Article

Not peer-reviewed version

Multidimensional Visualization of Sound-Sense Harmony for Shakespeare's Sonnets Classification

Rodolfo Delmonte * and Nicolò Busetto

Posted Date: 3 October 2024

doi: 10.20944/preprints202410.0286.v1

Keywords: SPARSAR = Specialized NLP system for English Poetry; Automatic Analysis of XVIth Century English Poetry; Multidimensional Visualization of Linguistic and Poetic Content; Creating Graphic Interface with Coloured Boxes of Different Size; Computing Sound-Sense Harmony; Comparing Phonetic and Phonological Features with Meaning; Automatic Lexical and Semantic Sentiment Analysis of Poetry; Appraisal Theory Framework

Article

Multidimensional Visualization of Sound-Sense Harmony for Shakespeare's Sonnets Classification

Rodolfo Delmonte [1,*] and Nicolò Busetto [2]

[1] Ca Foscari University of Venice (Italy) 1
[2] Meta, USA
* Correspondence: delmont@unive.it

Abstract: In this paper we focus on the association of sound and sense harmony in the collection of sonnets written by Shakespeare in the XVI° beginning of XVII° century and propose a new four-dimensional representation to visualize them by means of the system called ***SPARSAR***. To compute the degree of harmony and disharmony, we have automatically extracted the sound grids of all the sonnets and have combined them with the semantics and polarity expressed by their contents. We explain in details the algorithm and show the representation of the whole collection of 154 sonnets and comment them extensively. In a second experiment, we use data from manual annotation of the sonnets for Satire detection using the Appraisal Theory Framework, to gauge the system's accuracy in matching these data with the output of the automatic algorithm for sound-sense harmony. The results obtained with an 94.6% accuracy confirm the obvious fact that the poet has to account for both sound and meaning in the choice of words.

Keywords: SPARSAR = specialized NLP system for English poetry; automatic analysis of XVIth century English poetry; multidimensional visualization of linguistic and poetic content; creating graphic interface with coloured boxes of different size; computing sound-sense harmony; comparing phonetic and phonological features with meaning; automatic lexical and semantic sentiment analysis of poetry; appraisal theory framework

1. Introduction

In this article, we present in detail the visualization algorithm that affords a multidimensional representation of Shakespeare's Sonnets. In a previous paper [1], a totally new theory and technique to assess and appreciate poetry is proposed, the Algorithm for Sound and Sense Harmony (henceforth ASSH). The main tenet of the theory is the existence of a hidden and systematic plan realized by important poets like Shakespeare to organize poetic structures in accordance with the principles of an overall Sound and Sense Harmony.

What is meant here by "Sound-Sense Harmony" is a deep agreement between poetic and semantic devices: on the one side the presence of stressed syllables containing vowels belonging to the four sound classes that, phonologically speaking, may comprise contrasting vowel sounds, i.e., low, mid, vs. high-front, high-back, or part of them. The same applies to the presence of the three main classes of consonants, i.e., continuants and sonorants, and their contrasting opponents, obstruents and their partition into voiced vs. unvoiced. The "Sound Harmony" is then composed with Sense to make up the ASSH, where the choice of sounds is supposed to reflect partly or entirely, the contents of the poem - or else opposes its contents, as it may be represented by intended and expressed meaning and overall sentiment.

The choice to favour the presence of one class vs. another is to be interpreted as a way to highlight sense-related choices of words and phrases that will either accompany or contrast with Sounds. In particular, we associate different mood — following traditional judgements — to vowels and consonants according to their class, as follows:

1. Low and mid vowels evoke a sense of brightness, peace and serenity;

 doi:10.20944/preprints202410.0286.v1

2. High, front and back vowels evoke a sense of surprise, seriousness, rigor and gravity;
3. Obstruent and unvoiced consonants evoke a sense of harshness and severity;
4 Sonorant and continuant consonants evoke a sense of pleasure, softness and lightness.

Classes 1 and 4 are regarded in the same area of positive thinking, while classes 2 and 3 are more naturally accompanied by negative sentiment. Of course, it may be the case that crossed matches with classes belonging to opposite types will take place more or less frequently, indicating the need to reconcile opposite feelings in the same poem. This is what happens in most Shakespeare's sonnets, as will be shown in the sections below.

It is important to highlight the role of sounds in poetry, which is paramount for the creation of poetic and rhetoric devices. Rhyme, alliterations, assonances and consonances may contribute secondary and, in some cases, primary additional meaning by allowing words which are not otherwise syntactically or semantically related to share part if not all of their meaning by means of metaphors and other similar devices. Thus, most of the difficult work of every poet is devoted to the choice of the appropriate word to use for rhyming purposes, mainly, but also for the other important devices mentioned above.

In the case of Shakespeare, for the majority of the sonnets, he took care of choosing appropriate words for the rhymes contributing sounds to the four varieties, thus producing a highly varied sound harmony. However, as will appear more clearly in a section below, the most important result of the new evaluation is the existence of a full match between Dis-harmonic sonnets and the need to express sarcasm and irony, feelings that permeate the majority of Shakespeare's Sonnets.

The paper is organized as follows: in section 1.1 we present state of the art in the sound-sense harmony; in section 2, parameters and methods are described, including prosodic, poetic and semantic ones; in section 3, the multidimensional representation is presented and described in detail; in section 4 the previous work with ATF for satire detection is presented and compared; in section 5 a bried discussion and finally a Conclusion in section 6.

1.1. State of Art in the Association of Sound and Meaning in Poetry

The use of mood associated with sound in poetry by means of colours has a long tradition. Rimbaud composed a poem devoted to "Vowels" [2], where colours and moods were associated with each of the main five vowels. Roman Jakobson [3,4] and Mazzeo [5] wrote extensively about the connection between sound, mood and colour in a number of papers. Fónagy [6] wrote an article in which he explicitly connected the use of certain types of consonant sounds associated with certain moods: unvoiced and obstruent consonants are associated with aggressive mood; sonorants with tender moods. The most critical and important approach to the problem was presented by Mallarmé [7] in the article "Crise de Vers". Macdermott [8] identified a specific quality associated with "dark" vowels, i.e., back vowels, that of being linked with dark colours, mystic obscurity, hatred and struggle. The same applies to representing unvoiced and obstruent consonants as opposed to voiced and sonorants. But as Tsur [9,10], [p. 15] notes, this sound–colour association with mood or attitude has no real significance without a link to semantics. Here we are not claiming that there is always a direct link between the sound of a word and its meaning which would amount to saying that linguistic signs are not arbitrary (see Gérard Genette [11] and Dwight Bolinger, [12]). Word sounds in poetry are harmonized in order to convey abstract connections not always directly expressed by word meaning but related to emotions or moods. Current NLP technologies are boasting their ability to tell the "sentiment" expressed by the opinions of people commenting on some public event: this is what we are aiming to here – even though using different approaches, capturing the overall mood or "sentiment" by matching the analysis of the sounds to the polarity lexically or structurally contained in the words or phrases (see below).

To produce a feasible comparison, the negative sentiment expressed by the words' sense with sad-sounding rhymes poetic devices are matched as a whole, and the opposite for positive sentiment, and this is done by scoring and computing the ratios for each class of sounds. Here below is shown the way in which vowel and consonant sounds are organized:

 doi:10.20944/preprints202410.0286.v1

- Low, middle, high-front, high-back

where the two classes low and middle are identified as promoting positive feelings, and the two high as inducing negative ones: the ratios are produced by dividing the number of low-middle by the number of high.

As to the consonants, the sounds are organized into three main classes and two types:

- Obstruents (plosives, affricates), continuants (fricatives), sonorants (liquids, vibrants, approximants)

plus the distinction into

- Voiced vs. unvoiced.

In the case of consonants, the ratios are computed by dividing the sum of continuants and sonorants by the number of obstruents; a second parameter is the ratio obtained by dividing number of voiced by unvoiced. Whenever the value of the ratios is above 1, positive results are obtained; the contrary applies whenever values are below 1. In this way, counting results is immediate and very effective.

The lexical and semantic approach to deriving the sentiment of each sonnet operates a first subdivision of harmonic and dis-harmonic sonnets into negatively vs. positively marked sonnets. Measuring correlations reveals a constant contrasting attitude induced by the sound–sense agreement, which is interpreted as an underlying hidden intention to produce some form of ironic mood in Shakespeare's sonnets.

However, detecting irony requires a much deeper and accurate analysis of the semantic and the pragmatics of the sonnets. In a number of paper we published recently [13,14] an evaluation of the sonnets was proposed based on objective criteria and a comparison with the so-called Appraisal Theory Framework (ATF). To this aim, each sonnet has been manually annotated using the highly sophisticated labeling system proposed by the ATF (see [15,16])[1]. Matching the empirical approach and the automatic analysis confirms the overall underlying hypothesis: the sound–sense disharmony has a fundamental task, that of suggesting an underlying ironic attitude which is at the heart of all the sonnets [17,18][2]. ATF makes available a more fine-grained approach which takes non-literal language into due account, thus improving on previous methods of sentiment-based analysis.

The decision of adopting ATF is based on the fact that previous approaches to detect irony – including satire/sarcasm - in texts have failed to explain the phenomenon. Computational research on the topic has been based on the use of shallow features, so as in [19–22], in order to train statistical model with the hope that when optimized for a particular task, they would come up with a reasonably acceptable performance. However, they would not explain the reason why a particular Twitter snippet or short Facebook text has been evaluated as containing satiric/sarcastic expressions. Except perhaps for features based on text exterior appearance, i.e. use of specific emoticons, use of exaggerations, use of unusually long orthographic forms, etc. which however is not applicable to our poetic texts.

The other common approach used to detect irony, in the majority of the cases, is based on polarity detection. So-called Sentiment Analysis is in fact an indiscriminate labeling of texts either on a lexicon basis or on a supervised feature basis [23–25], where in both cases, it is just a binary decision that has to be taken. This is again not explanatory of the phenomenon and will not help in

1 The theory is organized around three subsystems, Engagement, Graduation and Attitude. We chose to select Attitude for our annotation guideline. The Attitude subsystem describes the author's feelings and intentions as they are conveyed within the text, and it is articulated into three main semantic regions or subclasses with their relative positive/negative polarity, namely: Affect, Judgement, Appreciation. More in this section below.

2 Weiser dedicates his papers to comment on the Sonnets from the point of view of the irony they contain. As a literary critic, he refers to a number of previous studies by other famous literary critics like Stephen Booth, Harold Bloom etc. who also regarded the sonnets highly ironic. The number of sonnets judged as such are all included in the list that we discuss in the sections below.

 doi:10.20944/preprints202410.0286.v1

understanding what is it that causes humorous reactions to the reading of an ironic piece of text. It certainly is of no help in deciding which phrases, clauses or just multiwords or simply words, contribute to create the ironic meaning[3].

In this paper, together with sound evaluation, in the multidimensional representation we will only consider the polarity extracted from the classification labels we are using in line with what is suggested by the ATF present in the Attitude subsystem, and they are Affect, Appreciation and Judgement both Negative and Positive – more on the use of ATF parameters below.

2. Materials and Methods

2.1. How the ASSH Works: Dimensions and Parameters

The current version of the system SPARSAR [28] has a visualization algorithm that computes a multidimensional representation. Poems are listed vertically into three separate categories organized into columns or sections: Negative Harmony on the left-hand side, Positive Harmony on the right-hand side, and Dis-harmonic sonnets in the center column. Sonnets placed between the center column and the other two columns have either Disharmonic sound or Disharmonic sense: the ones on the right or Positive column are partly Disharmonic and partly Positive. The same applies to sonnets placed om the left, which are partly Disharmonic and partly Negative. This is to be regarded as the first dimension because it is immediately recognizable. Negative or Positive sentiment may be determined either by the sound grid or by semantic polarity.

Then, as a second dimension, each list differentiates poems according to a ranking, derived from evaluations made on the basis of both poetic devices, prosodic devices, and linguistic contents. The algorithm makes a sum of the ranking in each respective parameter: the result is a smaller number for those poems that rank higher thus allowing a reranking that accounts for the total evaluation. In this way, poems are organized in a graded scale.

In Figure 1 we show a small sample of the multidimensional visualization. Each poem is enclosed in a coloured box/ellipse of varied dimension according to the values of six parameters, those referred to semantics, poetic devices and prosody: and this constitutes a third dimension again visually relevant. Those sonnets that satisfy entirely parameters requirements for each of the three main categories will be represented with a square or a rectangular shape according to their relevance in terms of parameters' satisfaction. Instead, we use ellipses for those sonnets that constitute an intermediate classification: this may happen every time there is no agreement between sounds and sense and sentiment labels, and the two labels don't coincide, as clarified below.

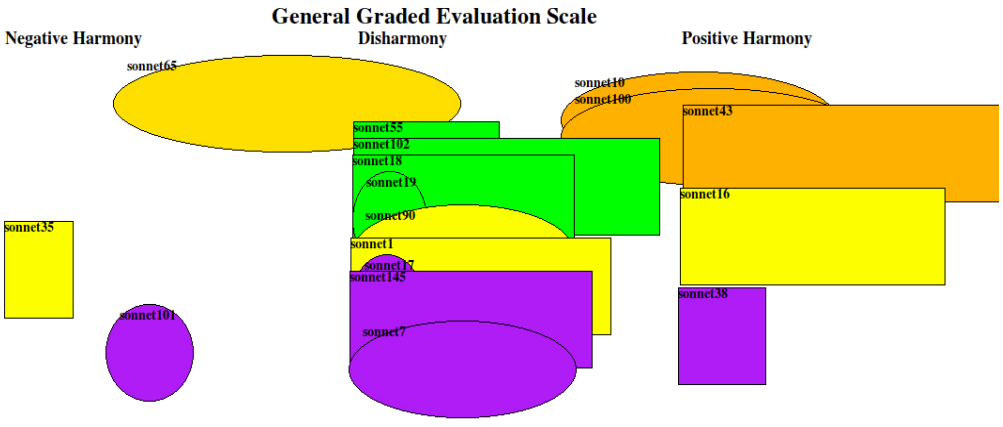


Figure 1. A short sample of the sound-sense multidimensional visualization.

[3] For a different approach to irony and satire detection see [26, 27]

 doi:10.20944/preprints202410.0286.v1

The fourth dimension regards sonnets that don't satisfy fully the sound-sense (dis)harmony in terms of parameters threshold values. Intermediate sonnets will be positioned always according either to sound or to polarity, but sound-sense harmony will cause their final position to be moved further on the center-right side for the negative sonnets, and the center-left side for the positive ones, thus making them occupy an intermediate position with the following or previous class-column.

The reason for using a multidimensional graphic representation is now starting to become more clearly motivated considering the number of parameters that contribute to shape and position each sonnet's image. In the following, we explore these parameters in detail. As will become clear, visualizing semantic and poetic content of a poem is by far a more effective way of communicating relations and differences than just using diagrams and tables.

As said above, the position in the graded scale is determined by what we called A General Description map which includes seven Macro Indices with a statistical evaluation of six parameters evaluated and used in the computation:

- Poetic Rhetoric Devices;
- Metrical Length;
- Semantic Density;
- Prosodic Structure Dispersion;
- Deep Conceptual Index;
- Rhyming Scheme Comparison.

The first parameter is obtained by the number of assonances, consonances present in the poem as a whole; the second parameter is obtained by scansion and comparison of number of syllables in each verse; the third parameter is obtained by checking semantic classification of each noun and verb as reported in our lexica and the difference between abstract/concrete, event/state; the fourth parameter is obtained by comparing the length of each verse in msecs; the fifth parameter is obtained by counting the number of abstract vs concrete references; finally the sixth parameter is obtained by counting the number of rhyming verses and repeated rhymes. See the following section for more information on all parameters.

2.2. Semantic and Conceptual Parameters

The Semantic Density Index - "Semantic Density Index" (SDI) - is derived from the computation of a number of features, some of which have negative import and others positive import.

With this definition we now refer to the possibility to classify poems according to their intrinsic semantic density in order to set apart those poems which are easy to understand from those that require a rereading and still remain somewhat obscure. An intuitive notion of SDI can be formulated as follow:

- easy to understand are those semantic structures which contain a proposition, made of a main predicate and its arguments
- difficult to understand are on the contrary semantic structures which are filled with nominal expressions, used to reinforce a concept and as such they are simply juxtaposed in a sequence
- also difficult to understand are sequences of adjectives and nominals used as modifiers, union of such items with a dash
- in the definition of semantic parameters presence of negation and modality contributes to increase complexity, for this reason we compute Polarity and Factuality at propositional level.

Additional features are obtained by measuring the level of affectivity by means of sentiment analysis, focusing on presence of negative items which contribute to make understanding more difficult.

At the end of the computation the index may end up to be positive if the poem is semantically "light", that is easy to read and understand; otherwise, it is computed as "heavy" which implies that it is semantically difficult.

 doi:10.20944/preprints202410.0286.v1

The index will have a higher value for those cases of high density and a lower value for the contrary. It is a linear computation and includes the following features:

- the ratio of number of words vs number of verbs;
- the ratio of number of verbal compounds vs non-verbal ones;
- the internal composition of non-verbal chunks:
- every additional content word increases their weight (functional words are not counted);
- the number of semantic classes.

To produce the ratio the score associated to each item is divided by the total number of tagged words and of chunks. In detail, these are the categories used:

- verbs found by the total number of tokens (the more the best);
- adjectives found by the total number of tokens (the more the worst);
- verb structures by the total number of chunks (the more the best);
- inflected vs uninflected verbal compounds (the more the best);
- nominal chunks rich in components: those that have more than 3 members (the more the worst);
- semantically rich (with less semantic categories) words by the total number of lemmas (the more the worst);
- rare words (the more the worst)[4];
- generic or collective referred concepts (the more the best);
- specific vs ambiguous semantic concepts (those classified with more than two senses on the basis of WordNet) (the more the worst);
- doubt and modal verbs, and propositional level negation (the more the worst);
- abstract and eventive words vs concrete concepts (the more the worst);
- we compute ATF based sentiment analysis with a count of negative polarity items (the more the worst).

The other important index we implemented is the Deep Conceptual index, which is obtained by considering the proportion of Abstract vs Concrete words contained in the poem.

This index is then multiplied with the Propositional Semantic Density which is obtained at sentence level by computing how many non verbal, and amongst the verbal, how many non inflected verbal chunks there are in a sentence.

A further conceptual index is captured using Martindale's Regressive Imagery Dictionary, RID [29]. Using the RID we obtain at first a very generic distinction into three psychologically-based set of concepts – primary (sensual, subjective), secondary (rational, abstract, objective), emotions. These three classes are internally further divided up into 65 classes which are highly specific. So using RID[5], a commonality of themes is established on the basis of the 65 classes and not simply on the threefold distinction ending up with another index to be added. To obtain an effective index, we normalize the absolute values of RID concepts found in every sonnet by dividing up each figure by the total number of entries for that category.

2.3. Poetic and Prosodic Parameters

Parameters related to the Rhyming Scheme (RS) multiply metrical structure and include:

- a count of metrical feet and its distribution in the poem;
- a count of rhyming devices and their distribution in the poem;

4 To tell whether a word belongs to the list of rare words we use a number of dictionaries of word-forms freely available to download. We classify rare words as those with a frequency value lower and equal to 3 occurrences.

5 In more detail the RID and its internal organization. RID has some 3200 so-called search patterns, which are roots and words as entries and they are so divided: 1800 belong to primary concepts, 728 to secondary concepts, 616 to emotions.

 doi:10.20944/preprints202410.0286.v1

- a count of prosodic evaluation based on durational values and their distribution.

This is how the ASSH works: after reading out the whole poem on a verse by verse basis and having produced all phonemic transcription the algorithm looks for rhetoric devices. Here assonances, consonances, alliterations and rhymes are analysed and then evaluated. Then it computes duration at verse level by associating theoretic mean durations at syllable level. Duration at syllable level is found by associating phonemes obtained by phonetic transcription into syllables. English syllables are at first matched with the dictionary: in case of failure, the parser tries to detect syllables on the basis of a principle of syllable wellformedness[6]. Syllable structure requires a nucleus to be in place, then a rhyme with an onset and offset. Syllables in our database may be up to 6 phonemes long. Durations have been recorded by means of a statistical study, with three different word positions: beginning, middle and end position. They have also been collected according to a prosodic criterion: stressed and unstressed syllables. Then, each syllable has been recorded with three durational values in msec.: minimum, mean and maximum duration length. Then, number of occurrences in the database has been associated and standard deviation for duration values has been obtained.

We use mean durational values to produce our prosodic model. We also select, whenever possible, positional and stress values. In case a syllable duration value is not available for those parameters, we choose the default unstressed value.

The analysis continues computing metrical structure, 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. Durations are then collected at stanza level and a statistics is produced. Statistical measures of metrical structure are used to evaluate its distribution in the poem: standard deviation reveals whether lines tend to have an equal duration or not.

Additional measure that the system is now able to produce are related to rhyming devices. Since we intended to take into account the structural internal rhyming scheme and rhymes persistence in the poems 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.

Parameters related to the Rhyming Scheme (hence RS) contribute a multiplier to the already measured metrical structure which is extracted from 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. The final score is tripled in case of structural persistence of more than one rhyming scheme or else if it is always the same repeated rhyme scheme it is doubled. With no rhyming scheme there will be no increase in the linear count of rhetorical and rhyming devices.

Now the RS is yet another plane or dimension on the basis of which a poem is evaluated. It is based on the regularity in the repetition of a rhyming scheme across the stanzas or simply the sequence of verses 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 computed are related to the Repetition Rate (hence 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). As a result, 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.

Different evaluations are given to full rhymes, which are added up to the number of identical phones, with respect to half-rhymes which on the contrary count only half that number. The final

6 We also use the list of word-forms divided into syllables freely made available as CMU Pronouncing Dictionary at the same website.

 doi:10.20944/preprints202410.0286.v1

value is obtained by dividing up the RR by the total number of lines multiplied by 100, and then summing the same number of total lines to the result. This is done 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.

3. Results

3.1. Multidimensional Visualization Commented

We will now show the multidimensional (dis)-harmony representation of the 154 Sonnets divided up into three separate diagrams, each containing approximately 50 poems and starting from the first sonnets. Each figure is assigned one separate page: this is done in order to allow the visualization attributes to be reasonably visible and not reduced too much to become unreadable. In an Appendix we provide an enlarged and a more focused image by dividing up each main figure shown here into four parts where sonnet numbers and their positions are more visible.

As said above, sonnets are aligned into three columns where sonnets with square and rectangular shape have parameters in full agreement, while sonnets in ellipses contain disagreements and also partially matching sentiment related parameters. In particular, in order to be situated in one of the three lists at first the semantic and sentiment-based sense evaluation is considered, then the sounds, which in case of poor match contribute to placing the figure closer to the column on the right-hand side, except for the Positively marked ones, which being the last column, will be placed on its left-hand side.

The first fifty-three poems in Figure 2. show a clear majority of sonnets expressing Dis–Harmony, (24), followed by those expressing Positive Harmony (7) plus those sonnets (10) positioned in the middle between positive and disharmony, indicating a positively based but not completely in harmony between sound and sense; and a small number of sonnets in the Negative Harmony section – only (3) fully compliant with negative harmony, including those sonnets (9) with elliptical shapes on the right-hand side. For instance, the first poem in the three ranks is sonnet 47 positioned on the Dis -Harmony side. Its box is elliptical and bigger than the other boxes below which is obtained by a high ranking in the evaluation of the six parameters mentioned above. If we check the ranking of the sonnet in the six parameters and make a sum we have the following result:

Poetic Rhetoric Devices	(11)
Metrical Devices	(3)
Semantic Density	(25)
Prosodic Structure	(2)
Deep Conceptual Index	(1)
Rhyming Scheme	(27)
Total Ranking	***(69)***

The smaller is the sum the bigger will be the corresponding shape, thus indicating high rankings. If we look now at Sonnet 38 which is lower in the Positive Harmony rank list and has been assigned a very small rectangular shape we have the following values and a much higher sum:

Poetic Rhetoric Devices	(41)
Metrical Devices	(33)
Semantic Density	(18)
Prosodic Structure	(15)
Deep Conceptual Index	(33)
Rhyming Scheme	(41)
Total Ranking	***(181)***

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 3 October 2024 doi:10.20944/preprints202410.0286.v1

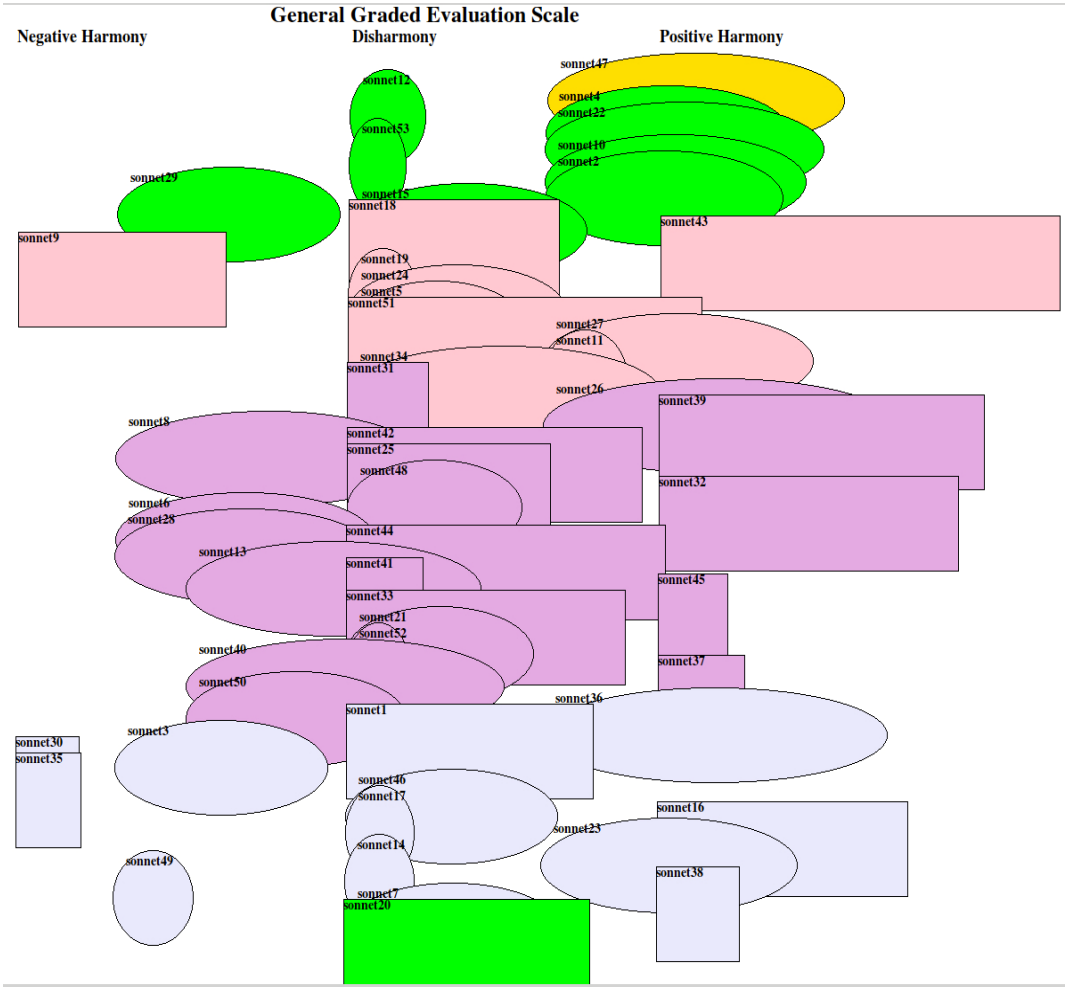


Figure 2. Multidimensional patterns for Sonnets 1-53.

As to the parameters determining their position in one of the three columns and the choice of the shape we are using the following labels: OK for POSITIVE, both for sound and sense; NO for NEGATIVE, again both for sound and sense. Then we use two intermediate labels to mark different levels of dis-harmony, EDGE and MIXED, where the first EDGE indicates a slight deviation and MIXED a more important dis-harmony. Hence fully dis-harmonic sonnets will be marked OK-NO, while slightly disharmonic EDGE-EDGE where however both sound and sense are in a similar condition; fully harmonic positive sonnets will be marked OK-OK and negative ones NO-NO.

Full parameters require an additional clarification as to where the threshold have been computed first of all. Sonnets are classified with five parameters where, a) is the vowels ratio between high-back/middle-low, b) the consonants ratio between obstruents/sonorants-fricatives, c) voicing ratio between voiced/unvoiced, d) the ratio between abstract/concrete a parameter we don't use in this evaluation, e) the sentiment/polarity parameter derived from ATF classification lexicon. Parameters are assigned a value according to threshold that we established by computing averages on the 154 sonnets. In Table 1 we show the numbers related to the threshold derived from the average values for the four parameters we consider in our analysis.

 doi:10.20944/preprints202410.0286.v1

Table 1. Ratio parameters averages and macro-averages divided by batches.

Parameters	Sonnets 1-54	Sonnets 55-105	Sonnets 106-154	Macro Averages
Vowels	0.749	0.843	0.809	0.8
Consonants	0.379	0.424	0.39	0.4
Voicing	0.645	0.653	0.653	0.65
Polarity	1.707	1.684	1.322	1.6

It is also interesting to show how the subdivision of the 154 sonnets into three separate batches allows insights into the evolution of each parameter. Values for vowels and consonants are increasing from batch one to two and then slightly decreasing, while values for polarity are steadily decreasing. Threshold are then established according to macro-averages with slight adjustments. Vowels' threshold is fixed to 0.75, consonants' threshold to 0.45, voicing threshold to 0.65, and finally polarity is assigned two thresholds: first one is established at 1.65 with higher values for positively marked and with a minimum threshold at 0.9 where lower values start for negatively marked sonnets, while the interval between the two threshold is marked as EDGE. The three sound related parameters a), b), c) may contribute singularly thus being encoded as EDGE, or together, three or two of them contributing with a similar result to the evaluation thus ending up either as OK if values are below the thresholds, or as NO is values are above.

As to our two sonnets, 47 and 38 these are their values:

Sonnet 47

a) 0.6042, b) 0.3469, c) 0.755, d) 8, e) 5.25, f) ok, g) edge

Sonnet 38

a) 0.6296, b) 0.4, c) 0.5636, d) 0.6667, e) 2.85, f) ok, g) ok

where parameter e) is higher than 1 thus indicating a majority of positive sentiment evaluations, this being in agreement with parameters a), b), all being below 0.75 for a) (vowel related), below or equal to 0.4 for b) (consonants related), but below 0.65 for c) (voicing related) in Sonnet 38, and above in Sonnet 47. The outcome of these results will contribute the final evaluation amounting to a couple of OK – which stands for full positive harmony -, in parameters f) and g) for Sonnet 38, but not fully in agreement for Sonnet 47, where EDGE – as said above - is used to indicate that one parameter in the sound grid is not within the threshold. Sonnet 7 is a fully Dis-harmonic sonnet that is projected in the main column. Here are its parameters where we see the contemporary presence of a negative and a positive parameter respectively for the sound grid and the sense polarity:

Sonnet 7

a) 1.1428, b) 0.255, c) 0.763, d) 1.57, e) 2.4, f) ok, g) no

In the second portion of the Sonnets collection that we see in Figure 3 below, with sonnets from 54 to 105, we have the same number of Positive sonnets – (7), while the additional sonnets to the left are now only (9), the number of negatives is (4) with an additional (11) sonnets to the right. The number of disharmonic sonnets is now (21). Consider again Sonnet 60 which has received a classification as EDGE-NO different from Sonnet 7 above.

Sonnet 60

a) 0.8462, b) 0.641, c) 0.949, d) 1.223, e) 1.067, f) edge, g) no

In this case two sound related parameters – b) and c) - are above thresholds thus receiving NO, while polarity parameter e) is on the EDGE.

 doi:10.20944/preprints202410.0286.v1

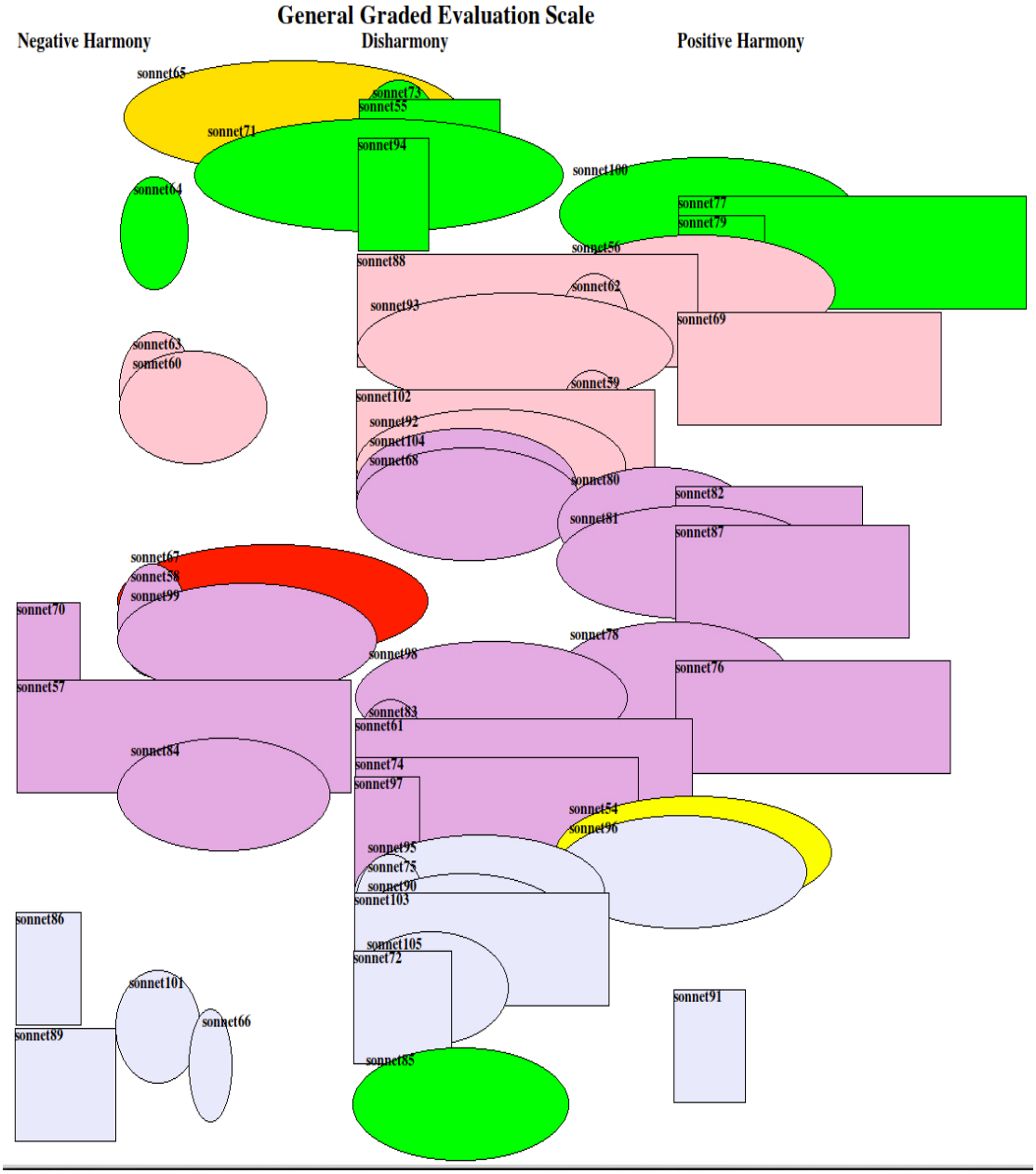


Figure 3. Multidimensional patterns for Sonnets 54-105.

Finally, in Figure 4 here below, a different result is obtained for sonnets going from 106 to 154, where this time, the majority is included in the left column – the one of Negative sonnets (7), plus those positioned on the right-hand side (19), thus reducing the number of Positive ones (4), plus those situated on its left-hand side (8). The number of disharmonic sonnets is now diminished to (11).

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 3 October 2024 doi:10.20944/preprints202410.0286.v1

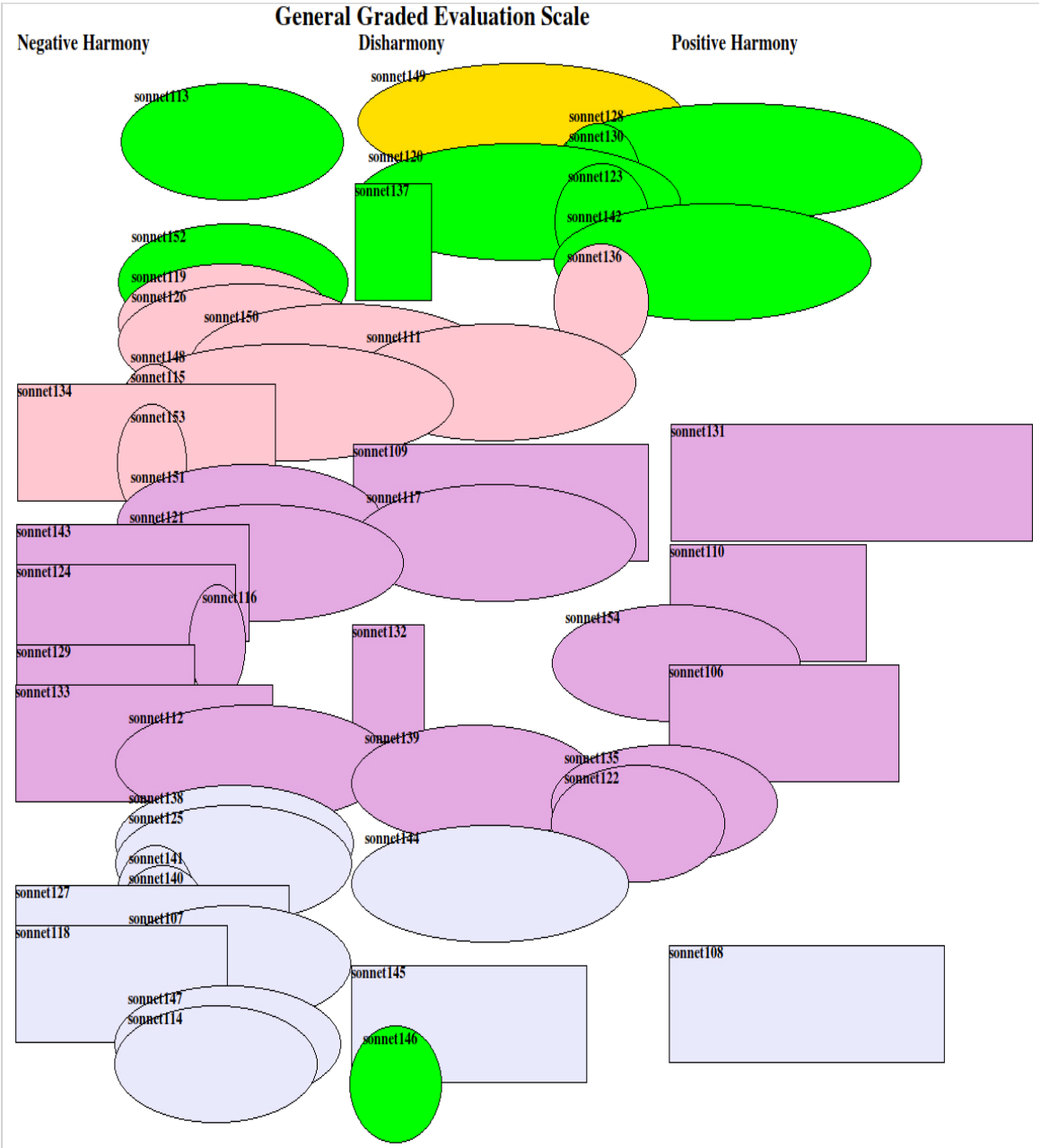


Figure 4. Multidimensional patterns for Sonnets 106-154.

In total we have 56 fully disharmonic sonnets, 53 fully and partially negatively marked, and 45 fully and partially positively marked sonnets.

As we already mentioned, Disharmony may in turn be determined by a negative or a positive sense in disagreement with its sound grid and these are the ones that appear on the right or on the left of the two main columns, and we list them below together with the sonnets that have full harmony. The number of sonnets with an exactly NEGATIVE harmony is equal to 14 and they are the following ones reported in their order of general relevance according to semantic/poetic/prosodic evaluation and divided into three batches (A-the first 53, B-the second 51, C-the third 50) corresponding to those placed on the extreme left in the three figures 2., 3., 4. presented above:

A-9 30 35

B-57 70 86 89

C-118 124 127 129 133 134 143

While the number of an exactly POSITIVE harmony becomes 18 – graded as before - and they are,

 doi:10.20944/preprints202410.0286.v1

A-16 32 37 38 39 43 48
B-69 76 77 79 82 87 91
C-106 108 110 131
and the number of sonnets with an exact DIS-HARMONY is 56 and they are :
A-1 5 7 12 14 15 17 18 19 20 21 24 25 31 33 34 41 42 44 46 48 51 52 53
B-55 63 68 72 73 74 75 83 85 88 90 92 93 94 95 97 98 102 103 104 105
C-109 111 117 120 132 137 139 144 145 146 149

In sum, 88 sonnets have been computed as fully compliant with the requirements of the ASSH : the remaining 66 are to be regarded as belonging to the class of Dis-harmonic with a bent or a bias on Positive or Negative polarity. These dis-harmonic sonnets may be subdivided into four separate sense related and biased classes:

Highly Negative 10; Negative Intermediate 8; Positive 55; Highly Positive 6

In particular, Highly Positive stands for sonnets with a ratio of Positive items superior to 2 (corresponding to the double of negatives); while Intermediate indicates sonnets with an identical count of Negative/Positive equal to 1 or below 1.2. Thus, we have 39 partially Disharmonic Negatives and 27 partially Disharmonic Positives sonnets: as a final count, only 32 sonnets have full harmony, while 122 are fully or partially disharmonic, that we show better in Table 2 below.

Table 2. Sonnets divided by batches and classification into five parameters.

Disharmony Class/Batches	Sonnets 1-53	Sonnets 54-105	Sonnets 106-154	Totals
Negatives	3	4	7	14
Disharmony Negatives	9	11	19	39
Positives	7	7	4	18
Disharmony Positives	10	9	8	27
Fully Disharmonic	24	21	11	56

According to our classification in sound-sense harmony, fully positive classification can be obtained from parameters similar to those of sonnet 82, while fully disharmonic classification can be obtained from parameters similar to those of sonnet 137, and the ones shown by sonnet 131 and 142 are regarded close to positive harmony, see below[7]:

Sonnet 82 a) 0.4314, b) 0.4167, c) 0.4898, d) 0.5, e) 3,ok,ok
Sonnet 131 a) 0.6122, b) 0.4339, c) 0.5882, d) 0.7143, e) 1.6,edge,ok
Sonnet 137 a) 0.5818, b) 0.25, c) 0.7292, d) 1.625, e) 1.0667,edge,edge
Sonnet 142 a) 0.6734, b) 0.3334, c) 0.4068, d) 0.3334, e) 1.3334,edge,ok

Where we see that while sonnet 82 has three times the number of positive elements with respect to negative ones, sonnet 131 has slightly more than half the same amount, and sonnets 137 and 142 have almost the same amount. In fact, in the whole collection of sonnets, there are only 9 sonnets whose ratios for sentiment/polarity overcomes 3 times negative ones, and 23 that have between 2 and 3 times the number of negative linguistic items. Here below we see parameters for sonnets not classified as positive but representing different cases of Disharmony, with the exception of 134 which is fully Negative.

Sonnet 130 a) 0.7059, b) 0.2903, c) 0.6226, d) 3, e) 1.5625,ok,edge
Sonnet 120 a) 0.5178, b) 0.2419, c) 0.5636, d) 0.6667, e) 0.4091,ok,no
Sonnet 126 a) 1.0606, b) 0.4893, c) 0.8095, d) 1.5, e) 1.1818,edge,no

[7] We checked our analysis mainly with Melchiori [30] and Schoenfeld [31]

 doi:10.20944/preprints202410.0286.v1

Sonnet 134 a) 0.9473, b) 0.2388, c) 0.8445, d) 0.625, e) 0.5,no,no

4. Producing a Gold Standard Annotation for Satire Interpretation

In this section we will describe the second experiment which is aimed at evaluating the ability of the ASSH to collaborate with ATF for detecting Irony, Sarcasm and Neutral sonnets. The previous data annotation process is described in two papers published lately [13,14], where we presented the modality by which we intended to produce a gold standard annotation of the sonnets to be used to gauge the ability of the Appraisal Theory Framework (ATF) to represent and tell separately irony from sarcasm, and the two from neutral class. To this aim, we collected literary critics' reviews of the sonnets and produce a short scheme with seven entries including: Sequence, Main Theme, Action, Metaphor, Neg.Eval., Pos.Eval. CONTRAST. The last entry was used as a trigger for the presence of opposite concrete or abstract concepts used by the poet to reinforce the arguments purported in the poem. We ended up with different levels of Contrast and assumed that sonnets with highest contrast could belong to the class of Sarcasm. No contrast indicating Neutral, and lower level of contrast Irony. The results of the evaluation with ATF reached 80% accuracy. This annotation has been used in this new experiment to compare with the output of the ASSH and see how well it performs. In this case, we were mainly interested in verifying whether the CONTRAST criterion could be preserved and/or partially substituted by the class of Dis-harmonic sonnets. It is important to note that Contrast was characterized in ATF terms by presence of a high level of Judgement Negative to signal Sarcasm, while presence of Negative Affect and Negative Appreciation were a signal of presence of Irony, accompanied in both cases by a significant presence of words/phrases of the corresponding Positive classes. This second experiment may end up by successful match with previous results; or by failing to achieve a good match with previous results, or else find mistakes and thus propose corrections. It is important to note that manual annotation with such a highly abstract semantic framework like the one proposed by ATF may result in presence of slight inconsistencies that will however blur the final results.

4.1. Irony and the ASSH

We are now interested in matching the output of the ASSH with a subdivision of sonnets into Irony vs Sarcasm and these two vs Neutral. To obtain this result, besides the use of general lexica with our system, we added to the ASSH the ATF lexicon produced from the xml manual annotation of the sonnets. In our previous analysis, choice was based partly on manual classification and partly on the comments made by literary critics collected and evaluated into three classes of CONTRAST – from high to low where high was related to Sarcasm. We will use the diagrams produced by the previous classification with only some slight modification. We start by commenting the classification of Irony, that we report here below in Figure 5, and matching it with the multidimensional representation presented above: from the comparison, we see that the great majority of the 49 sonnets, e.g. 37, have been classified as Disharmonic thus empirically confirming their possible match with ironically classified ones. The remaining 12 are listed below and are mainly fully Positive apart from sonnets 124, and 129. According to our previous experiment, Irony should be characterized by a low level of « contrast » which is obtained by presence of Negative Affect and absence or very low presence of Negative Judgement, together with a comparable presence of positive parameters.

Besides information of the ATF class label associated to a word or a phrase, we highlight the importance of the use of the sound-sense parameters presented before, which are now characterized by thresholds to allow the overall interpretation of the sonnet's mood in favour of one class over another. In particular, as shown above, the ASSH assigns sonnets to the class of Positives in case parameter e) is greater than 1. This use of the threshold is however insufficient to match with Irony which requires a certain comparable amount of negative elements. Positive items must be higher than the number of negative items that must be present in a comparable number, otherwise the « contrast » criterion – in this case lower than the « high » level associated with Sarcasm - will not apply. Correct classification consists of Disharmonic classified sonnets which are the majority; wrong classification regards all sonnets classified as realizing either fully Negative or fully Positive Harmony. As will

 doi:10.20944/preprints202410.0286.v1

appear in the detailed analysis below, the manual classification of Irony with ATF produced a small number of wrong results. First of all, as can be noted, correctly, Judgement Negative is present only in 25 over 49 sonnets, and in a percentage lower than 20% and in some cases lower than 10%.

As said above, the automatic classification made by the ASSH has assigned 37 sonnets to Disharmony and 7 to other classes.

Here are the parameters for the 7 wrongly matched sonnets:

Sonnet 16 a) 0.5882, b) 0.44, c) 0.3667, d) 2.3334, e) 1.6428,ok,ok
Sonnet 32 a) 0.6383, b) 0.4445, c) 0.4655, d) 1.8, e) 2.25,ok,ok
Sonnet 43 a) 0.6207, b) 0.3065, c) 0.51667, d) 2.875, e) 4,ok,ok
Sonnet 45 a) 0.7272, b) 0.4468, c) 0.556, d) 1.7, e) 4,ok,ok
Sonnet 87 a) 0.5192, b) 0.4, c) 0.4445, d) 1.75, e) 5.25,ok,ok
Sonnet 91 a) 0.6808, b) 0.35, c) 0.4528, d) 1.1667, e) 2,ok,ok
Sonnet 124 a) 0.75, b) 0.5116, c) 0.9736, d) 3, e) 0.8421,no,no
Sonnet 129 a) 0.7317, b) 0.6757, c) 0.9412, d) 0.2, e) 0.875,no,no

Looking at the e) parameter, the negative harmony of sonnets 124 and 129 is due to presence of sound negative parameters in b), and c) being above threshold: now, if we consider the value of the e) parameter, a low level of contrast would however be partially guaranteed, the proportion indicated by the ratio being negative but very close to the threshold – 0.9 – to become EDGE, and thus still comparable. In addition, the ATF classification shows a high level of AffectNegative and a very low level of JudgementNegative, together with Positive classes. The remaining 5 wrong associations are all positive. But, if we check the validity of parameters considered as positive, we see that sonnet 16 has parameter e) very close to the threshold – 0.65 – to become POSITIVE and parameter b) close to the threshold to become NEGATIVE, which in both cases makes a low contrast. The same applies to sonnet 32 with respect to parameter b), very close to the threshold.

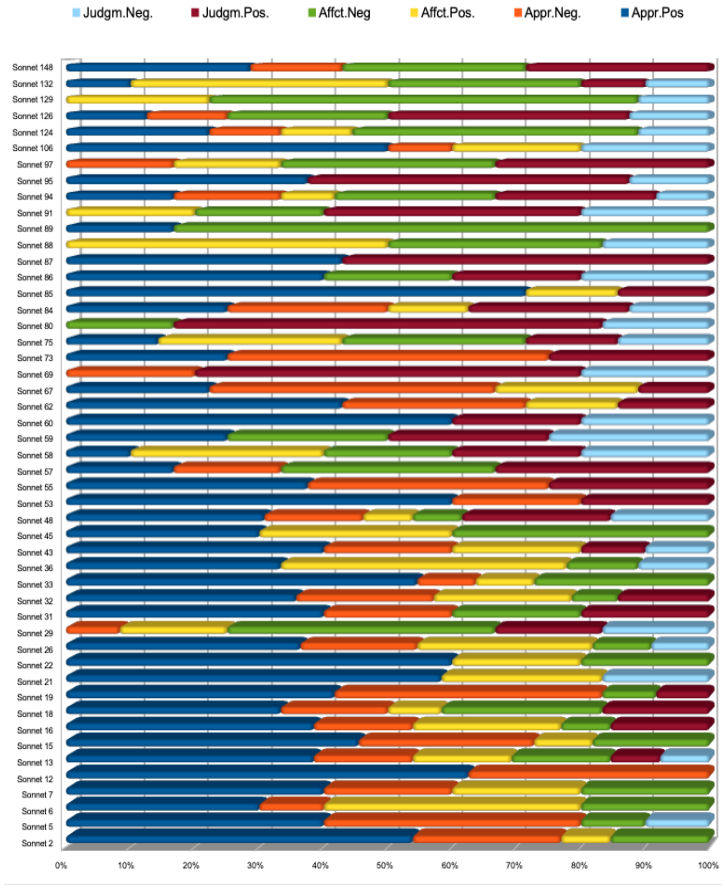


Figure 5. Distribution of ATF classes in 49 sonnets according to manual evaluation for Irony classification.

 doi:10.20944/preprints202410.0286.v1

However, sonnet 91 has the Judgement Negative class higher than 20%, which added to Affect Negative contributes a reasonable level of contrast. Thus, a first evaluation with 7 wrong matches over 49 corresponds to 83.7% of accuracy. However, if we apply the slight corrections of the thresholds to parameters indicated above, we have only 3 completely wrong matches – sonnets 43, 45, and 87, which makes accuracy go up to 94%.

4.2. Sarcasm and the ASSH

Let's now look at those sonnets that in previous classification were identified as Sarcastic. Here below the diagram in Figure 6. As said previously, we expect sonnets classified as Sarcastic to have a majority of Negative Judgement annotations, which in our case amounts to saying that these sonnets should not be listed under Positive Harmony but rather either under Negative or simply as Disharmonic. As can be easily noted, the amount of Negative Judgement is remarkable and always higher than 20% with only one exception, sonnet 153 which however is counterbalanced by highly negative sound parameters as shown below:

Sonnet 153 a) 0.7778, b) 0.5112, c) 0.7442, d) 2.2223, e) 1.3077,edge,no

If Contrast with polarity parameter e) is not present in a sufficient amount, Contrast may be guaranteed as long as also sound parameters are matched correctly. The classification by the ASSH has assigned 34 sonnets to the class of Disharmony. However, as happened with Irony classification, there are a small number of sonnets – only 8 - that received a wrong class, where by "wrong" classification we identify at first the couples: NO NO and OK OK. This is because – we repeat – in this way "contrast" is not guaranteed, but it may be present in the ATF classification.

Here below the 8 wrong matches:

Sonnet 9 a) 0.8085, b) 0.2414, c) 0.6667, d) 2, e) 0.6667,no,no
Sonnet 35 a) 1, b) 0.2, c) 0.7619, d) 2, e) 0.619,no,no
Sonnet 70 a) 0.5334, b) 0.5897, c) 0.8, d) 0.3, e) 0.9473,no,no
Sonnet 82 a) 0.4314, b) 0.4167, c) 0.4898, d) 0.5, e) 3,ok,ok
Sonnet 131 a) 0.6122, b) 0.4339, c) 0.5882, d) 0.7143, e) 1.6,ok,ok
Sonnet 133 a) 1.075, b) 0.4, c) 0.7551, d) 3.5, e) 0.3,no,no
Sonnet 134 a) 0.9474, b) 0.2388, c) 0.8445, d) 0.625, e) 0.5,no,no
Sonnet 143 a) 0.6981, b) 0.6939, c) 1.0223, d) 1.75, e) 0.6667,no,no

We consider at first how close parameters of these eight sonnets are to related restrictions for being computed as OK-OK or NO-NO. Regarding the first case, the two sonnets involved are sonnet 82 and 131: only 82 is fully compliant with the evaluation, while 131 has the e) parameter on the EDGE. If we look at the ATF parameters for 131, they show a perfect balance which however means a high level of contrast with JudgementNegative; sonnet 82 on the contrary, has a majority of positive classes and little JudgementNegative, thus a mistaken manual annotation and no improvement with the lexical automatic classification. As to the negatives, we see that only three sonnets have all parameter values correctly positioned over the threshold we established for being regarded fully negative: they are, sonnet 133, 134, 143, but the ATF evaluation shows a remarkable contrast with sufficient presence of JudgementNegative. The remaining 3 sonnet – 9, 35 and 70 - might as well be computed with a weak form of Dis-harmony: sonnet 70 has parameter e) very close to the threshold 0.95 to become EDGE, in addition "contrast" is reflected in the ATF parameters; sonnet 9 has parameter c) voicing, very close to the threshold 0.65 to become EDGE.

In addition to the eight sonnets above, Sarcasm classification contains also five sonnets which have the positively marked couple OK EDGE, and they are:

Sonnet 4 a) 0.7112, b) 0.4259, c) 0.3667, d) 0.5385, e) 1.2857,edge,ok
Sonnet 10 a) 0.3684, b) 0.3167, c) 0.4909, d) 0.3, e) 1.2143,edge,ok
Sonnet 106 a) 0.5625, b) 0.575, c) 0.5814, d) 1.2, e) 1.7778,edge,ok
Sonnet 142 a) 0.6735, b) 0.3334, c) 0.4068, d) 0.3334, e) 1.3334,edge,ok
Sonnet 154 a) 0.7112, b) 0.3585, c) 0.8334, d) 2.1428, e) 3,edge,ok

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 3 October 2024 doi:10.20944/preprints202410.0286.v1

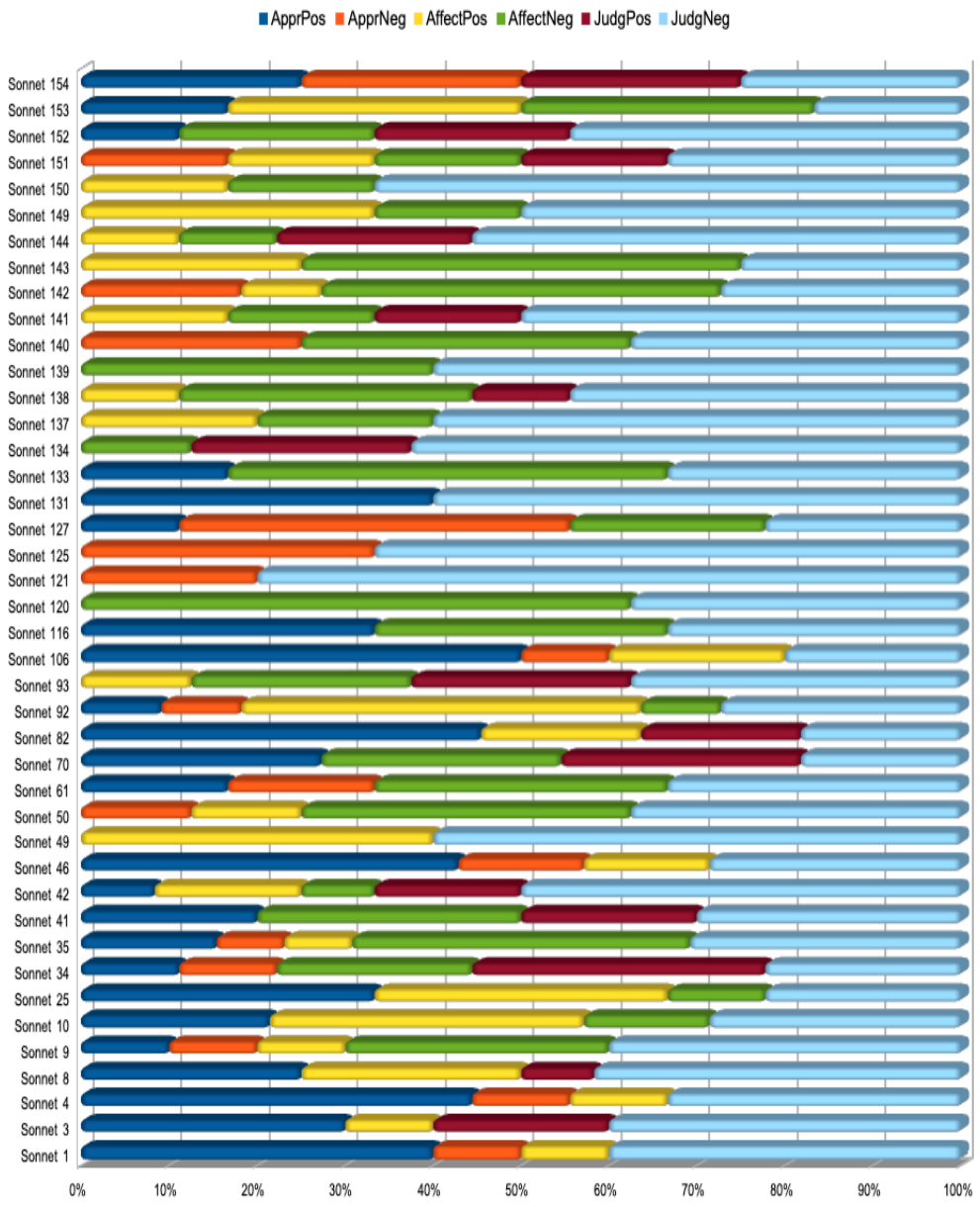


Figure 6. Distribution of ATF classes for 42 sonnets according to manual evaluation for Sarcasm classification.

If we exclude sonnet 154, the four remaining sonnets have all positive sound and edge value associated with polarity; 154 on the contrary, has positive polarity and sound classified as edge. Now if we look at the corresponding ATF classification, we can see that the only sonnet that counterbalances the positive sound is sonnet 142 where Judgement Negative is fairly highly represented. As to the other four, there seems to be no corresponding contrast in the ATF classification so we must regard them wrongly manually classified as Sarcastic: but the newly lexically based classification modifies drastically the previous one with a sufficient presence of Positives and Negatives and in particular with JudgementNegative, that in the case of Sonnet 10 is higher than any other class Positive and Negative.

So eventually, ASSH match with Sarcasm classification is on a first count 29 correct vs 13 wrong resulting in 69% accuracy, with further refinement we rescued five sonnets, and they are 9, 35, 70, 131, 142. Thus the preliminary result is 34 over 42 corresponding to 81% accuracy. If then we correct the manual ATF classification with the lexically implemented one, excluding sonnet 82, we end up with 41 over 42 correct classification and 97.6% accuracy.

 doi:10.20944/preprints202410.0286.v1

4.3. Neutral Sonnets and the ASSH

Coming now to classifications made for Neutral sonnets, we realized that this class constitutes the strongest motivation for equipping the Sound-Sense Harmony with the lexicon provided by the manual annotation with ATF. Experiments using only polarity showed that the ASSH was inherently unable to tell apart the three classes we required: Irony, Sarcasm and Neutral. Sentiment analysis itself is insufficient for such a fine-grained classification. The results are presented here below, but at first Figure 7 shows the diagram obtained from manual classification.

As can be easily noticed, neutral sonnets are characterized by an almost total absence of JudgementNegative and a majority of Positively marked classes. Disharmonic classification may appear but they should avoid having a strong "contrast", as represented by the couple OK NO. In general, this is what has happened with ASSH classification, with some exceptions though. We report here below the sonnet numbers preceded by a double star to indicate a major error:

NEGATIVE

147 146 119 118 115 114 113 112 101 99 90 71 66 65 64 63 40 30 28

DISHARMONY

145 **117 **111 109 107 **105 **104 103 102 **98 **83 72 **68 51 44 **24 **17 **14

POSITIVE-36

136 135 128 123 122 110 108 100 96 81 79 78 77 76 56 54 47 39 38 37 27 23 20 11

Ten sonnets have been classified by ASSH as having strong contrast, OK-NO. The remaining sonnets have received classifications containing EDGE accompanied either by NO, OK or again EDGE all representing low cases of contrast. If we look at the ATF classification only two sonnets have some form of contrast with Positive and Negative classes, and they are sonnet 68 and sonnet 17. All the remaining sonnets, i.e. 14, 24, 83, 98, 104, 105, 111, and 117 have sometimes only positive labels or a marked majority of them. Reconsidering the output of the ASSH we can see that sound related parameter are sometimes very close to the threshold to become EDGE. Here they are:

Sonnet 14 a) 0.6667, b) 0.7187, c) 0.9706, d) 2, e) 2.375,ok,no
Sonnet 17 a) 0.6316, b) 0.4615, c) 0.7659, d) 0.6667, e) 1.8,ok,no
Sonnet 24 a) 0.7659, b) 0.4339, c) 0.8085, d) 5.25, e) 6,ok,no
Sonnet 68 a) 1.1945, b) 0.5, c) 0.575, d) 1.4445, e) 2.875,ok,no
Sonnet 83 a) 1.2162, b) 0.5952, c) 0.6038, d) 0.7778, e) 1.9,ok,no
Sonnet 98 a) 1.1282, b) 0.4347, c) 0.6956, d) 0.9167, e) 2.2223,ok,no
Sonnet 104 a) 1.4359, b) 0.4528, c) 1.6667, d) 1.0834, e) 2.1667,ok,no
Sonnet 105 a) 0.8, b) 0.6551, c) 1.0312, d) 0.25, e) 3,ok,no
Sonnet 111 a) 0.9474, b) 0.4584, c) 0.4528, d) 0.6667, e) 1.6428,ok,no
Sonnet 117 a) 0.9512, b) 0.4878, c) 0.5531, d) 0.6667, e) 2.5,ok,no

This applies to sonnets 17, 24, 68, 111, 117. Summing both ATF values with slightly corrected ASSH values we come up with 4 totally wrong and 6 sonnets that may still be regarded as Neutral. Apart from these 10 mistakes, the remaining 53 sonnets have all been classified in accordance with criteria established for NEUTRALITY, which however would not be fully transparent for ASSH classification. Contrary to what happened with Irony and Sarcasm, in this case no or low CONTRAST should be present in the sonnet classification.

 doi:10.20944/preprints202410.0286.v1

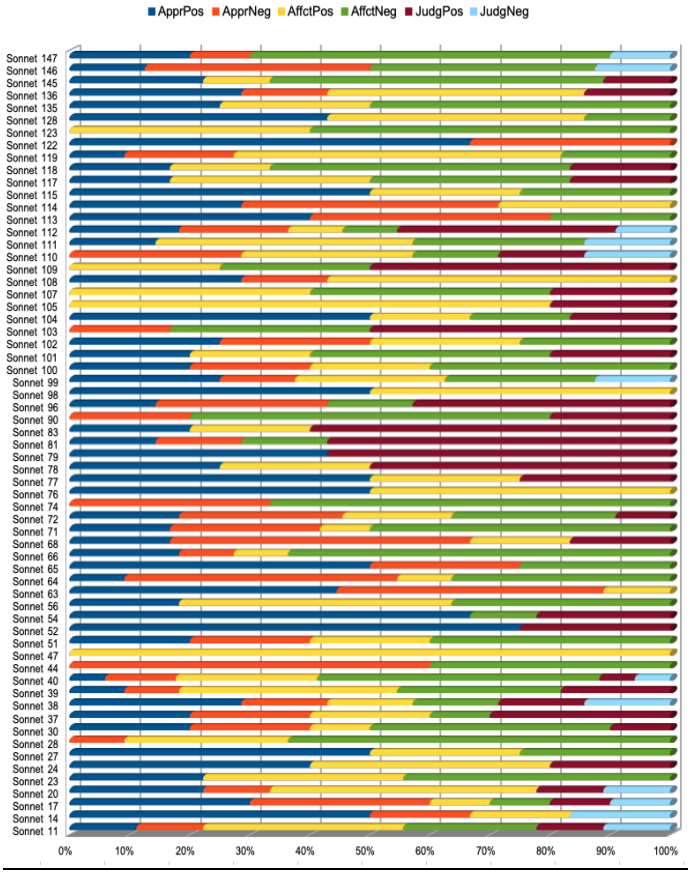


Figure 7. Distribution of 63 Sonnets evaluated as Neutral from manual ATF annotation.

As regards ATF classification, general criteria for Neutrality are then the following ones: a. Majority of Positively Marked Sonnets: High level of Judgement Positive, High Level of Appraisal Positive, High Level of Affect Positive; Negative categories most represented: Affect and Appraisal. As can be easily see from Figure 7, the level of Judgement Negative in the 12 sonnets containing it, is well below 20%. The contribution of the sound grid of the poems contained in this partition is very important because it confirms the semantics. This result amounts to 84% accuracy, however if we consider that 6 sonnets over 10 have fully or highly Positive ATF classification, the number of wrong sonnets falls down to only four, with a final result of 93.6% accuracy.

Eventually, we have an average overall accuracy – considering Irony, Sarcasm and Neutrality – of 94.6%, where Neutrality has been the hardest class to approximate.

5. Discussion

The contribution of this paper is to present a new paradigm by means of which to assess poetic content, that we called Sound-Sense Harmony. The algorithm that produces the evaluation of each poem is based on the output of a symbolic system for text analysis called *SPARSAR*, mainly used for automatic poetry reading, which has been enriched with the graphical output that shows a multidimensional representation of the contents of a poem or a collection of poems, thus allowing a visual immediate comparison. In addition, the paper shows that deep semantic processing like the one contained in the ATF can be made to compare with poetic sound-based devices to produce a more comprehensive analysis. In sum, the paper features two experiments: the first one uses directly the output of the ASSH to produce the multidimensional representation; the second one, adds the information coming from the lexicon containing the result of the manual annotation of the sonnets with ATF, and tries to use the sound-sense harmony with the task of telling apart Sarcasm from Irony and these two from Neutral.

 doi:10.20944/preprints202410.0286.v1

More importantly the experiments documented in this paper are all carried out with a symbolic system *SPARSAR*, developed in the last 20 years or so and fully evaluated (but see Delmonte [1] for more information). Deep Neural Networks and Large Language Models are unable to carry out complex tasks like the one proposed here: first of all, they cannot produce a correct and complete phonetic transcription of the text contained in the 154 sonnets, whereby the sound is derived to be subsequently matched with the semantic annotation. But here again there is no possibility for DNNs and LLMs to annotate automatically the sonnets using Appraisal Theory Framework (hence ATF) schemas. For these reasons we stick to the symbolic system, but are unable to make comparisons with other systems for the difficulty of the task. Besides, to our knowledge, there is no previous attempt at combining phonetic information with ATF semantic classification, and in general with semantic classification, to produce a more comprehensive interpretation of the content of a poem. We also used the automatic lexically-driven ATF classification to correct some slight inconsistencies resulted from the manual annotation showing that using a simple matching algorithm at word and phrase level on the basis of a well-wrought lexicon can produce improvements on manual annotation when the task is semantically very heavy and hard to apply consistently.

6. Conclusion

In this article we presented experimental results of the sound-sense harmony obtained and shown by the ASSH, i.e. a multidimensional representation of the poetic and linguistic content of Shakespeare's Sonnets. The algorithm computes the representation on the basis of an extended number of parameters encompassing poetic devices, prosodic features, linguistic lexical, syntactic features, semantic and conceptual features. The final output is a ranking of each sonnet on a graded scale organized into three separate columns where sonnets may take two possible shapes: a square or an ellipse of varying size. Automatic results have then been validated by comparison with a manual annotation of the sonnets done in order to classify them into three different types: ironic, sarcastic, neutral. Matching was aimed at checking the feature "CONTRAST" and its correspondence with Dis-harmony: it has been fairly successful obtaining an overall 94.6% accuracy. Computing Sound-Sense Harmony with the help of the fine-grained classification framework made available by ATF has thus shown that poetic devices are not disjoint from the meaning the poem intends to convey: in sum, poetic style requires a unity of intents in the choice of words, both sound and meaning should be taken into account.

Supplementary Materials: The following supporting information can be downloaded at the website of this paper posted on Preprints.org,

Author Contributions: The ATF classification task has been carried out partly by Nicolò Busetto, co-author of a number of papers describing the work done.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: We make available data of the complete analysis of the 154 sonnets as supplementary material.

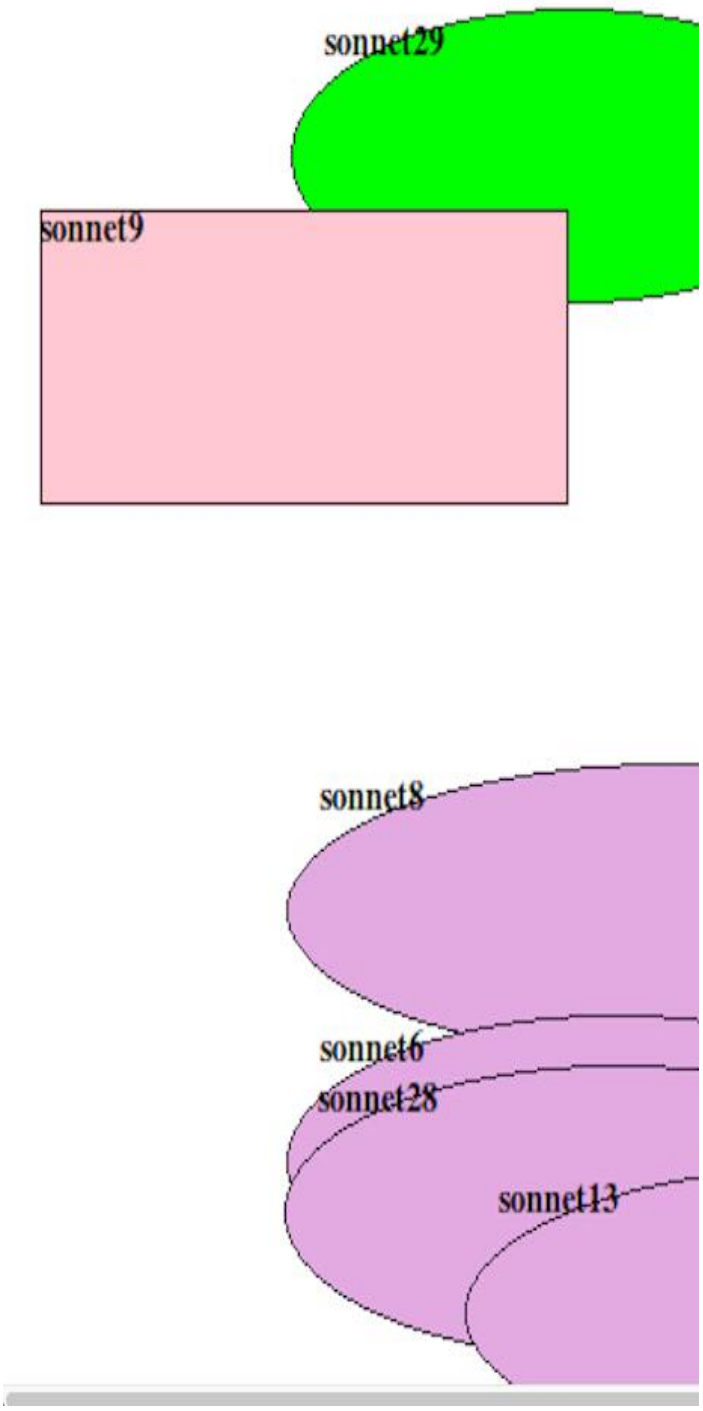
Conflicts of Interest: The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript; or in the decision to publish the results.

Appendix A

This appendix contains the enlarged version of the Figures 2,3,4, inserted in the text above, where multidimensional sound-sense representation was shown. We divided the 154 sonnets into three batches to allow for the representation details to be visible but sonnet numbers still appear too small to be read easily. Here below we show enlarged versions of the same figures by dividing each figure into four subsections: a low section containing the less relevant sonnets with three columns Negative-Disharmonic-Positive, then a separate figure for each column.

 doi:10.20944/preprints202410.0286.v1

Negative Harmony



 doi:10.20944/preprints202410.0286.v1

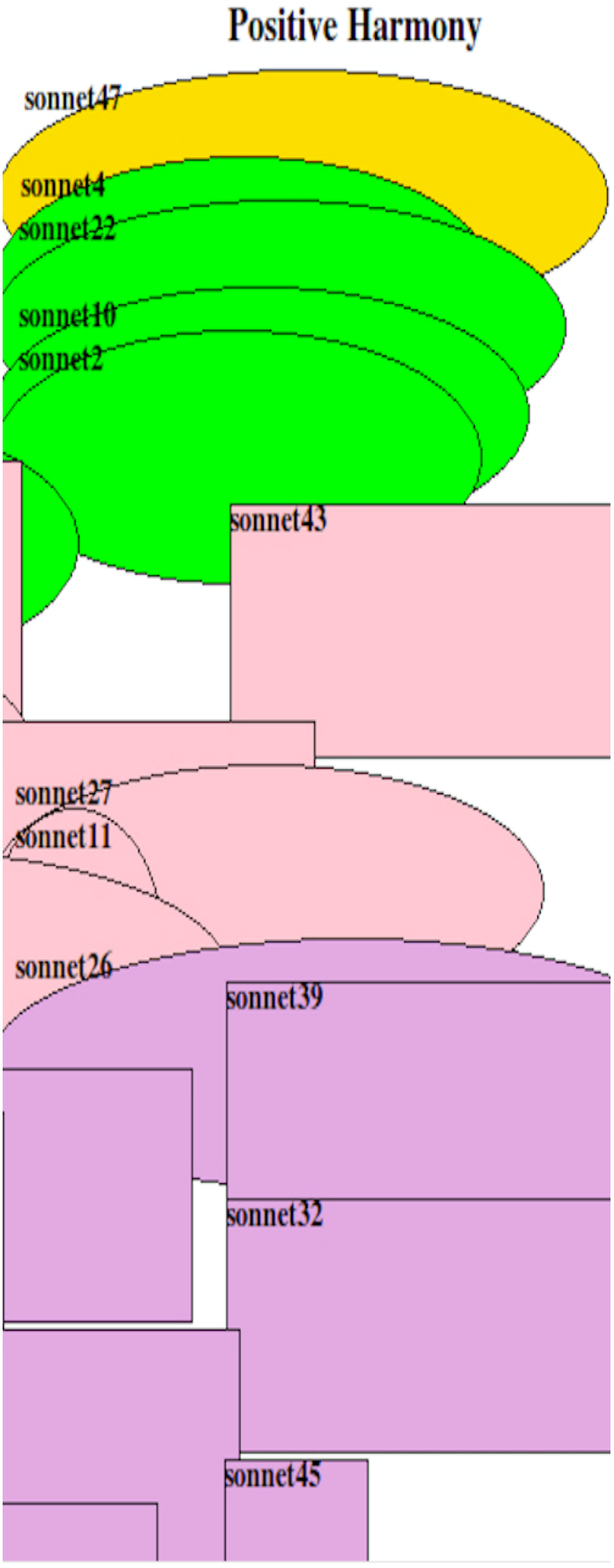


Figure A1. A2. Upper part of multidimensional representation for sonnets 1-53 with Negative and Positive harmony.

 doi:10.20944/preprints202410.0286.v1

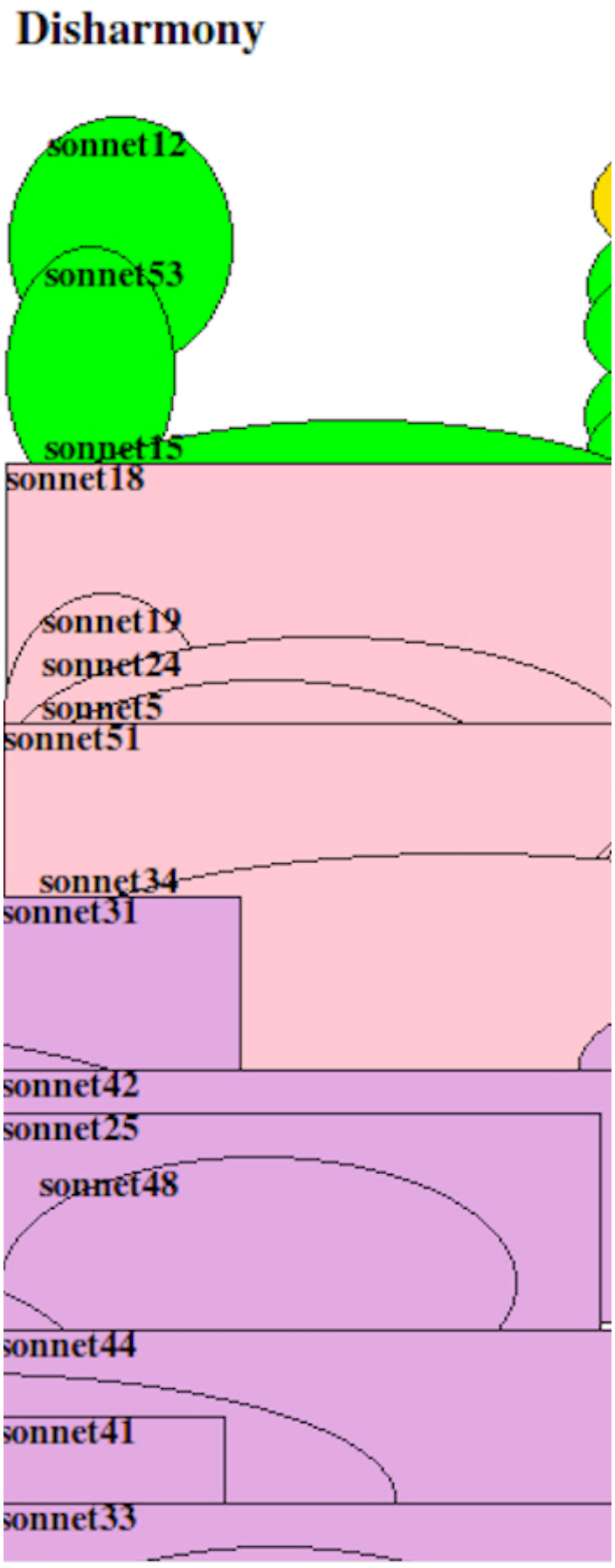


Figure A3. Upper part multidimensional representation for sonnets 1-53 with Disharmonic sonnets.

 doi:10.20944/preprints202410.0286.v1

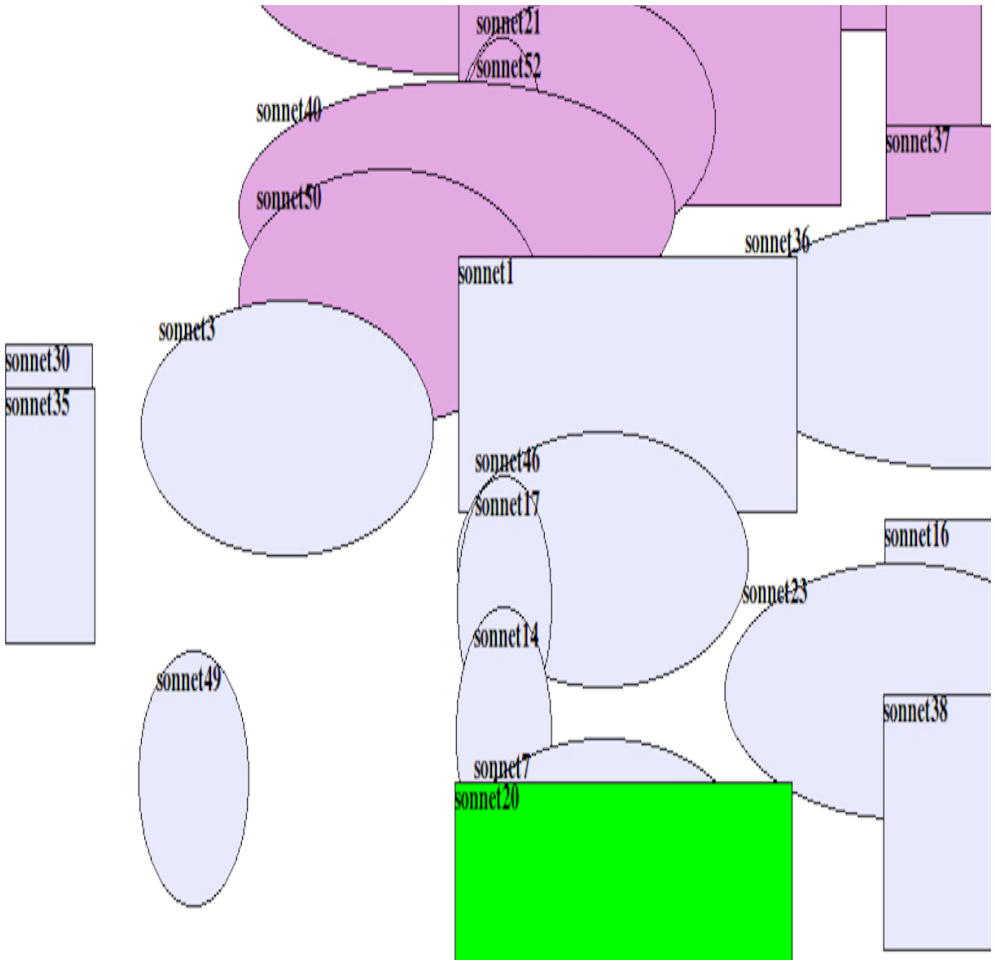
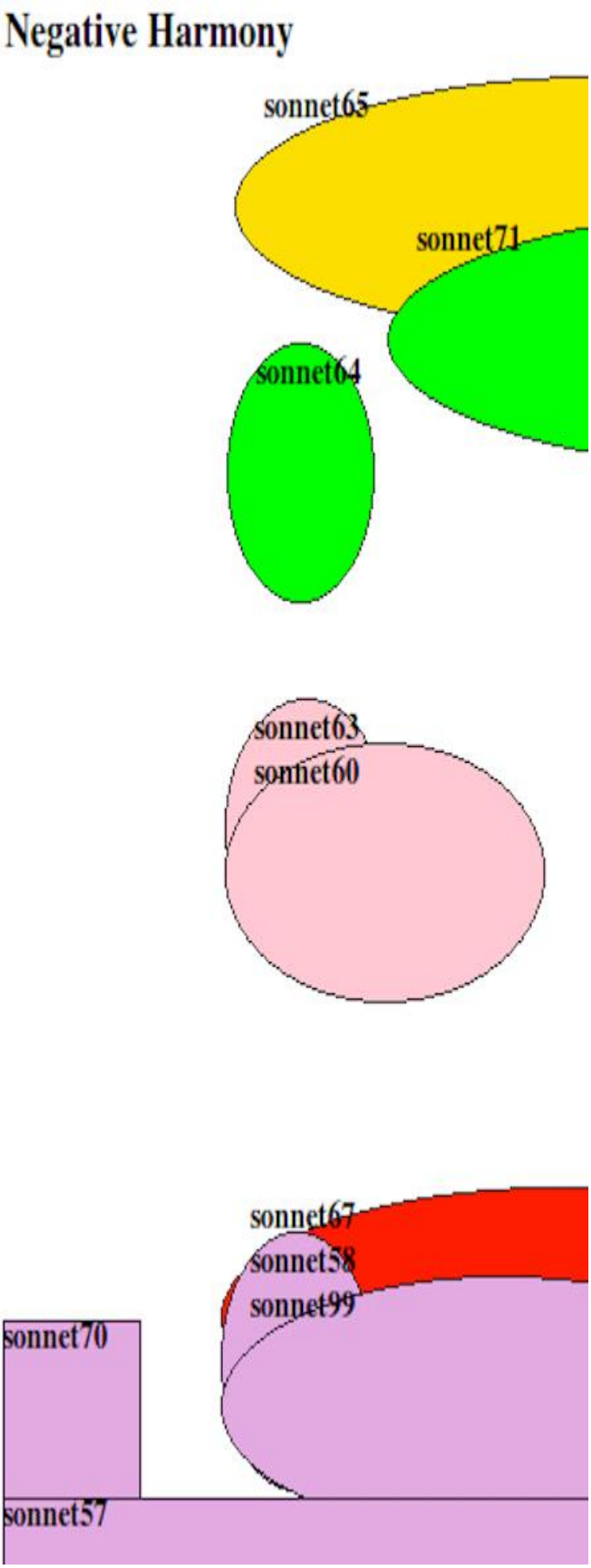


Figure A4. Lower part of multidimensional representation for sonnets 1-53 with three columns.

 doi:10.20944/preprints202410.0286.v1



 doi:10.20944/preprints202410.0286.v1

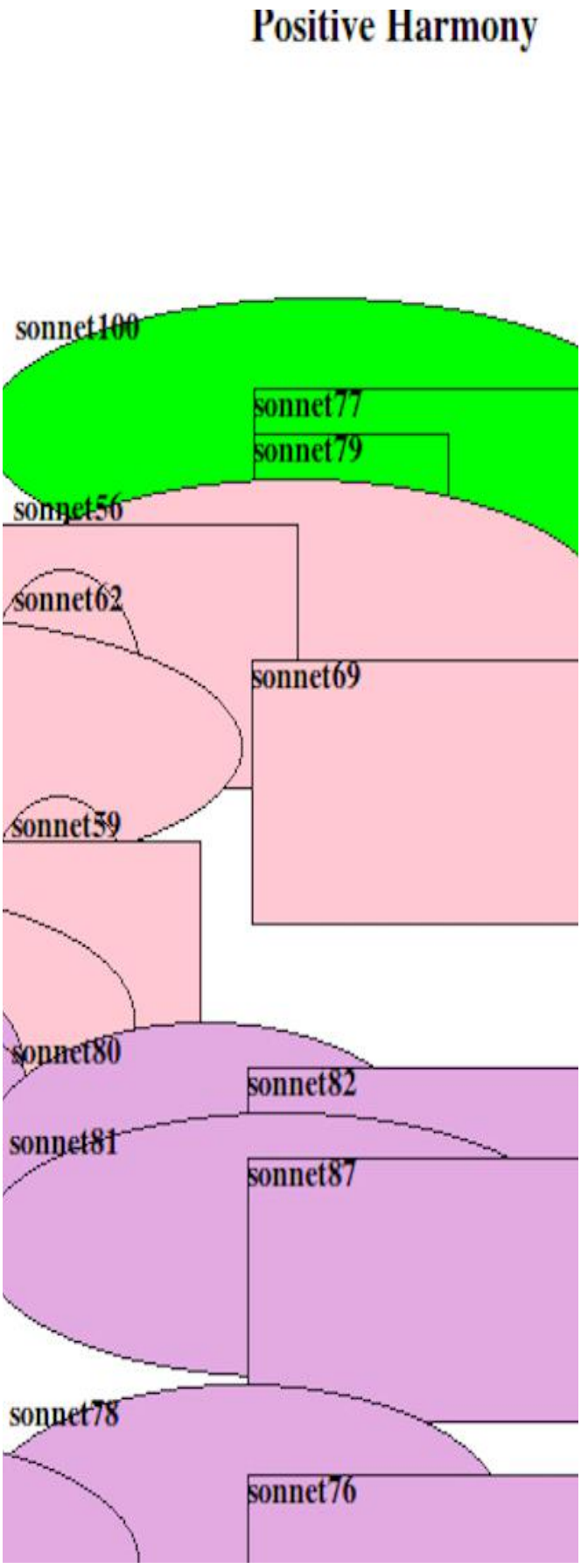


Figure A5. A6. Upper part of multidimensional representation for sonnets 54-105 with Negatives and Positives.

 doi:10.20944/preprints202410.0286.v1

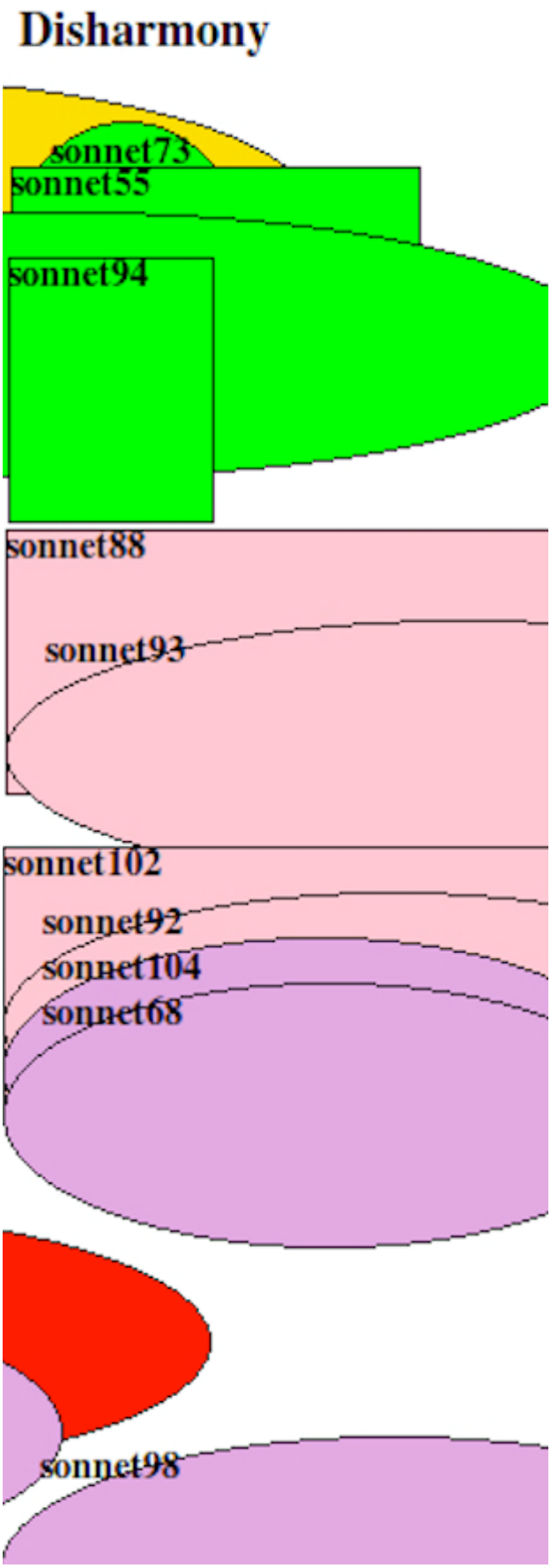


Figure A7. Upper part of multidimensional representations for sonnets 54-105 with Disharmony.

 doi:10.20944/preprints202410.0286.v1

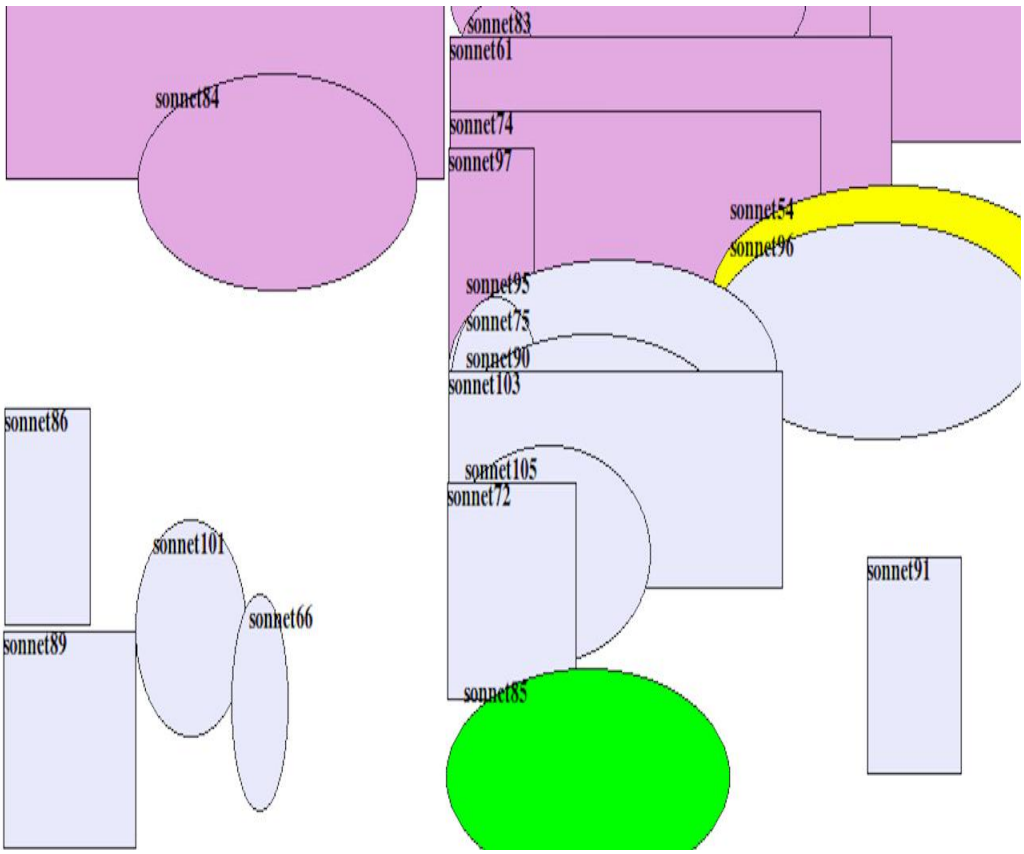
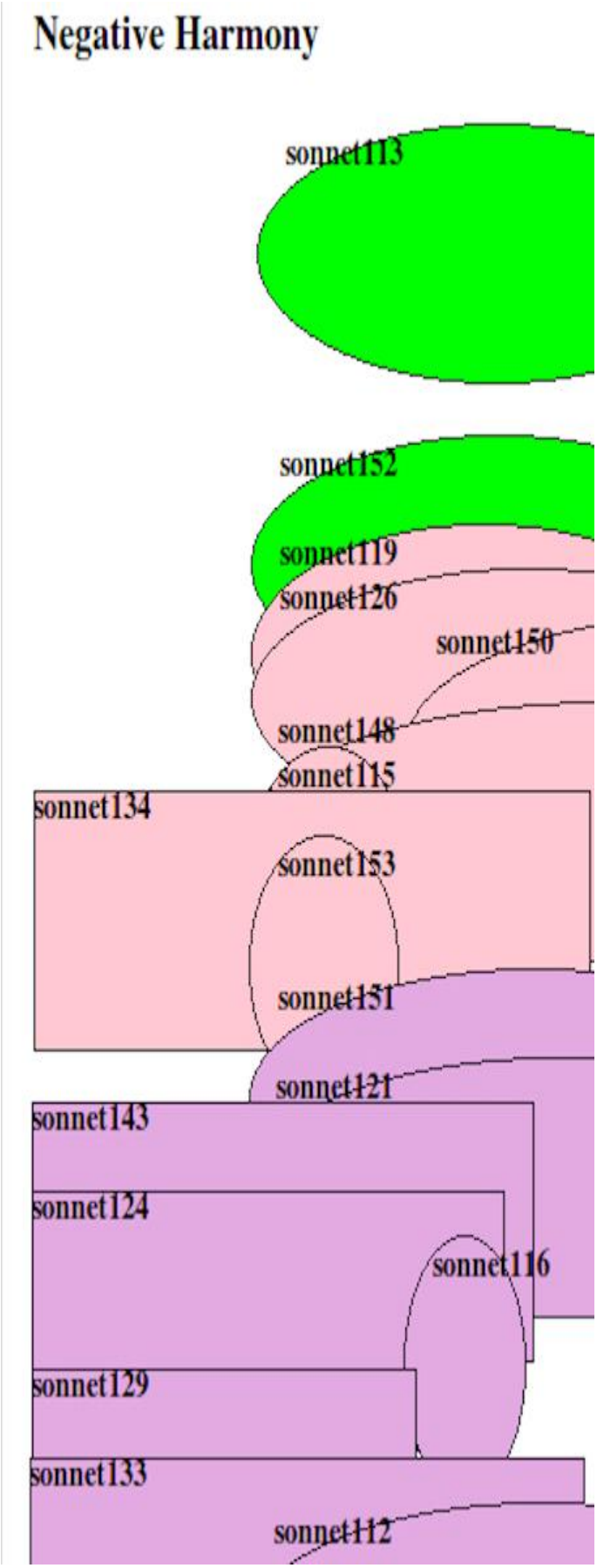


Figure A8. Lower part of multidimensional representation for sonnets 54-105 with three columns**.**

 doi:10.20944/preprints202410.0286.v1



 doi:10.20944/preprints202410.0286.v1

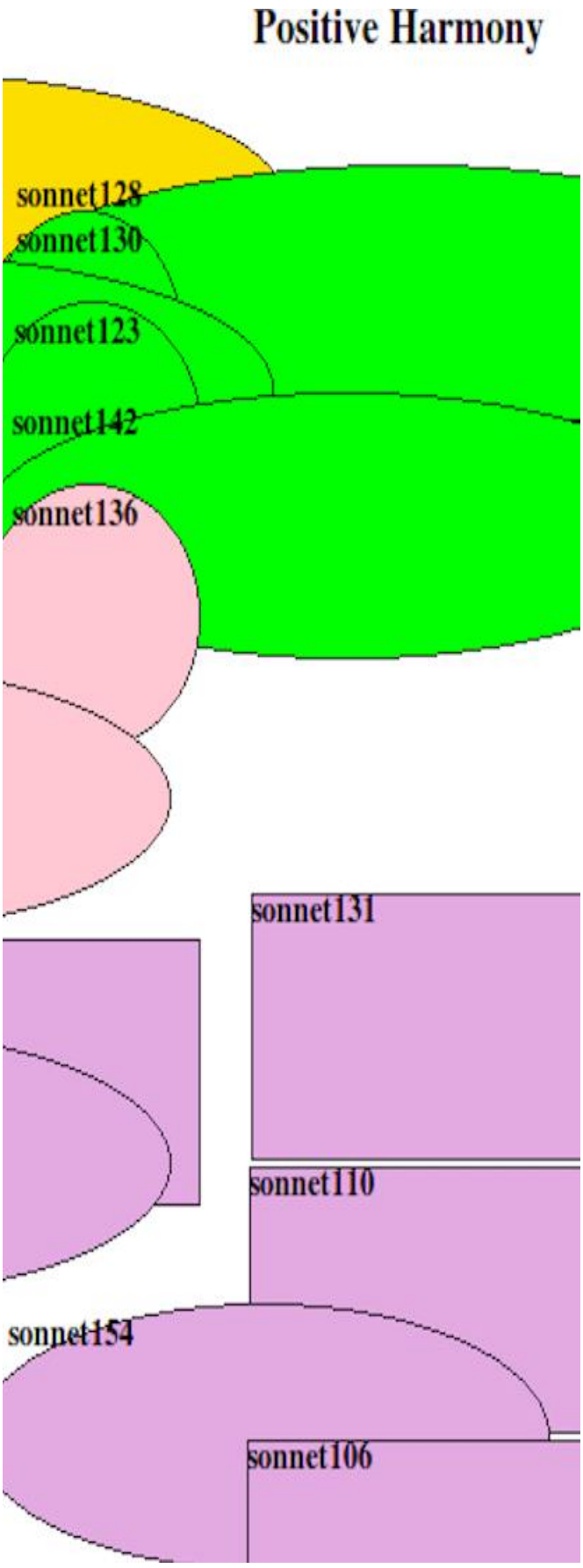


Figure A9/A10. Upper part of multidimensional representation for sonnets 106-154 with Negatives and Positives

 doi:10.20944/preprints202410.0286.v1

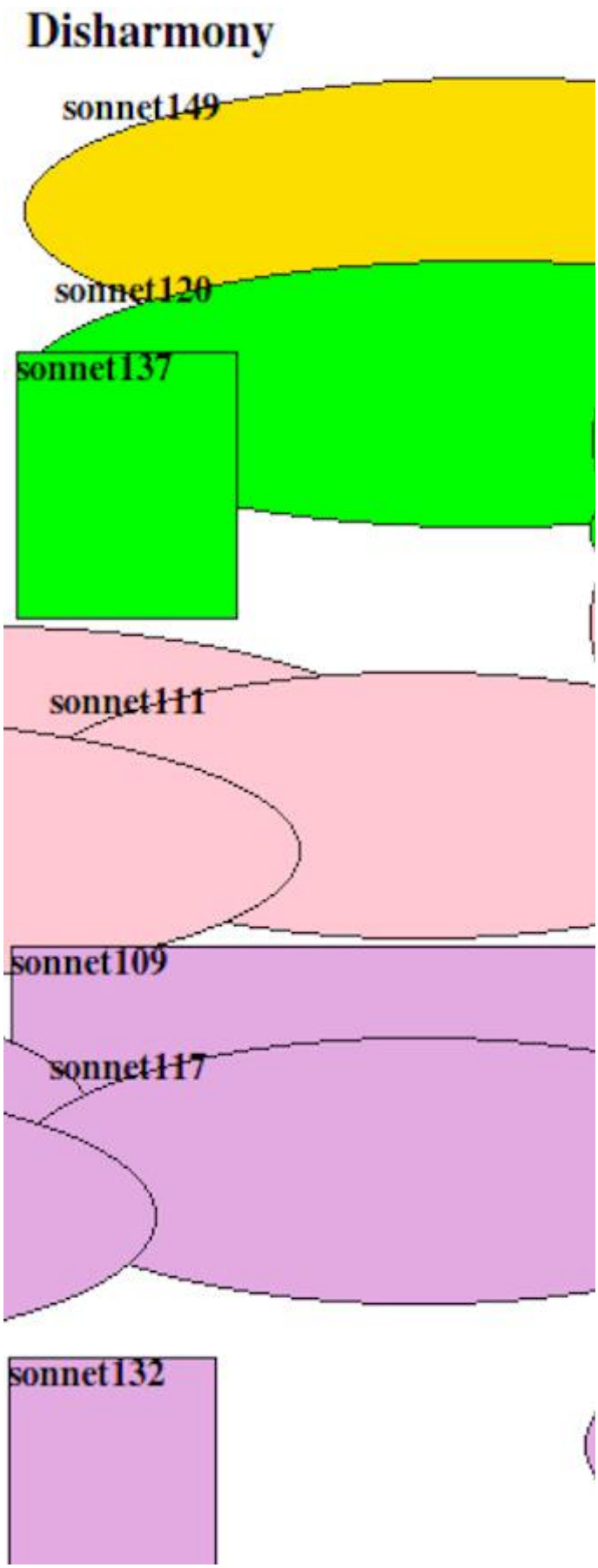


Figure A11. Upper part of multidimensional representation for sonnets 106-154 with Disharmony.

 doi:10.20944/preprints202410.0286.v1

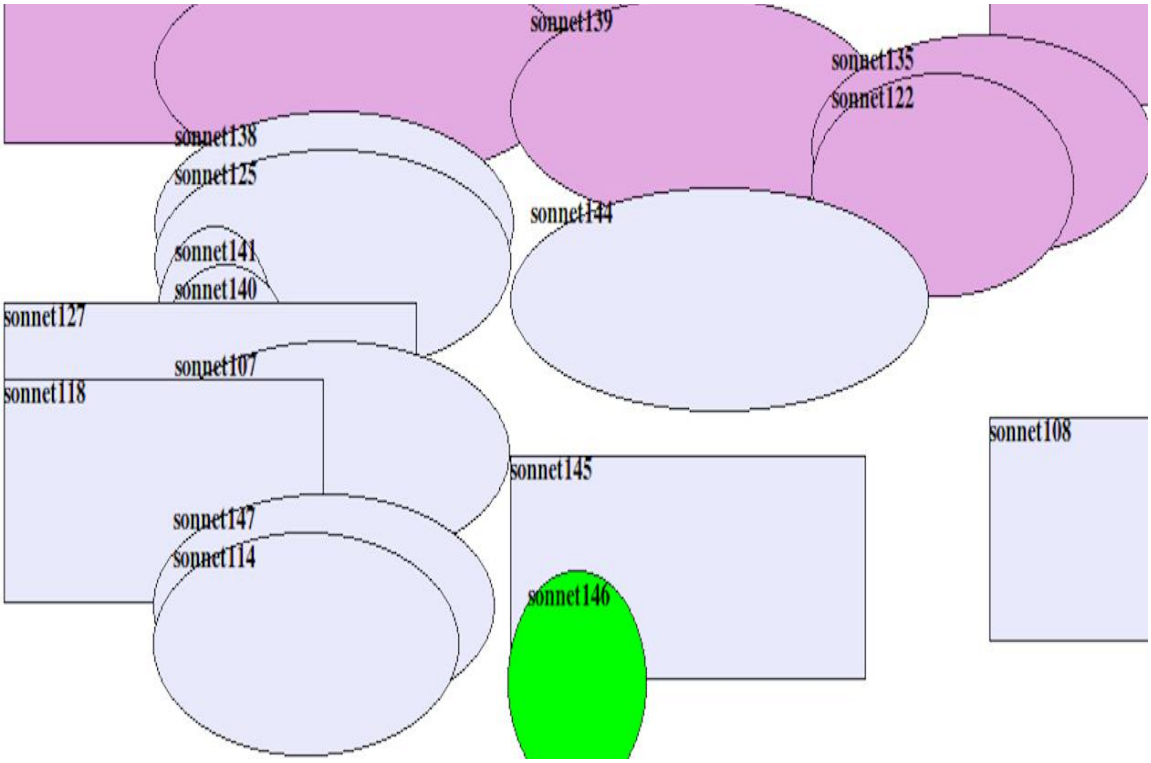


Figure A12. Lower part of multidimensional representation for sonnets 106-154 with three columns**.**

References

1. Rodolfo Delmonte. Computing the Sound-Sense Harmony - a Case Study of William Shakespeare's Sonnets and Francis Webb's Most Popular Poems, *Information,* 2023; 14(10):576, https://doi.org/10.3390/info14100576.
2. Rimbaud, Arthur, 1871. Voyelles, 38th poem of the collection Poésies, Lutèce, Paris, 1883.
3. Jakobson, R. Six lectures on sound and meaning (Trans.: J. Mepham). Cambridge: MIT Press; 1978, (Original work published in 1976).
4. Jakobson, R., & Waugh, L. The sound shape of language. Bloomington: Indiana University Press; 1978
5. Mazzeo, M. Les voyelles colorées: Saussure et la synesthésie. Cahiers Ferdinand de Saussure; 2004; 57, pp. 129–143.
6. Fónagy, Iván The Functions of Vocal Style", in Seymour Chatman (ed.), Literary Style: A Symposium. London: Oxford UP; 1971, pp. 159-174.
7. Mallarmé, Stéphane, 1897, Crise de Vers, in Eugene Pasquelle, Divagations, pp. 235-251.
8. Macdermott M.M. Vowel Sounds in Poetry: Their Music and Tone Colour, Psyche Monographs, No.13, London: Kegan Paul; 1940.
9. Tsur, Reuven. What Makes Sound Patterns Expressive: The Poetic Mode of Speech-Perception. Durham N. C.: Duke UP; 1992.
10. Tsur Reuven, Poetic Rhythm: Structure and Performance: An Empirical Study in Cognitive Poetics; Sussex Academic Press; 2012, pp. 472.
11. Genette, Gérard, 1976. Mymologiques, Voyage en Cratylie, Le Seuil, Paris. Mymologics – English Edition, 1995, University of Nebraska.
12. Bolinger, Dwight, 1949, The sign is not arbitrary. THESAURUS, Tomo V., No. 1,2,3, pag.52-62.

 doi:10.20944/preprints202410.0286.v1

13. Busetto N.,Delmonte R., (2019b), Annotating Shakespeare's Sonnets with Appraisal Theory to Detect Irony, in Proceedings of the Sixth Italian Conference on Computational Linguistic, Bari, Italy, November 13-15, downloadable at http://ceur-ws.org/Vol-2481/paper11.pdf
14. Delmonte, R., & Busetto, N. (2019a), Detecting irony in Shakespeare's sonnets with SPARSAR. In Proceedings of the sixth italian conference on computational linguistics; 2019. Available online: https://dblp.org/rec/bib/conf/clic-it/DelmonteB19.
15. J. Martin and P.R. White. Language of Evaluation, Appraisal in English. Palgrave Macmillan, London and New York; 2005.
16. M. Taboada and J. Grieve. Analyzing appraisal automatically. In Proceedings of the AAAI Spring Symposium on Exploring Attitude and Affect in Text: Theories and Applications; AAAI Press; 2004, pp. 158–161.
17. David K. Weiser. (1983). The 'sonnets'. 84(4):456–469. Shakespearean irony: Neuphilologische Mitteilungen,
18. David K. Weiser. (1987). Mind in Character – Shake- speare's Speaker in the Sonnets. The University of Missouri Press.
19. Burfoot, C., & Baldwin, T. (2009). Automatic satire detection: are you having a laugh? Proceedings of the ACL-IJCNLP 2009 Conference Short Papers (pp. 161--164). Suntec, Singapore: Association for Computational Linguistics.
20. Carvalho, P., Sarmento, L., Silva, M., & de Oliveira, E. (2009). Clues for detecting irony in user- generated contents: oh...!! it's so easy;-). Proceeding of the 1st international CIKM workshop on Topic-sentiment analysis for mass opinion (pp. 53-56). Hong Kong: ACM.
21. Reyes, A., & Rosso, P. (2011). Mining subjective knowledge from customer reviews: a specific case of irony detection. WASSA '11 Proceedings of the 2nd Workshop on Computational Approaches to Subjectivity and Sentiment Analysis (pp. 118-124). Stroudsburg, PA, USA: Association for Computational Linguistics.
22. Owais, S., Nafis, T., & Khanna, S. (2015). An Improved Method for Detection of Satire from User- Generated Content. International Journal of Computer Science and Information Technologies, Vol. 6 (3), pp. 2084-2088.
23. Gianti, A., Bosco, C., Patti, V., Bolioli, A., & Di Caro, L. (2012). Annotating Irony in a Novel Italian Corpus for Sentiment Analysis. In L. Devillers, B. Schuller, A. Batliner, P. Rosso, E. Douglas- Cowie, R. Cowie, & C. Pelachaud (Ed.), Proc. of the 4th International Workshop on Corpora for Research on Emotion Sentiment & Social Signals (ES3@LREC'12) (pp. 1-7). Istanbul, Turkey: ELRA.
24. Hernandez Farias, D., Sulis, E., Patti, V., Ruffo, G., & Bosco, C. (2015). ValenTo: Sentiment Analysis of Figurative Language Tweets with Irony and Sarcasm. Proceedings of the 9th International Workshop on Semantic Evaluation (SemEval 2015) (pp. 694-698). Denver, Colorado, USA: Association for Computational Linguistics.
25. Özdemir, C., & Bergler, S. (2015). CLaC-SentiPipe: SemEval2015 Subtasks 10 B,E, and Task 11. Proceedings of the 9th International Workshop on Semantic Evaluation (SemEval 2015), (pp. 479-485). Denver, Colorado: Association for Computational Linguistics.
26. Kim, S.-M. and E. Hovy (2004). Determining the sentiment of opinions. In Proceedings of the 20th international conference on computational linguistics - COLING, Stroudsburg, PA, USA, pp. 1367–1373. ACL.
27. Kao Justine and Dan Jurafsky. A Computational Analysis of Style, Affect, and Imagery in Contemporary Poetry. In Proceedings NAACL Workshop on Computational Linguistics for Literature; 2012.

 doi:10.20944/preprints202410.0286.v1

28. Rodolfo Delmonte Visualizing Poetry with SPARSAR - Poetic Maps from Poetic Content. In Proceedings of NAACL-HLT Fourth Workshop on Computational Linguistics for Literature, Denver, Colorado, June 4, 2015. Association for Computational Linguistics, pp. 68–78.
29. RID: https://github.com/jefftriplett/rid.py/blob/master/rid.py (accessed on 18/07/2024).
30. Giorgio Melchiori. Shakespeare's Sonnets. Adriatica Editrice, Bari; 1971.
31. Michael Schoenfeldt. Cambridge introduction to Shakespeare's poetry. Cambridge University Press, Cambridge; 2010.