanied by a module that works at sentence level and produces a whole set of analysis both at quantitative, syntactic and semantic level. The second module is a rule-based system that converts each poem into phonetic characters, it divides words into stressed/unstressed syllables and computes rhyming schemes at line and stanza level. To this end it uses grapheme to phoneme translations made available by different sources, amounting to some 500K entries, and include CMU dictionary, MRC Psycholinguistic Database, Celex Database, plus our own database made of some 20,000 entries. Out of vocabulary words are computed by means of a prosodic parser we implemented in a previous project (Bacalu & Delmonte, 1999a,b).

The system has no limitation on type of poetic and rhetoric devices, however it is dependent on language: Italian line verse requires a certain number of beats and metric accents which are different from the ones contained in an English iambic pentameter. Rules implemented can demote or promote word-stress on a certain syllable depending on selected language, line-level syllable length and contextual information. This includes knowledge about a word being part of a dependency structure either as dependent or as head. A peculiar feature of the system is the use of prosodic measures of syllable durations in msec, taken from a database created in a previous project(Bacalu & Delmonte, 1999a,b). We produce a theoretic prosodic measure for each line and stanza using mean durational values associated to stressed/ unstressed syllables. We call this index, "prosodic-phonetic density index", because it contains count of phones plus count of theoretic durations: the index is intended to characterize the real speakable and audible consistency of each line of the poem. A statistics is issued at different levels to evaluate distributional properties in terms of standard deviations, skewness and kurtosis. The final output of the system is a parameterized version of the poem which is then read aloud by a TTS

[1] The system is available at sparsar.wordpress.com and will soon be made interactive via a webservice.

system: parameters are generated taking into account all previous analysis including sentiment or affective analysis and discourse structure, with the aim to produce an expressive reading.

This paper extends previous conference and demo work (SLATE, Essem, EACL), and concentrates on the second module which focuses on poetic rhythm. The paper is organized as follows: the following section 2 is devoted to present the main features of the prosodic-phonetic system with some example; we then present a conclusion and future work.

2 The prosodic-phonetic module of the system

As R.Tsur(2012) comments in his introduction to his book, iambic pentameter has to be treated as an abstract pattern and no strict boundary can be established. The majority of famous English poets of the past, while using iambic pentameter have introduced violations, which in some cases – as for Milton's Paradise Lost – constitute the majority of verse patterns. Instead, the prosodic nature of the English language needs to be addressed, at first. English is a stress-timed language as opposed to Spanish or Italian which are syllable-timed languages. As a consequence, what really matters in the evaluation of iambic pentameters is the existence of a certain number of beats – 5 in normal cases, but also 4 in deviant ones. Unstressed syllables can number higher, as for instance in the case of exceptional feminine rhyme or double rhyme, which consists of a foot made of a stressed and an unstressed syllable (very common in Italian), ending the line - this is also used by Greene et al. 2010 to loosen the strict iambic model. These variations are made to derive from elementary two-syllable feet, the iamb, the trochee, the spondee, the pyrrich. According to the author, these variations are not casual, they are all motivated by the higher syntactic-semantic structure of the phrase. So there can be variations as long as they are constrained by a meaningful phrase structure.

In our system, in order to allow for variations in the metrical structure of any line, we operate on the basis of syntactic dependency and have a stress demotion rule to decide whether to demote stress on the basis of contextual information. The rule states that word stress can be demoted in dependents in adjacency with their head, in case they are monosyllabic words. In addition, we also have a promotion rule that promotes function words which require word stress. This applies typically to ambiguously tagged words, like "there", which can be used as expletive pronoun in preverbal position, and be unstressed; but it can also be used as locative adverb, in that case in postverbal position, and be stressed. For all these ambiguous cases, but also for homographs not homophones, tagging and syntactic information is paramount.

Our rule system tries to avoid stress clashes and prohibits sequences of three stressed/three unstressed syllables, unless the line syntactic-semantic structure allow it to be interpreted otherwise. Generally speaking, prepositions and auxiliary verbs may be promoted; articles and pronouns never. An important feature of English vs. Italian is length of words in terms of syllables. As may be easily gathered, English words have a high percentage of one-syllable words when compared to Italian which on the contrary has a high percentage of 3/4-syllable words.

In the two tables below we show percentages of

	1-syll. words	2-syll. words	Total 1+2	Total words	Percent
English CELEX	34269	102204	136,473	213,266	63%
English CMU	15945	55426	71371	115,000	62%
Italian PHONit	1496	15258	16,754	120,000	13.96%
Italian SIWL	30	2432	2462	31291	7.9%
Italian ITDict	3012	3989	7001	56000	12%
Totals	53256	164051	217307	535,557	40.58%

Table 1. English/Italian Quantitative 1- 2-Syllable Word Statistics

	Tot 3-5 syll. words	Total words	Perc.
Italian PHONit	97,485	120,000	81.23%
Italian SIWL	22861	31291	73.06%
Italian ITDict	44098	56000	78.75%
Totals	217307	535,557	40.58%

Table 2. Italian Quantitative 3- 5-Syllable Word Statistics

syllable length as contained in phonetic dictionaries of Italian vs English[2].

2.1 Computing Metrical Structure and Rhyming Scheme

Any poem can be characterized by its rhythm which is also revealing of the poet's peculiar style. In turn, the poem's rhythm is based mainly on two elements: meter, that is distribution of stressed and unstressed syllables in the verse, presence of rhyming and other poetic devices like alliteration, assonance, consonance, enjambments, etc. which contribute to poetic form at stanza level. This level is combined then with syntax and semantics to produce the adequate breath-groups and consequent subdivision: these will usually coincide with line stop words, but they may continue to the following line by means of enjambments.

What is paramount in our description of rhythm, is the use of the acoustic parameter of duration. The use of acoustic duration allows our system to produce a model of a poetry reader that we implement by speech synthesis. The use of objective prosodic rhythmic and stylistic features, allows us to compare similar poems of the same poet and of different poets both prosodically and metrically. To this aim we assume that syllable acoustic identity changes as a function of three parameters: internal structure in terms of onset and rhyme which is characterized by number of consonants, consonant clusters, vowel or diphthong; position in the word, whether beginning, end or middle; primary stress, secondary stress or unstressed.

The analysis starts by translating every poem into its phonetic form - see Figure 1. in the Appendix. After reading out the whole poem on a line by line basis and having produced all phonemic transcription, we look for rhetoric devices. Here assonances, consonances, alliterations and rhymes are analysed and then evaluated. Then we compute metrical structure, that is the alternation of beats: this is computed by considering all function or grammatical words which are monosyllabic as unstressed. We associate a “0” to all unstressed syllables, and a value of “1” to all stressed syllables, thus including both primary and secondary stressed syllables. We try to build syllables starting from longest possible phone sequences to shortest one. This is done heuristically trying to match pseudo syllables with our syllable list. Matching may fail and will then result in a new syllable which has not been previously met. We assume that any syllable inventory will be deficient, and will never be sufficient to cover the whole spectrum of syllables available in the English language. For this reason, we introduced a number of phonological rules to account for any new syllable that may appear. To produce our prosodic model we take mean durational values. We also select, whenever possible, positional and stress values. We also take advantage of syntactic information computed separately to highlight chunks’ heads as produced by our bottomup parser. In that case, stressed syllables take maximum duration values. Dependent words on the contrary are “demoted” and take minimum duration values.

Durations are then collected at stanza level and a statistics is produced. Metrical structure is used to evaluate statistical measures of its distribution in the poem. As a final result, we found out that it is difficult to find lines with identical number of syllables, identical number of metrical feet and identical metrical verse structure. If we consider the sequence “01” as representing the typical iambic foot, and the iambic pentameter as the typical verse metre of English poetry, there is no poem strictly respecting it in our transcription. On the contrary we find trochees, “10”, dactyls, “100”, anapests, “001”and spondees, “11”. At the end of the computation, the system is able to measure two important indices: “mean verse length” and “mean verse length in no. of feet” that is mean metrical structure.

Additional measures that we are now able to produce are related to rhyming devices. Since we intended to take into account structural internal rhyming schemes and their persistence in the poem we enriched our algorithm with additional data. These measures are then accompanied by information derived from two additional component: word repetition and rhyme repetition at stanza level. Sometimes also refrain may apply, that is the repetition of an entire line of verse. Rhyming schemes together with metrical length, are the strongest parameters to consider when assessing similarity between two poems.

Eventually we reconstruct the internal structure of metrical devices used by the poet: in

[2] For English we use the CMU syllable dictionary, the MRC Psycholinguistic Database, the database contained in the CELEX LDC distribution CD. For Italian, we used our own material amounting to some 100K phonetically transcribed lemmata and wordforms taken from a frequency list computed on 500K tokens of text. We also use PhoneItalia data (see Goslin et al., 2013)

some cases, also stanza repetition at poem level may apply. We then use this information as a multiplier. The final score is tripled in case of structural persistence of more than one rhyming scheme; it is doubled for one repeated rhyme scheme. With no rhyming scheme there will be no increase in the linear count of rethorical and rhyming devices. To create the rhyming scheme we assign labels to each couple of rhyming line and then match recursively each final phonetic word with the following ones, starting from the closest to the one that is further apart. Each time we register the rhyming words and their distance. In the following pass we reconstruct the actual final line numbers and then produce an indexed list of couples, Line Number-Rhyming Line for all the lines, stanza boundaries included. Eventually, we associate alphabetic labels to the each rhyming verse starting from A to Z. A simple alphabetic incremental mechanism updates the rhyme label. This may go beyond the limits of the alphabet itself and in that case, double letters are used.

What is important for final evaluation, is persistence of a given rhyme scheme, how many stanzas contain the same rhyme scheme and the length of the scheme. A poem with no rhyme scheme is much poorer than a poem that has at least one, so this needs to be evaluated positively and this is what we do. Rhetorical and rhyming devices are then used, besides semantic and conceptual indices, to match and compare poems and poets.

SPARSAR visualizes differences by increasing the length and the width of each coloured bar associated to the indices (see Figure 2. in the Appendix). Parameters evaluated and shown by coloured bars include: Poetic Rhetoric Devices (in red); Metrical Length (in green); Semantic Density (in blue); Prosodic Structure Dispersion (in black); Deep Conceptual Index (in brown); Rhyming Scheme Comparison (in purple). Their extension indicates the dimension and size of the index: longer bars are for higher values. In this way it is easily shown which component of the poem has major weight in the evaluation.

Parameters related to the Rhyming Scheme (RS) multiply metrical structure which includes: a count of metrical feet and its distribution in the poem; a count of rhyming devices and their distribution in the poem; a count of prosodic evaluation based on durational values and their distribution. RS is based on the regularity in the repetition of a rhyming scheme across the stanzas or simply the sequence of lines in case the poem is not divided up into stanzas. We don't assess different RSs even though we could: the only additional value is given by the presence of a Chain Rhyme scheme, that is a rhyme present in one stanza which is inherited by the following stanza. Values to be computed are related to the Repetition Rate (RR), that is how many rhymes are repeated in the scheme or in the stanza: this is a ratio between number of verses and their rhyming types. For instance, a scheme like AABBCC, has a higher repetition rate (corresponding to 2) than say AABCDD (1.5), or ABCCDD (1.5). The RR is a parameter linked to the length of the scheme, but also to the number of repeated schemes in the poem: RS may change during the poem and there may be more than one scheme. A higher evaluation is given to full rhymes, which add up the number of identical phones, with respect to half-rhymes which on the contrary count only half that number. We normalize final evaluation to balance the difference between longer vs. shorter poems, where longer poems are rewarded for the intrinsic difficulty of maintaining identical rhyming schemes with different stanzas and different vocabulary.

In Figure 3. in the Appendix, general graded evaluation is shown for the first 53 Shakespeare's sonnets. Position in the space is determined by values of each of the six macro-indices as well as the overall skewness and kurtosis. Most valued sonnets are placed at the top and in the middle of the space, thus indicating the even distribution of their parameters. It is interesting to see that best ranked sonnet is no.29, which has always been regarded as one of the best of the collection.

3 Evaluation and Conclusion

We have done a manual evaluation by analysing a randomly chosen sample of 50 poems out of the 500 analysed by the system. The evaluation has been made by a secondary school teacher of English literature, expert in poetry. We asked the teacher to verify the following four levels of analysis: 1. phonetic translation; 2. syllable division; 3. feet grouping; 4. metrical rhyming structure. Results show a percentage of error which is around 5% as a whole, in the four different levels of analysis, thus subdivided: 1.8 for parameter 1; 2.1 for parameter 2; 0.3 for parameter 3; 0.7 for parameter 4.

References

Bacalu, C. & Delmonte R. (1999a), Prosodic Modeling for Syllable Structures from the VESD - Venice English Syllable Database, in Atti 9° Convegno GFS-AIA, Venezia.

Bacalu, C. & Delmonte R. (1999b), Prosodic Modeling for Speech Recognition, in Atti del Workshop AI*IA - "Elaborazione del Linguaggio e Riconoscimento del Parlato", IRST Trento, pp.45- 55.

Baayen, R. H., R. Piepenbrock, and L. Gulikers. 1995. *The CELEX Lexical Database (CD-ROM).* Linguistic Data Consortium.

Byrd, Roy J. and Martin S. Chodorow. 1985. Using an online dictionary to find rhyming words and pronunciations for unknown words. In Proceedings of the 23rd Annual Meeting of ACL, 277–283, Chicago, Illinois, USA, July.

Delmonte R. (1999), A Prosodic Module for Self-Learning Activities, Proc.MATISSE, London, 129-132.

Delmonte, R. 2013. SPARSAR: a System for Poetry Automatic Rhythm and Style AnalyzeR, SLATE 2013, Demonstration Track.

Delmonte R., Computing Poetry Style, in C.Battaglino, C.Bosco, E.Cambria, R.Damiano, V.Patti, P.Rosso(eds.), Proceedings of 1st International Workshop ESSEM 2013, CEUR Workshop Proc., n.1096, 148-155, Aachen.

Delmonte, R. & Anton Maria Prati, SPARSAR: An Expressive Poetry Reader. 2014. Proceedings EACL, Demonstration Track.

Delmonte R., Sara Tonelli, Marco Aldo Piccolino Boniforti, Antonella Bristot, Emanuele Pianta, 2005. VENSES – a Linguistically-Based System for Semantic Evaluation, in J. Quiñonero-Candela et al.(eds.), 2005. Machine Learning Challenges. LNCS, Springer, Berlin, 344-371.

Genzel, Dmitriy, Jakob Uszkoreit, and Franz Och. 2010. "Poetic" statistical machine translation: Rhyme and meter. In Proceedings of EMNLP.

Greene, Erica, Tugba Bodrumlu, Kevin Knight. 2010. Automatic Analysis of Rhythmic Poetry with Applications to Generation and Translation, in Proceedings of the 2010 Conference on EMNLP, 524–533, MIT, Massachusetts, USA, 9-11 October 2010.

Goslin, Jeremy, Claudia Galluzzi, Cristina Romani, (2013), PhonItalia: a phonological lexicon for Italian, in Behavior Research Methods, vol. 45, no. 3, pp.17.

Hayward, M. (1991). A connectionist model of poetic meter. Poetics, 20, 303-317.

Hayward, M. (1996). Application of a connectionist model of poetic meter to problems in generative metrics. Research in Humanities Computing 4. (pp. 185-192). Oxford: Clarendon P.

Kaplan, D., & Blei, D. (2007). A computational approach to style in American poetry. In *IEEE Conference on Data Mining.*

Keppel-Jones, David, 2001. The Strict Metrical Tradition: Variations in the Literary Iambic Pentameter from Sidney and Spenser to Matthew Arnold, Mcgill Queens Univ. Pr., 280.

Sonderegger, Morgan, 2011. Applications of graph theory to an English rhyming corpus. Computer Speech and Language, 25:655–678.

Sravana Reddy & Kevin Knight, 2011. Unsupervised Discovery of Rhyme Schemes, in Proceedings of the 49th Annual Meeting of ACL: shortpapers, 77-82, Portland, Oregon, June 19-24, 2011.

Reuven Tsur, 2012. Poetic Rhythm: Structure and Performance: An Empirical Study in Cognitive Poetics, Sussex Academic Press, 472.

APPENDIX 1.

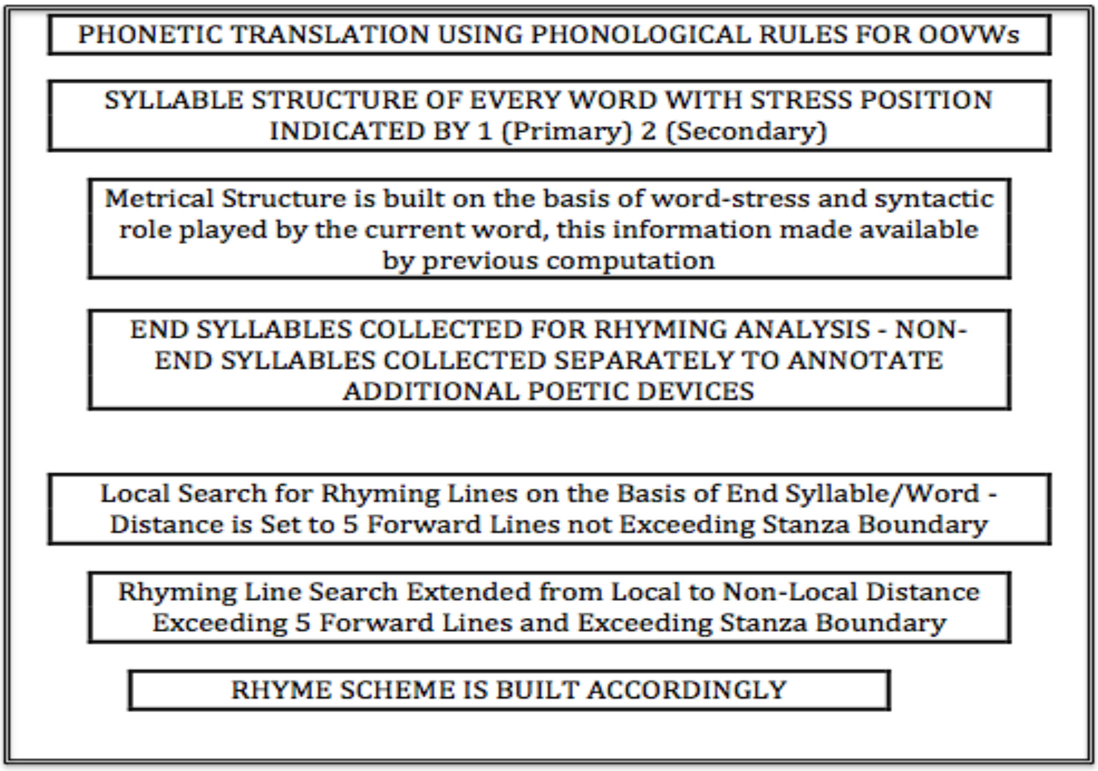

Figure 1. The Rhythm and Rhyme Module of SPARSAR Poetic Analyzer

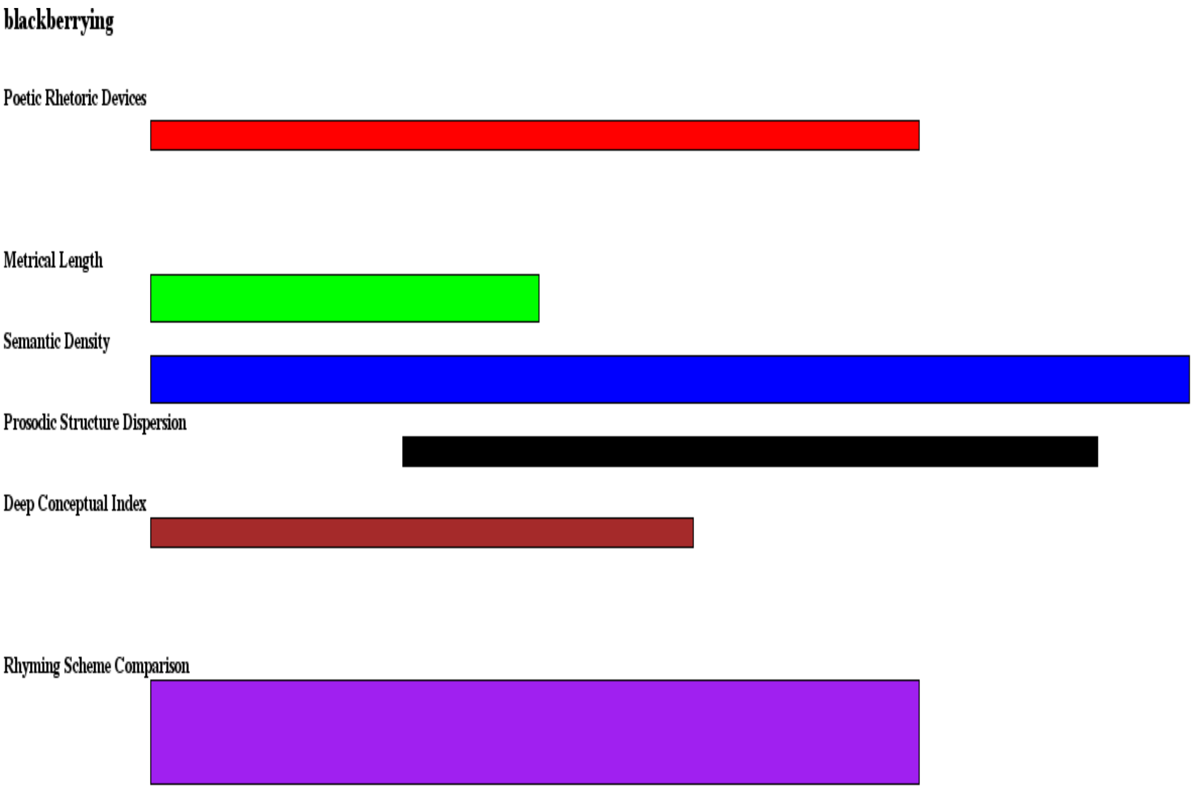

Figure 2. SPARSAR's six macroindices for Sylvia Plath's Blackberrying

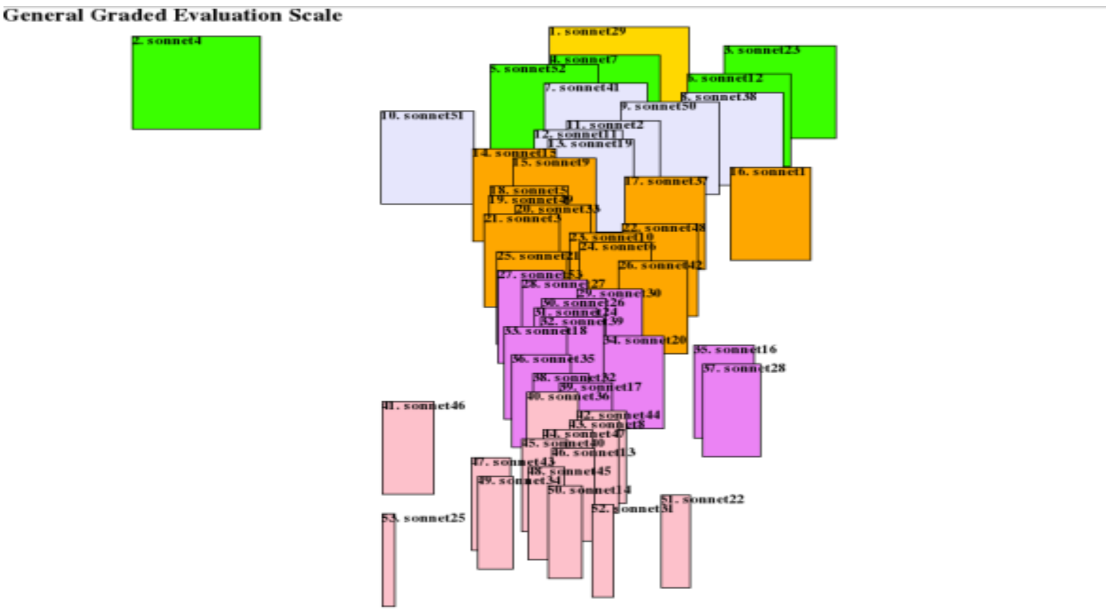

Figure 3. Graded Evaluation of 53 sonnets by William Shakespeare