Computing Poetry

Rodolfo Delmonte
Department of Language Studies & Department of Computer Science
Ca’ Foscari University - 30123, Venezia, Italy
delmont@unive.it

Abstract: We present SPARSAR, a system for the automatic analysis of poetry (and text) style which makes use of NLP tools like tokenizers, sentence splitters, NER (Name Entity Recognition) tools, and taggers. Our system in addition to the tools listed above which aim at obtaining the same results of quantitative linguistics, adds a number of additional tools for syntactic and semantic structural analysis and prosodic modeling. We use a constituency parser to measure the structure of modifiers in NPs; and a dependency mapping of the previous parse to analyse the verbal complex and determine Polarity and Factuality. Another important component of the system is a phonological parser to account for OOVWs, in the process of grapheme to phoneme conversion of the poem. We also measure the prosody of the poem by associating mean durational values in msecs to each syllable from a database and created an algorithm to account for the evaluation of durational values for any possible syllable structure. Eventually we produce six general indices that allow single poems as well as single poets to be compared. These indices include a Semantic Density Index which computes in a wholly new manner the complexity of a text/poem.

Keywords: NLP, Sentiment and Affective Analysis, Factuality and Subjectivity Analysis, Prosodic Structure, Semantic and Syntactic Processing, Metrical Structure

1 Introduction

We present SPARSAR, a system for poetry (and text) style analysis by means of parameters derived from deep poem (and text) analysis. We use our system for deep text understanding called VENSES [10] for that aim. SPARSAR [8] works on top of the output provided by VENSES and is organized in three main modules which can be used also to analyse similarities between couples of poems by the same or different poet and similarities between collections of poems by a couple of poets. These modules produce six general indices which are derived from quantitative evaluation of features derived from the analysis. They include a Semantic Density Index, a Deep Conceptual Index, a Metrical Distance Index, a Prosodic Distribution Index and a Rhyming Scheme Comparison Index. A General Evaluation Index is then produced and used to compare poems and poets with one another and establish a graded list on the basis of the parameters indicated above.

In addition to what is usually needed to compute text level semantic and pragmatic features, poetry introduces a number of additional layers of meaning by means of metrical and rhyming devices. For these reasons more computation is required in order to assess and evaluate the level of complexity that a poem objectively contains. An ambitious project would include computing metaphors and relate the imagery of the poem to his life and his Weltanschaung. This is however not our current aim. In particular, as far as metaphors are concerned, we dealt with this topic in another paper [9]. We also dealt with general quantitative measurements of poetic style in the past [3, 4, 5]. Some of the most interesting data will be reported here below. The paper is organized as follows: here below a subsection contains a short state of the art limited though to latest publications; section 2 present a quantitative linguistic approach to poetry analysis; section 3 presents SPARSAR, the system and its architecture highlighting the syntactic and semantic
modules; section 4 is dedicated to Prosody, Rhyming and Metrical Structure; section 5 presents the Poetry Reader Modeling component and finally there is conclusion.

1.1 State of the Art

Our interest in writing a program for the automatic analysis of poetry style and content derives from D.Kaplan’s program called American Poetry Style Analyzer, (hence APSA) for the evaluation and visualization of poetry style. Kaplan’s program works on the basis of an extended number of features, starting from word length, type and number of grammatical categories: verb, adjective, noun, proper noun; up to rhythmic issues related to assonance, consonance and rhyme, slant rhyme vs perfect rhyme. The output of the program is a graphic visualization for a set of poems of their position in a window space, indicated by a coloured rectangle where their title is included. D.M.Kaplan worked on a thesis documented in a number of papers [14,15].

I will base my analysis on the collected works of an Australian poet, Francis Webb who died in 1974. Webb was considered one of the best poet in the world at the time of the publication of the first edition of his Collected [18]. In the past, stylistic quantitative analysis of literary texts was performed using concordancers and other similar tools that aimed at measuring statistical distribution of relevant items. Usually adjectives, but also verbs, nouns and proper nouns were collected by manual classification [5,16,17]. This approach has lately been substituted by a computational study which makes heavy use of NLP tools, starting from tokenizers, sentence splitters, NER (Name Entity Recognition) tools, and finally taggers and chunkers. One such tools is represented by David Kaplan’s “American Poetry Style Analyzer” (hence APSA) which was our inspiration and which we intend to improve in our work. We used APSA to compare collected poems of different poets and show the output in a window, where each poet is represented by a coloured rectangle projected in space [9] but see below. The spatial position is determined by some 85 parameters automatically computed by the system on the raw poetic texts. However, the analysis is too simple and naive to be useful and trustful and in fact, a paper by Kao & Jurafsky [13] who also used the tool denounces that. In that paper, Jurafsky works on the introduction of a semantic classifier to distinguish concrete from abstract nouns, in addition to the analysis that the tool itself produces. Kaplan himself denounced shortcomings of his tool when he declared he did not consider words out of the CMU phonetic vocabulary apart from plurals and other simple morphological modifications 1. Eventually we decided to contribute a much deeper analyzer than the one available by introducing three important and missing factors: phonological rules for OOVWs, and syntax and semantics. In APSA the position that a poem will take in the window space is computed by comparing values associated automatically to features. The most interesting component of the program is constituted by the presence of weights that can be associated to parameter, thus allowing the system resilience and more perspicuity. Apart from that, there is no way for the user to know why a specific poem has been associated to a certain position in space. This was basically the reason why we wanted to produce a program that on the contrary allowed the user to know precisely why two or more poems were considered alike - on the basis of what attribute or feature - and in which proportion.

2 A quantitative analysis

One of the most valuable element in the evaluation of a poet careful treatment of meaning and words derives from the analysis of the so-called Vocabulary Richness – also referred to as Type-Token Ratio. This is derived from the ratio computed on the basis of Types and Tokens. Tokens

---

1 The syllabified version of the CMU dictionary dates 1998 and is called cmudict.0.6, which is the fifth release of cmudict.
are the total number of words considered, in this case all the words used by Webb as documented in his Collected Poems, in their actual number of occurrences, single or repeated. Types are those words but taken only once. Words are in this case corresponding to wordforms and not to lemmata, that is there are still types that duplicate the same words in the case of nouns, because they include both singular and plural forms, if existent in the frequency count. As for verbs, there are usually 4 different forms for regular verbs and there may be up to 6 different wordforms for the same lemma. To compute Vocabulary Richness (hence VR) one needs then two quantities: total number words or tokens, after erasing punctuation; total number of types that is tokens taken only once. I did this for a number of different poets and for a number of corpora available which contain non poetic writing, this in order to get a comparison. Of course we expect there to be a great difference between the two realms: simple normal writing should have a much lower value than poetry. It is also important to note that the Type/Token Ration by itself is not a good indicator of VR. Two other pieces of information should be considered: number of Hapax/Dis/Tris Legomena and their distribution in the corpus. This is something we have done in another publication [5], where we studied in detail Vocabulary Growth and distribution. I will include this study in another Table 3 below.

<table>
<thead>
<tr>
<th>Poets/Occurrences</th>
<th>Tokens</th>
<th>Types</th>
<th>VR</th>
</tr>
</thead>
<tbody>
<tr>
<td>Francis Webb</td>
<td>66965</td>
<td>12363</td>
<td>18.64</td>
</tr>
<tr>
<td>Anne Sexton</td>
<td>36501</td>
<td>5471</td>
<td>15.73</td>
</tr>
<tr>
<td>Emily Dickinson</td>
<td>31873</td>
<td>4503</td>
<td>14.13</td>
</tr>
<tr>
<td>T.S.Eliot</td>
<td>29144</td>
<td>5026</td>
<td>17.24</td>
</tr>
<tr>
<td>Sylvia Plath</td>
<td>28239</td>
<td>6166</td>
<td>21.84</td>
</tr>
<tr>
<td>Elizabeth Bishop</td>
<td>19047</td>
<td>4156</td>
<td>21.82</td>
</tr>
<tr>
<td>Robert Frost</td>
<td>21306</td>
<td>3251</td>
<td>15.26</td>
</tr>
<tr>
<td>Walt Whitman</td>
<td>76047</td>
<td>10946</td>
<td>14.39</td>
</tr>
<tr>
<td>W.B.Yeats</td>
<td>131485</td>
<td>10666</td>
<td>8.11</td>
</tr>
<tr>
<td>Wall Street Journal</td>
<td>1061166</td>
<td>28219</td>
<td>2.71</td>
</tr>
<tr>
<td>Total/Mean Poets</td>
<td>440607</td>
<td>62548</td>
<td>14.20</td>
</tr>
</tbody>
</table>

Table 1. Quantitative evaluation of Vocabulary Richness (1)

However, it is quite easy to see that Sylvia Plath has the highest ratio or VR, followed by Elizabeth Bishop, then Webb and Eliot. Other poets are placed lower in the graded scale and of course written prose is placed at the lowest. Even if VR computed with absolute number is not very indicative, we can easily see that Webb has the highest number of Types of all poets. In the following table we further deepen the study of Vocabulary Richness by examining Low Frequency Word distribution in the 6 poets. From Table 2. we can easily see that Plath corpus is now the poet with highest number of Hapax.

<table>
<thead>
<tr>
<th>Poets/Occurrences</th>
<th>Hapax</th>
<th>Bis</th>
<th>Tris</th>
<th>Rare</th>
<th>Types</th>
<th>Hapax</th>
<th>Rare</th>
</tr>
</thead>
<tbody>
<tr>
<td>Francis Webb</td>
<td>6662</td>
<td>1961</td>
<td>909</td>
<td>9532</td>
<td>12363</td>
<td>0.575</td>
<td>0.77</td>
</tr>
<tr>
<td>Anne Sexton</td>
<td>3144</td>
<td>924</td>
<td>395</td>
<td>4463</td>
<td>5471</td>
<td>0.575</td>
<td>0.81</td>
</tr>
<tr>
<td>Emily Dickinson</td>
<td>1716</td>
<td>1164</td>
<td>403</td>
<td>3283</td>
<td>4503</td>
<td>0.381</td>
<td>0.72</td>
</tr>
<tr>
<td>T.S.Eliot</td>
<td>2239</td>
<td>1365</td>
<td>366</td>
<td>3970</td>
<td>5026</td>
<td>0.445</td>
<td>0.79</td>
</tr>
<tr>
<td>Sylvia Plath</td>
<td>3686</td>
<td>982</td>
<td>384</td>
<td>5052</td>
<td>6166</td>
<td>0.598</td>
<td>0.82</td>
</tr>
<tr>
<td>Elizabeth Bishop</td>
<td>2471</td>
<td>631</td>
<td>334</td>
<td>3436</td>
<td>4156</td>
<td>0.594</td>
<td>0.82</td>
</tr>
<tr>
<td>Robert Frost</td>
<td>1730</td>
<td>548</td>
<td>240</td>
<td>2518</td>
<td>3251</td>
<td>0.532</td>
<td>0.77</td>
</tr>
<tr>
<td>Walt Whitman</td>
<td>5318</td>
<td>1753</td>
<td>845</td>
<td>7916</td>
<td>10946</td>
<td>0.486</td>
<td>0.72</td>
</tr>
<tr>
<td>W.B.Yeats</td>
<td>4698</td>
<td>1821</td>
<td>874</td>
<td>7393</td>
<td>10666</td>
<td>0.440</td>
<td>0.79</td>
</tr>
</tbody>
</table>
This second Table evaluates Vocabulary Richness on the basis of words repeated only once, twice or three times, and Rare Words, that is their sum. From this second evaluation we see the primacy of Sylvia Plath and Elizabeth Bishop with the highest percentages of less repeated words. Values indicated in the column Hapax, when subtracted from 1, will give the so-called Repetition Rate, which is complementary to number of once words. However the computation still suffers from the influence of a Zipfian variable, that is, the fact that by increasing the number of occurrences or tokens, number of types are meant to decrease in an incremental fashion. To better understand this point we should have made available for all poets, the Table of Vocabulary Increase which is however only available for Webb and has appeared in one of my previous publications (Delmonte 1980). We report this table together with the values of Standard Deviations for the six Phases or Folds into which Webb’s corpus had been subdivided. Consider that these data are slightly different from current ones, because they were derived from previous version of Collected Poems.

<table>
<thead>
<tr>
<th>Types/Phase</th>
<th>Ph1</th>
<th>Ph2</th>
<th>Ph3</th>
<th>Ph4</th>
<th>Ph5</th>
<th>Ph6</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tokens</td>
<td>10540</td>
<td>10540</td>
<td>10540</td>
<td>10540</td>
<td>10540</td>
<td>10538</td>
<td>63238</td>
</tr>
<tr>
<td>Types</td>
<td>3606</td>
<td>3526</td>
<td>3260</td>
<td>3096</td>
<td>3626</td>
<td>3205</td>
<td>20319</td>
</tr>
<tr>
<td>New Words</td>
<td>1422</td>
<td>1276</td>
<td>1189</td>
<td>999</td>
<td>1401</td>
<td>1144</td>
<td>7431</td>
</tr>
<tr>
<td>Hapax Leg.</td>
<td>1246</td>
<td>1151</td>
<td>1089</td>
<td>897</td>
<td>1263</td>
<td>1021</td>
<td>6667</td>
</tr>
</tbody>
</table>

Table 3. Vocabulary Increase in Webb’s Poetic Corpus

<table>
<thead>
<tr>
<th>Types/Phase</th>
<th>Ph1</th>
<th>Ph2</th>
<th>Ph3</th>
<th>Ph4</th>
<th>Ph5</th>
<th>Ph6</th>
<th>Means</th>
<th>Var.C.</th>
<th>St.Dev.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Types</td>
<td>.884</td>
<td>.181</td>
<td>-.239</td>
<td>-1.154</td>
<td>.783</td>
<td>-.455</td>
<td>3386.5</td>
<td>.168</td>
<td>207.5</td>
</tr>
<tr>
<td>New Words</td>
<td>.915</td>
<td>.270</td>
<td>-.158</td>
<td>-1.454</td>
<td>1.031</td>
<td>-.612</td>
<td>1236.5</td>
<td>.133</td>
<td>147.3</td>
</tr>
<tr>
<td>Hapax Leg.</td>
<td>.287</td>
<td>.183</td>
<td>-.165</td>
<td>-.380</td>
<td>-.314</td>
<td>-.238</td>
<td>1111.2</td>
<td>.115</td>
<td>127.4</td>
</tr>
</tbody>
</table>

Table 4. Vocabulary Increase in Webb’s Poetic Corpus expressed in Standard Error

As can be easily gathered from these Tables 3 and 4, Vocabulary Increase has been quite high in the first two Phases of Webb’s poetic production, and in the second last. It suffered a slow down in the other three Phases, but was nonetheless quite sustained.

2.1 Images and Metaphors

Francis Webb’s poetry is full of metaphors, similes and other rhetorical devices that we will look into more deeply in this section. Extended or explicit metaphors are enacted in a poem basically by the use of two grammatical linguistic markers: LIKE and AS. These two conjuctions are the intermediary linguistic and grammatical means to assert the poet’s presence, his/her command of the communicative goal the poem may have in the reader. In other words they make patent, apparent the presence of the point of view of the poet. In Table 5, we present data from a number of contemporary poets which have been partly made object of enquiry in previous works by myself [4,5]. An important contribution by Kaplan’s APSA, was making available corpora of a number of American poets which we may now compare to Webb in order to better evaluate the linguistic properties of his work. In the same Table I also added figures for the use of the two other conjuctions, AND/OR which are commonly used to conjoin or disjoin alternate linguistic items
in normal prose writing. This is not so in poetry and Webb uses the two conjunctions as a substitute to a null marker of adjacency: two or more concepts are thus juxtaposed to communicate synaesthetic images by means of asyndetons. There is no room here to show examples but it is quite normal that poets coordinate or disjoint images rather than concepts or semantic propositions in their poems. The other important component that is used by Webb to produce metaphors are list structures, that is images or concepts separated by commas repeated in a sequence of two or three. We counted both single words and also two words or bigrams, again separated by commas in a sequence. This is an artifice that achieves the same result of asyndetic juxtaposition by proposing parallel structures to the reader.

<table>
<thead>
<tr>
<th>Poets/Markers</th>
<th>LIKE</th>
<th>AS</th>
<th>AND</th>
<th>OR</th>
<th>LISTS</th>
<th>Total Like/As</th>
<th>Ratio WMM</th>
<th>Total And/Or</th>
</tr>
</thead>
<tbody>
<tr>
<td>Francis Webb</td>
<td>221</td>
<td>345</td>
<td>2355</td>
<td>251</td>
<td>298</td>
<td>566</td>
<td>30.36</td>
<td>139.81</td>
</tr>
<tr>
<td>Anne Sexton</td>
<td>170</td>
<td>148</td>
<td>643</td>
<td>63</td>
<td>66</td>
<td>318</td>
<td>20.21</td>
<td>44.88</td>
</tr>
<tr>
<td>Emily Dickinson</td>
<td>49</td>
<td>135</td>
<td>390</td>
<td>78</td>
<td>61</td>
<td>184</td>
<td>13.02</td>
<td>33.12</td>
</tr>
<tr>
<td>T.S.Eliot</td>
<td>26</td>
<td>54</td>
<td>651</td>
<td>107</td>
<td>124</td>
<td>80</td>
<td>4.64</td>
<td>43.91</td>
</tr>
<tr>
<td>Sylvia Plath</td>
<td>111</td>
<td>93</td>
<td>404</td>
<td>44</td>
<td>118</td>
<td>204</td>
<td>9.34</td>
<td>20.51</td>
</tr>
<tr>
<td>Elizabeth Bishop</td>
<td>124</td>
<td>83</td>
<td>656</td>
<td>110</td>
<td>61</td>
<td>207</td>
<td>9.49</td>
<td>35.11</td>
</tr>
<tr>
<td>Robert Frost</td>
<td>86</td>
<td>167</td>
<td>72</td>
<td>78</td>
<td>49</td>
<td>253</td>
<td>16.58</td>
<td>9.83</td>
</tr>
<tr>
<td>Walt Whitman</td>
<td>37</td>
<td>245</td>
<td>1757</td>
<td>301</td>
<td>925</td>
<td>282</td>
<td>19.60</td>
<td>143.02</td>
</tr>
<tr>
<td>Total</td>
<td>824</td>
<td>1270</td>
<td>6928</td>
<td>1032</td>
<td>1702</td>
<td>2094</td>
<td>2094</td>
<td>7950</td>
</tr>
<tr>
<td>Mean</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>261.7</td>
<td>261.7</td>
<td>993.7</td>
</tr>
<tr>
<td>Weighted Mean</td>
<td></td>
<td>15.60</td>
<td>15.60</td>
<td></td>
<td></td>
<td>59.23</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 5. Grammatical Markers for Metaphor and Synaesthesia

In Table 1, we show absolute and weighted frequency values of the occurrence of the conjunctions used to objectively build metaphors. If we look at totals, we understand that Francis Webb is the poet that uses most of explicit markers, like/as. Together with Whitman, they are the two poets using most similar devices. However absolute counts need to be interpreted and relativized in order to acquire a comparable value, and this is done by weighting absolute figures by the number of types making up their corpus of poems. This is what we do in the two final columns where we report the ratio of absolute totals by the value of Vocabulary Richness that is computed in the section below, dedicated to quantitative measurements. Weighted Metaphor Markers (or WMM for short) indicate clearly Webb’s primacy in the richness of explicit metaphoric constructs as testified by the amount of LIKE and AS occurrences. In fact, the same applies to the other two conjunctions: Webb is by far the most frequent user followed by Whitman. List structure on the contrary are found in a greater number in Whitman followed by Webb. Another syntactic construction often used to produce metaphoric meaning is by copulative structure, where the conjuncts like/as are omitted, thus generating a “strange” property assignment process to the subject of the copula.

3 SPARSAR - Automatic Analysis of Poetic Structure and Rhythm with Syntax, Semantics and Phonology

SPARSAR[8] produces a deep analysis of each poem at different levels: it works at sentence level at first, than at verse level and finally at stanza level (see Figure 1 below). The structure of the system is organized as follows: at first syntactic, semantic and grammatical functions are evaluated. Then the poem is translated into a phonetic form preserving its visual structure and its
subdivision into verses and stanzas. Phonetically translated words are associated to mean duration values taking into account position in the word and stress. At the end of the analysis of the poem, the system can measure the following parameters: mean verse length in terms of msec. and in number of feet. The latter is derived by a verse representation of metrical structure. Another important component of the analysis of rhythm is constituted by the algorithm that measures and evaluates rhyme schemes at stanza level and then the overall rhyming structure at poem level. As regards syntax, we now have at our disposal, chunks and dependency structures if needed. To complete our work, we introduce semantics both in the version of a classifier and by isolating verbal complex in order to verify propositional properties, like presence of negation, computing factuality from a crosscheck with modality, aspectuality – that we derive from our lexica – and tense. On the other hand, the classifier has two different tasks: distinguishing concrete from abstract nouns, identifying highly ambiguous from singleton concepts (from number of possible meanings from WordNet and other similar repositories). Eventually, we carry out a sentiment analysis of every poem, thus contributing a three-way classification: neutral, negative, positive that can be used as a powerful tool for evaluation purposes.

As said above, we have been inspired by by Kaplan’s tool APSA, and started developing a system with similar tasks, but which was more transparent and more deeply linguistically-based. The main new target in our opinion, had to be an index strongly semantically based, i.e. a “Semantic Density Index” (SDI). With this definition I now refer to the idea of classifying 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 are justaposed in a sequence
- also difficult to understand are sequences of adjectives and nominals used as modifiers, union of such items with a dash.

There are other elements that I regard very important in the definition of semantic parameters and are constituted by presence of negation and modality: this is why we compute Polarity and Factuality. Additional features are obtained by measuring the level of affectivity by means of sentiment analysis, focussing on presence of negative items which contribute to make understanding more difficult.

The Semantic Density Index is derived from the computation of a number of features, some of which have negative import while others positive import. 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.

At the end we come up with a number of evaluation indices that include: a Constituent Density Index, a Sentiment Analysis Marker, a Subjectivity and Factuality Marker. We also compute a Deep Conceptual Index, see below.
The procedure is based on the tokenized sentence, which is automatically extracted and may contain many verses up to a punctuation mark, usually period. Then I use the functional structures which are made of a head and a constituent which are measured for length in number of tokens. A first value of SDI comes from the proportion of verbal compounds and non-verbal ones. I assume that a "normal" distribution for a sentence corresponds to a semantic proposition that contains one verbal complex with a maximum of four non verbal structures. More verbal compounds contribute to reducing the SDI.

The other contribution comes from lemmatization and the association of a list of semantic categories, general semantic classes coming from WordNet or other similar computational lexica. These classes are also called supersense classes. As a criterion for grading difficulty, I consider more difficult to understand a word which is specialized for a specific semantic domain and has only one such supersense label. On the contrary, words or concepts easy to understand are those that are ambiguous between many senses and have more semantic labels associated to the lemma. A feature derived from quantitative linguistic studies is the rare words, which are those words that appear with less than 4 occurrences in frequency lists. I use the one derived from Google GigaWord.

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

![SPARSAR Three Level Algorithm](image)

**Figure 1.** The SPARSAR three-level system
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. Eventually a single index is associated to the poem which should be able to differentiate those poems which are easy from the cumbersome ones.

What I do is dividing each item by the total number of tagged words and of chunks. In detail, I divide verbs found by the total number of tokens (the more the best); I divide adjectives found by the total number of tokens (the more the worst); I divide verb structures by the total number of chunks (the more the best); I divide inflected vs uninflected verbal compounds (the more the best); I divide nominal chunks rich in components: those that have more than 3 members (the more the worst); I divide semantically rich (with less semantic categories) words by the total number of lemmas (the more the worst); I count rare words (the more the worst); I count generic or collective referred concepts (the more the best); I divide specific vs ambiguous semantic concepts (those classified with more than two senses) (the more the worst); I count doubt and modal verbs, and propositional level negation (the more the worst); I divide abstract and eventive words vs concrete concepts (the more the worst); I compute sentiment analysis with a count of negative polarity items (the more the worst).

Another 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.

### 3.1 Indices and Graphical Display

Eventually, together with the indices indicated above, I devised four complex indices and the final best grading index to evaluate a poem. Another important index is the one represented by the Lexical Selection index which is a combination of Vocabulary Richness and the Rare Words, where the first is the ratio obtained between Types and Tokens, the second is the ratio obtained between Rare Words and Types. That is, with the first index we know how many single wordforms and no repetition are included in the poem, and with the second we learn how many types are rare words, i.e. are words that have frequency equal or below 3 in the general index of billion words obtained from the web and other such sources.

In the figures obtained by Kaplan’s Poetry Analyzer we increased the weight associated to all rhyming devices which include the following: Alliteration Frequency; Assonance Frequency; Consonance Frequency; Slant End Rhyme Frequency; Semi End Rhyme Frequency; Perfect End Rhyme Frequency; Identity End Rhyme Frequency; Sound Devices Frequency; Partial End Rhyme Frequency; Full End Rhyme Frequency. Without this modification the distribution of the poems in the space would look totally different, and this can be seen by comparing the two pictures below in Fig. 2 and 3: in the first one rhyming devices are set to zero; on the contrary, in the second picture the same parameters are risen to half their value. As can be seen, none of the poems preserves the same position apart from “The Gunner”: curiously enough, the poem has also the same position in our system evaluation. The question is that this poem is a strongly rhymed structure of the AABB with full end rhymes and there is no reason why by modifying weights this feature does not also modify its position in the window. On the contrary, if we look at the output of our system, which will be shown in a section below, the rhyming feature is particularly highlighted and contributes a much higher value of metrical measure parameter. “The Gunner” in fact is assigned totally different positions according to parameters evaluated. As can be easily gathered, collapsing all parameters in one single value that is used to position the poem in the vector space is not the best practice because it obscures the specificity of the poem in
terms of semantic and conceptual import, rather than rhetoric or rhyming devices. In Figures 2 and 3 below I show the output of APSA system for those poems regarded as deviant in a previous graph, and produce an analysis where the poem positioned in centre space represent the norm and this is *Five Days Old*. The other poems are: *A Sunrise, The Gunner, The Explorer’s Wife, For My Grandfather, Idyll, Middle Harbour, Politician, To a Poet, The Captain of the Oberon, Palace of Dreams, The Room, Vancouver by Rail, Henry Lawson, Achilles and the Woman.*

![Figure 2. Francis Webb 15 deviant poems not filtered](image)

![Figure 3. Francis Webb 15 deviant poems filtered by weights](image)

In the APSA graphical spatial representation, the positioning of all poems starting from top of the window is almost totally different when modifying weights associated to Rhymes and other Rhetorical Devices. They can be listed as follows, where we collected labels associated to each
colored rectangle in the vertical order in which they appear. We then list in the third column the graded scale produced by our system. As can be easily gathered from Table 6, only four poems are graded the same by the two systems. If we look at Table 7, we discover then that poems are graded quite differently by SPARSAR depending on the indices involved.

<table>
<thead>
<tr>
<th>Poems Graded the same in at least two scales</th>
<th>A. Figure 2. before weights</th>
<th>B. Figure 3. after weights</th>
<th>C. SPARSAR grading</th>
</tr>
</thead>
<tbody>
<tr>
<td>For My Grandfather</td>
<td>Achilles</td>
<td>Middle Harbour</td>
<td></td>
</tr>
<tr>
<td>Achilles</td>
<td>Politician</td>
<td>Bridgehead</td>
<td></td>
</tr>
<tr>
<td>Five Days Old</td>
<td>Idyll</td>
<td>For My Grandfather</td>
<td></td>
</tr>
<tr>
<td>The Room</td>
<td>The Captain</td>
<td>To a Poet</td>
<td></td>
</tr>
<tr>
<td>Bridgehead</td>
<td>The Explorer's Wife</td>
<td>The Room</td>
<td></td>
</tr>
<tr>
<td>The Captain A/C</td>
<td>The Captain</td>
<td>Middle Harbour</td>
<td></td>
</tr>
<tr>
<td>Henry Lawson B/C</td>
<td>Middle Harbour</td>
<td>Henry Lawson</td>
<td></td>
</tr>
<tr>
<td>The Explorer's Wife</td>
<td>To a Poet</td>
<td>Idyll</td>
<td></td>
</tr>
<tr>
<td>Vancouver A/C</td>
<td>Vancouver by Rail</td>
<td>Bridgehead</td>
<td></td>
</tr>
<tr>
<td>Henry Lawson</td>
<td>Five Days Old</td>
<td>A Sunrise</td>
<td></td>
</tr>
<tr>
<td>Palace of Dreams</td>
<td>For My Grandfather</td>
<td>The Explorer's Wife</td>
<td></td>
</tr>
<tr>
<td>The Gunner A/B/C</td>
<td>The Gunner</td>
<td>The Gunner</td>
<td></td>
</tr>
<tr>
<td>Politician</td>
<td>A Sunrise</td>
<td>Achilles</td>
<td></td>
</tr>
<tr>
<td>Idyll</td>
<td>Vancouver by Rail</td>
<td>Five Days Old</td>
<td></td>
</tr>
<tr>
<td>To a Poet</td>
<td>The Room</td>
<td>Palace of Dreams</td>
<td></td>
</tr>
<tr>
<td>A Sunrise</td>
<td>Palace of Dreams</td>
<td>Politician</td>
<td></td>
</tr>
</tbody>
</table>

Table 6. Comparing graded positions in space in APSA and in SPARSAR

<table>
<thead>
<tr>
<th>POEMS</th>
<th>Poetic Rhetoric Devices</th>
<th>Metrical Length</th>
<th>Semantic Density</th>
<th>Prosodic Structure Distribution</th>
<th>Deep Conceptual Index</th>
<th>Identical</th>
</tr>
</thead>
<tbody>
<tr>
<td>Middle Harbour</td>
<td>2</td>
<td>5</td>
<td>9</td>
<td>7</td>
<td>7</td>
<td>X 7</td>
</tr>
<tr>
<td>Bridgehead</td>
<td>10</td>
<td>9</td>
<td>5</td>
<td>6</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>Grandfather</td>
<td>7</td>
<td>8</td>
<td>10</td>
<td>13</td>
<td>12</td>
<td></td>
</tr>
<tr>
<td>To a Poet</td>
<td>5</td>
<td>13</td>
<td>8</td>
<td>4</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>The Room</td>
<td>4</td>
<td>15</td>
<td>1</td>
<td>3</td>
<td>3</td>
<td>X 3</td>
</tr>
<tr>
<td>The Captain</td>
<td>13</td>
<td>11</td>
<td>6</td>
<td>5</td>
<td>4</td>
<td></td>
</tr>
<tr>
<td>Henry Lawson</td>
<td>1</td>
<td>16</td>
<td>7</td>
<td>1</td>
<td>6</td>
<td>X 1</td>
</tr>
<tr>
<td>Idyll</td>
<td>6</td>
<td>4</td>
<td>12</td>
<td>12</td>
<td>13</td>
<td>X 12</td>
</tr>
<tr>
<td>Vancouver</td>
<td>9</td>
<td>2</td>
<td>11</td>
<td>14</td>
<td>8</td>
<td></td>
</tr>
<tr>
<td>A Sunrise</td>
<td>12</td>
<td>7</td>
<td>13</td>
<td>15</td>
<td>15</td>
<td>X 15</td>
</tr>
<tr>
<td>Explorer’s Wife</td>
<td>8</td>
<td>6</td>
<td>2</td>
<td>16</td>
<td>9</td>
<td></td>
</tr>
<tr>
<td>The Gunner</td>
<td>3</td>
<td>1</td>
<td>16</td>
<td>11</td>
<td>5</td>
<td></td>
</tr>
<tr>
<td>Achilles</td>
<td>11</td>
<td>3</td>
<td>14</td>
<td>9</td>
<td>14</td>
<td>X 14</td>
</tr>
<tr>
<td>Five Days Old</td>
<td>15</td>
<td>14</td>
<td>3</td>
<td>10</td>
<td>10</td>
<td>X 10</td>
</tr>
<tr>
<td>Palace of Dreams</td>
<td>14</td>
<td>10</td>
<td>4</td>
<td>8</td>
<td>16</td>
<td></td>
</tr>
<tr>
<td>Politician</td>
<td>16</td>
<td>12</td>
<td>15</td>
<td>2</td>
<td>11</td>
<td></td>
</tr>
</tbody>
</table>

Table 7. Comparing graded positions generated from indices computed by SPARSAR
Rhetoric Devices, Metrical and Prosodic Structure

The second module takes care of rhetorical devices, metrical structure and prosodic structure. This time the file is read on a verse by verse level by simply collecting strings in a sequence and splitting verses at each newline character. In a subsequent loop, whenever two newlines characters are met, a stanza is computed. In order to compute rhetorical and prosodic structure we need to transform each word into its phonetic counterpart, by accessing the transcriptions available in the CMU dictionary. The Carnegie Mellon Pronouncing Dictionary is freely available online and includes American English pronunciation\(^2\). Kaplan reports the existence of another dictionary which is however no longer available.\(^3\) The version of the CMU dictionary they are referring to is 0.4 and is the version based on phone/phoneme transcription. Kaplan & Blei in their longer paper specifies that “No extra processing is done to determine pronunciation ... so some ambiguities are resolved incorrectly.” [15:42]. In fact what they are using is the phoneme version of the dictionary and not the syllabified one, which has also been increased by new words. We had available a syllable parser which was used to build the VESD database of English syllables [1]. So we started out with a much bigger pronunciation dictionary which covers 170,000 entries approximately.

Remaining problems to be solved are related to ambiguous homographs like “import” (verb) and “import” (noun) and are treated on the basis of their lexical category derived from previous tagging and Out Of Vocabulary Words (OOVW). As happens in Kaplan’s system, if a word is not found in the dictionary, we also try different capitalizations, as well as breaking apart hyphenated words, and then we check at first for ’d, ’s, and s’ endings and try combining those sounds with the root word. The simplest case is constituted by differences in spelling determined by British vs. American pronunciation. This is taken care of by a dictionary of graphemic correspondences. However, whenever the word is not found we proceed by morphological decomposition, splitting at first the word from its prefix and if that still does not work, its derivation suffix. As a last resource, we use an orthographically based version of the same dictionary to try and match the longest possible string in coincidence with our OOVW. Then we deal with the remaining portion of word again by guessing its morphological nature, and if that fails we simply use our grapheme-to-phoneme parser. Here below are some of the OOVWs we had to reconstruct by means of our recovery strategy which is indicated by showing the input word rejected by the dictionary lookup, then the word found by subtraction and the final output obtained by recomposition:

\[
\begin{array}{ll}
\text{wayfarer} & \rightarrow \text{wayfare-[w_ey1f_eh1_r_r]} \\
\text{gangrened} & \rightarrow \text{gangrene-[g_ae1_nr_ah0_n_d]} \\
\text{krog} & \rightarrow \text{krog-g_r_aal_g} \\
\text{copperplate} & \rightarrow \text{copper-k_aal_p_er_p_1_ey1_t]} \\
\text{splendor} & \rightarrow \text{splendour-[s_p_l_eh1_n_d_ee]} \\
\text{filmy} & \rightarrow \text{film-f_ih1_l_miy]} \\
\text{seraphic} & \rightarrow \text{seraphine-}[s_e_r_a_ph_l_y1_k] \\
\text{unstarred} & \rightarrow \text{starred-[ah_n_s_t_aal_r_d]} \\
\end{array}
\]

Other words we had to reconstruct are: shrive, slipstream, fossicking, unplotted, corpuscle, thither, wraiths, etc. In some cases, the problem that made the system fail was the syllable which was not available in our database of syllable durations, VESD. This problem has been coped with by manually inserting the missing syllable and by computing its duration from the component phonemes, or from the closest similar syllable available in the database. We only had to add 12 new syllables for a set of approximately 1000 poems that the system computed.

---

\(^2\) It is available online at <http://www.speech.cs.cmu.edu/cgi-bin/cmudict/>.

\(^3\) Previously, data for POS were merged in from a different dictionary (MRC Psycholinguistic Database, <http://lcb.unc.edu/software/multimrc/multimrc.zip>, which uses British English pronunciation)
4.1 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, enjambements, etc. which contribute to poetic form at stanza level. Traditionally, poetic meter is visualized by a sequence of signs, typically a straight line is used to indicate vowels of stressed syllables and a half circle is positioned on vowels of unstressed ones. The sequence of these sings makes up the foot and depending on number of feet one can speak of iambic, trochaic, anapestic, dactylic, etc. poetic style.

English poetry has been for centuries characterized by iambic pentameter, that is a sequence of five feet made of a couple of unstressed + stressed syllables. Modern English poetry on the contrary – after G.M. Hopkins – has adopted a variety of stanza schemes. A poetic foot can be marked by a numerical sequence as for instance in [11,12] who uses “0” for unstressed and “1” for stressed syllables to feed a connectionist model of poetic meter from a manually transcribed corpus. There he also tries to state the view that poets are characterized by their typical meter and rhythm, which work as their fingerprint. We also agree with this view, however, we would like to be more specific on the notion of rhythm that we intend to purport. We do that in two ways: by considering stanzas as structural units in which rhyming – if existent - plays an essential role. Secondly and foremost, in our view, a prosodic acoustic view needs to be implemented as well, if any precise definition of rhythm and style is the goal. Syllables are not just any combination of sounds, and their internal structure is fundamental to the nature of the poetic rhythm that will ensue. This is partly amenable to the use and exploitation of poetic devices, which we also intend to highlight in our system. But what is paramount in our description of rhythm, is the use of the acoustic parameter of duration. The use of duration will allow our system to produce a model of a poetry reader that we intend to implement in the future by speech synthesis. In our demo we will show how poems can be characterized by the use of rhythmic and stylistic features in a highly revelatory manner, by comparing metrically similar poems of the same poet and of different poets. 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 consonants, consonant clusters, vowel or diphthong - position in the word, whether beginning, end or middle - primary stress, secondary stress or unstressed

These data have been collected in a database called VESD (Venice English Syllable Database) to be used in the Prosodic Module of SLIM, a system for prosodic self-learning activities. Syllables have been collected from WSJCAM, the Cambridge version of the continuous speech recognition corpus produced from the Wall Street Journal, distributed by the Linguistic Data Consortium (LDC). We worked on a subset of 4165 sentences, with 70,694 words which constitute half of the total number of words in the corpus amounting to 133,080. We ended up with 113,282 syllables and 287,734 phones. The final typology is made up of 44 phones, 4393 syllable types and 11,712 word types. From word-level and phoneme-level transcriptions we produced syllables automatically by means of a syllable parser. The result was then checked manually. The analysis in SPARSAR starts by translating every poem into its phonetic form: we used the CMU Pronouncing Dictionary for North American English to translate words into phoneme sequences, augmented with words derived from work above – see also [6,7]. In a second pass 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. Duration values are derived by comparison with phonologically closest ones – for this we use place, manner of articulation as parameters. We assign mean duration values in msecs to all syllables considering position and stress. We also take advantage of syntactic information computed separately to highlight chunks’ heads as produced by our bottomup parser. In that case, stressed syllables takes maximum duration value.

After reading out the whole poem on a verse by verse basis and having produces all phonemic transcription, we look for rhetoric devices. Here assonances, consonances, alliterations and rhymes are analysed and then evaluated. We introduce an important prosodic element: we produce a prosodic model of the poem and compute duration at verse level. This is done by associating durations at syllable level. In turn, these data are found by associating phonemes into syllables with our parser, which works on the basis of the phonological criterion of syllable wellformedness. Syllable structure requires a nucleus to be in place, then a rhyme with an onset and offset[1]. 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. Each syllable has been recorded with three durational values in msec.: minimum, mean and maximum duration length, with a standard deviation. To produce our prosodic model we take mean durational values. We also select, whenever possible, positional and stress values. Of course, if a syllable duration value is not available for those parameters we choose the default value, that is unstressed. 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.

Durations are then collected at stanza level and a statistics is produced. Metrical structure is used to evaluate statistical measures for its distribution in the poem. As can be easily gathered from our transcription, it is difficult to find verses 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, in our transcription it is easy to see that there is no line strictly respecting it. 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 measure that we are now able to produce are related to rhyming devices. Since we intended to take into account structural internal rhyming scheme 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 also stanza repetition at poem level may apply: in other words, we need to reconstruct the internal structure of metrical devices used by the poet. We then use this information as a multiplier. The final score is then tripled in case of structural persistence of more than one rhyming scheme; for only one repeated rhyme scheme, it is doubled. With no rhyming scheme there will be no increase in the linear count of rhythmic and rhyming devices. Creating the rhyming scheme is not an easy task. We do that by a sequence of incremental steps that assign labels to each couple of rhyming line and then matches their output. To create rhyme schemes we need all last phonetic words coming from our previous analysis. We then match recursively each final phonetic word with the following ones, starting from the closest to the one that is 6 lines far
apart. Each time we register the rhyming words and their distance, accompanied by an index associated to verse number. Stanza boundaries are not registered in this pass. The following pass must reconstruct the actual final verse numbers and then produce an indexed list of couples, Verse Number-Rhyming Verse for all the verses, 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 letter are used. I distinguish between poems divided up into stanzas and those that have no such a structure. Then I get stanzas and their internal structure in term of rhyming labels. Eventually what I want to know is the persistence of a given rhyme scheme, how many stanza 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 I do. In the final evaluation, it is possible to match different poems on the basis of their rhetorical and rhyming devices, besides their semantic and conceptual indices.

Parameters related to the Rhyming Scheme (RS) contribute a multiplier to the already measured metrical structure which as we already noted is extracted from the following counts: 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. 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 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). So the RR is one parameter and is 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 evaluation are 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. The final value is obtained by dividing up the RR by the total number of lines and multiplying 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.

We show now a number of examples from Francis Webb’s poetry[18]. First example is “Morgan’s Country” which is written in Dante’s “terza rima” in tercets, the rhyming pattern is “a-b-a, b-c-b, c-d-c, etc.” using chain rhymes. The poem is made up of 8 stanzas with the previous scheme and two ending stanzas which break it: the nineth stanzas introduces rhymes which are not present in the other stanzas, while the tenth stanza is made of two verses or a couplet, where the second verse repeats words contained in the first stanza of the poem, which we show here below:

First stanza:

This is Morgan’s country: now steady, Bill.
(Stunted and grey, hunted and murderous.)
Squeeze for the first pressure. Shoot to kill.

Two last stanzas:

Seven: and a blaze fiercer than the sun.
The wind struggles in the arms of the starved tree,
The temple breaks on a threadbare mat of grass.

Eight: even under the sun’s trajectory
This country looks grey, hunted and murderous.

Another interesting poem is “Before Two Girls” which is made up of six stanzas each of nine
verses. The internal structure of the stanzas are characterized by iambic pentameters and
sometimes alexandrines, apart from the sixth line which is always trimeter, that is three feet or six
syllables. Here below I show the second and third stanzas:

For Jack and Jill, tumbling in the kennel or the stews
Take dust to bed with them under a consumptive moon
Choking in the heavens, or under a smoggy sun
And frayed counterpane of cloud. But the moon for man
Still goads the spavined panther, the ocean, claws
By proxy every coast
To reclaim for a flowing all that she can of dust.
And dust we become. Was it a whisper or a nothing
Came from our dust before His touch or His breathing?

Moist spot on the lung burning at compline, and burning
In the thorax of dawn: all deathbeds are gathering in
Around these two, and the rheum and haze of sin.
Darkness the sun, the moon. Time has begun.
For He is His pattern: one prayer, twisting and turning
Among the amphibious pleading
Forms and faces, compels the waters and the bleeding
Of His Heart: as for ever the mangled hands and breast
May ponder into flower the hot tumuli of dust.

Rhymes are not always isomorphic nor isometric as can be seen in the previous two stanzas. There
are perfect or full rhymes for instance in the lines ending in “burning”/”turning” and also
“in”/”sin”. But then there are half-rhymes in the lines ending in “coast”/”dust”, “sun”/”man”,
“stews”/”claws”. Rhyme patterns vary in the first five lines, but are consistent in the last four. The poem repeats always the scheme ABCCADDEE; but the second stanza has another pattern in the
first five lines: ABCDAEEFF and we report it below. It must be noted however that the poem
contains a lot of repetition, which are centered around the word “dust”, which is repeated 12 times
and in one case it is made to rhyme with “lust”. And the second stanza repeats “to be”, to half-
rhyme with the previous ending in “knowingly”:

If they have finished the dishes and are leaning
Together in some timidity for the eye
Of the box–camera to wink at them, knowingly
Some chinaware chuckles, It had to be, to be.
They forsook the amorous dust for immaculate Meaning,
Only to find that same dust
Swarming fog–silken, making faces in its lust
(But faces melting into pity, irresolution, loss)
All about the pure intimacy of the Faucet, the Cross.
So, in this case, the computation of rhyming schemes will end up with two different values. But there also poems with no rhyme scheme at all, or only limited to some of the lines in a stanza as for instance in “Ball’s Head Again”. In this poem there are four stanzas each one made up of a different number of lines: 13, 15, 22 and 23 lines. Each stanza uses different internal rhyming schemes which are however not repeated in the following stanza. So the final count the system makes is that there are 4 different rhyming schemes which have internal rhymes. Also the internal metrical structure is highly diversified or heterometric: it goes from pentameters to monometer.

“Nuriootpa” is a limerick dedicated to the wine valley: its rhyming scheme is AABBA as required, and also the contents are structures as the limerick requires. First line presents a character the second adds something he/she did, then two following lines insert some comments and elaborations, and there’s a final line which ends the story. Here is one stanza:

Men with ancient communal brooms last night
Went over their pride and joy to doll it up aright.
Interloping box and bottle,
Some stray native head of cattle,
Were worried away or chivvied out of sight.

All lines have an internal metrical structure based on the iambic pentameter but with 11 or sometimes 12 syllables, apart from the fifth line which is trimeter with seven syllables. Special structure is represented by “Foreword” and “Yellowhammer”. The first poem is organized in three stanzas each 14 lines long. The internal rhyming scheme is represented by a complex but highly repetitive alternate structure, ABABCDBBCDEFEF, where full rhymes are sometimes accompanied by half-rhymes. Here is one stanza:

We do not forget how to kill.
Our classroom is clean, quite old,
So demurely practical
And exquisitely patrolled
By so many teachers—by one
Teacher. His accent, occasion
May vary; not themes of his world.
Tamed, compromising, cold,
The Outside with its lightning and sun
Surrenders all playtime passion
To sidle through well–scrubbed glass.
Few foreign trifles are hidden
Under our desks, and to pass
Notes is strictly forbidden.

All lines have a trimeter feet structure with 7/8 syllables. “The Yellowhammer” also has three 13 lines stanzas and a final 11 lines stanza, with complex rhyming schemes: ABCDCEFEGFE, ABCDDBEEFHGCDEFGHIHG, ABCDCEFEGCDEFGHIHG, ABCDCEFEGCDEFGHIHG. Another interesting case is represented by “The Runner” which is organized in two octaves, but rhymes do not follow the “ottava rima” scheme. They are organized in couplets in full and half-rhymes. Here is one stanza:

Watch, and for a moment pace him on.
Clipped are the wings of space from him, and gone
Thrust from the hips, self–conscious overstride.
His face hangs yellow, curtainless and void
As a cracked window in a headlong shack,
Brushed by the terrible hammer of the track,
The little spider of torment kicks and swings
In the grey, collapsing bubble of his lungs.

Finally we report another case of octave stanza poem, “Around Costessey – The Horses”, where the rhyme scheme is again different from the expected alternate one, and is: AABCCBDD, where there are four rhyme types rather than only three as the traditional “ottava rima”. The metrical structure is iambic pentameter as usual, with more than ten syllable however; and there are two lines which are trimeter, with 7 syllables, the fourth and the last line. Here is one stanza:

The vegetative soul is the dedicated rhetorician:
Yellow knuckles of gorse are eloquent; motion
Is the psyche entire whose fullness is a naked growing
Ungirt with passion or reflection.
Grass meanders intoxicate in green simple action,
Little hills troll the pastoral catches, allowing
Hosannas of Saints in sober gesture alive
As flowering cherry along a drive.

### 4.1.1 Computing Metrical Structure

Here below some excerpts from a short poem “The Runner” where each word is followed by its phone representation after a slash:

0 the/dh_ah runner/r_ah1_n_er
1 watch/w_ae1_l ch and/ah_n_d for/f_ao_r a/ah moment/m_ow1_m ah_n_t pace/p_ey1_s him/hh_ih1_m on/aa_n
2 clipped/k_l_ih1_p t are/aa_r the/dh_ah wings/w_ih1_ng z of/ah_v space/s_p_ey1_s from/f_r_ah1_m
3 him/hh_ih1_m and/ah_n_d gone/g_ao1_n
4 thrust/th_r_ah1_s t from/f_r_ah1_m the/dh_ah hips/hh_ih1_p s self/s_eh1_l f conscious/k_ae1_l n sh ah_s
5 overstride/ow_v_s t_r ay1_l d
6 his/hh_ih1_z face/f_ey1_s hangs/hh_ah1_l ng z yellow/y_eh1_l ow curtainless/k_eh1_l t ah n l eh1_s and/ah_n_d
7 void/v_ey1_l d
8 a/ah cracked/k_r_ah1_k t window/w_ih1_n_d ow in/ih1_n a/ah headlong/hh eh1_l d l ao2_ng
9 shack/sh_ae1_k
10 brushed/b_r_ah1_sh t by/b ay the/dh_ah terrible/t eh1_r ah b_ah_l hammer/hh_ah1_l er of/ah_v the/dh_ah
11 track/t r ay1_l k
12 the/dh_ah little/l_ih1_t ah l spider/s_p ay1_l d er of/ah_v torment/t ao1_r m eh2_n t kicks/k_ih1_k s and/ah_n_d
13 swings/s_w _ih1_ng z

stanza no. 1 verse durations [1998.27, 2534.2, 2871.87, 2651.1, 1951.35, 2004.04, 2653.2140.34]
total syllable length 18804.170000000002
mean verse length 2350.5212500000002
standard deviation 365.4796116821 skewness 0.13034352150269 kurtosis -1.8094344468253

verse no. 1. number of feet 4 total syllable length 12 metrical structure [1,0,0,0,0,0,1,0,0,1,1,0]
verse no. 2. number of feet 6 total syllable length 18 metrical structure [1,0,0,0,0,1,0,0,1,0,0,1,0]
verse no. 3. number of feet 6 total syllable length 17 metrical structure [1,0,1,0,0,0,1,0,1,1,0,0,0,1,0]
verse no. 4. number of feet 7 total syllable length 13 metrical structure [1,1,1,1,0,1,1,0,0,1,1,0]
verse no. 5. number of feet 6 total syllable length 13 metrical structure [0,0,1,1,0,1,0,1,0,1,1,0]
verse no. 6. number of feet 4 total syllable length 14 metrical structure [1,0,0,1,0,0,1,1,0,0,0,1,0]
verse no. 7. number of feet 6 total syllable length 20 metrical structure [0,0,1,1,0,0,0,0,1,0,1,1,0,0,0,1,0]
verse no. 8. number of feet 7 total syllable length 14 metrical structure [1,0,0,1,0,1,0,0,1,1,0,1,0,1,1,1]
Eventually we produce the overall stanza feet pattern that we then use to compute the overall metrical structure.

stanza feet pattern [5,5,5,5,7,5,5,5]
mean verse length 2275.3781249999997
standard deviation 349.05328046433 skewness 0.503463090507 kurtosis -1.506168632999
mean verse length in no. of feet 5.5
standard deviation 0.9660917830792 skewness 0.2218067893147 kurtosis -1.010204081632

Additional measure that we are now able to produce are related to rhyming devices and they are for this poem:

Absolute Count of Rhyming Devices = 129.5
Ratio of Absolute Count by No of Stanzas = 0.017857142857142856
Ratio of RhDevs Absolute Count by Corrected No of Verses = 8.09375
Ratio of RhDevs Absolute Count by Repetition Rate = 198.14207650273224
Ratio of RhDevs Absolute Count by Rare Word Repetition Rate = 136.5380437262087
Ratio of RhDevs Absolute Count by No Tokens = 56.93040864574422

And then the Rhyming Schemes:
Stanza-based Rhyme Schemes [1-[1-a,2-b,3-c,4-c,5-d,6-d,7-e,8-e],2-[10-f,11-f,12-g,13-g,14-h,15-h,16-i,17-i]]
Different Rhyme Schemes [[a,b,c,d,d,e,e],[f,f,g,g,h,h,i,i]]
Number of Different Rhyme Schemes 2
Multiplier Factor 3.5
Different Stanzas RSs [8]
Different Rhyme Types [4,5]

5 SPARSAR Linguistic and Statistic Graphical Modeling

Eventually, the system produces seven complex indices which include a final best grading index to evaluate a poem in comparison with others. These indices allow comparison amongst a number of poems by the same or different authors, and the selection of the two poems which are the closest according to the indices computed.

SPARSAR visualizes differences by increasing the length and the width of the coloured bars. Here below is the output of the system for two poems “Bells” and “A Death at Winson Green”. Parameters evaluated are shown by coloured bars and 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. Two poems are then easily comparable visually in that manner, as can be seen from the two figures below where the two poems indicated are shown.
Bells

Poetic Rhetoric Devices

Metrical Length

Semantic Density

Prosodic Structure Dispersion

Deep Conceptual Index

Rhyming Scheme Comparison

Figure 7. Sparsar Indices for the poem Bells

Figure 8. Sparsar Indices for the poem A Death at Winson Green
In the figures below we show comparisons of different poems by the same author as we did using APSA system. In our case, the representation is two dimensional rather than one dimensional as in APSA where the position in space indicates conformity or deviancy as would cluster analysis. We use size of the box to indicate the importance of the different indices: the bigger the size the more important and the higher score they have got. Position in space on the contrary is used to indicate distributional properties of the different indices: if the indices are well balanced then the position would be center otherwise it would move sideways. This movement is determined by values computed from Skewness and Kurtosis indices and are basically related to Standard Deviation. As we can gather from the position of Five Days Old, it is computed as central also in our space, but the remaining poems are all scattered in different positions from what APSA did.

Figure 9. Sparsar Indices for 15 deviant poems

In particular, in our figure a poem like To a Poet is positioned at the bottom rather than at the top as in APSA, and The Explorer’s Wife is positioned at the top rather than at the bottom. Also The Gunner is in high position and not in a low position. The same applies to A Sunrise. In other words we have totally different appraisal of the content of the poems in the two systems. As a final exercise, I computed the comparative image of 53 poems by Sylvia Plath and the same amount by Francis Webb. This should result in a general image which is suggestive of the way in which the poetic world has been organized as a whole. The spread of the poems in the space is much wider in Plath’s figure than in Webb’s. I computed the same poems with APSA tool – which we don’t show here - and didn’t get any valuable comparison apart from one or two poems in Webb’s scheme which were regarded deviants. As for Plath’s poems no matching has been found. In fact all the values – we don’t show them here for lack of space - computed for Plath’s poems are higher than Webb’s thus indicating that Plath’s poetry is much richer and more consistent than Webb's. But of course, what can be obtained by these measurements is only part of the content of each poem: metaphors and images can only be evaluated by critical manual perusal of each stanza and each line. Also the intermingling of imagery rhyme and rhythm can only be appreciated by deep and profound reading of the poems.
Figure 10. Sparsar General Graded Evaluation of 53 poems by Webb

Figure 11. Sparsar General Graded Evaluation of 53 poems by Plath
6 Comparing two poets

Another important feature of the system is the possibility to compare the works of two poets and to receive as a result the closest two poems if any such poem has been so evaluated by the system. Using the six macro indices to measure the semantic and prosodic content of each poem, the system is able at the end of the computation to make comparisons and this is done by transforming into correlation coefficients each pair of macro indices associated to any two poems from the list of the two authors to be compared. We use the formula for Pearson’s correlation coefficients. Here below are the best correlations coefficients obtained from a comparison of the most similar poems chosen from Francis Webb and Sylvia Plath collected poems:

Nessun Dorma - Blackberrying : 0.9995063248666134
A Death at W.G - The old women : 0.9983739250672603
A Death at W.G - Blackberrying : 0.9978765513783355
The Explorer’s Wife - Blackberrying : 0.9970534742059991
Nessun Dorma – The old women : 0.9964034793378926
On first hearing - Dialogue between : 0.992759946803719
A Death at W.G – The Goatsucker : 0.992550806997867
Nessun Dorma – The Goatsucker : 0.9923992745484463
The Explorer’s Wife – The old women : 0.9903907589788684

And here below are the two graphical representations for the two best choices Nessun Dorma and Blackberrying, where we can see that the correlation coefficients measure take really close shots of actual values computed by the system.

Figure 12. Sparsar’s evaluation for Plath’s poem Blackberrying
7 TTS and Modeling Poetry Reading

Now another important part of the work regards using the previous analyses to produce intelligible, correct, appropriate and possibly pleasant poetry reading by a Text To Speech system. In fact, the intention was more ambitious and was producing an “emotional” reading of a poem or at least expressive reading in the sense also intended by work reported under [22, 23]. The TTS we are now referring to is the one freely available under Mac OSX in Apple’s devices. In fact, the output of our system can be used to record .wav or .mpeg files that can then be played by any sound player program. The information made available by the system is sufficiently deep to allow for Mac TTS interactive program to adapt the text to be read and model it accurately. If a poem is inputted to the TTS the result is terrible: the internal rules compute stanza boundaries as sentence delimiters. The TTS is also blind to verse boundaries. More importantly, the TTS reads every sentence with the same tone, thus contributing an unpleasant repeated overall boring sense which is not certainly corresponding to the contents read.

At first we intended to use syllable durations as they have been produced by our system, but this proved to be unfeasible: the TTS accepts durations only at phoneme level in the TUNE mode. This however requires the production of a complete intonational contour associated to each vowel phoneme which is out of the system’s scope and ability.

We then discovered that the internal commands can actually modify sensibly the content of the text to be read. The voices now available are very pleasant and highly intelligible. So we decided to produce a set of rules that took into account a number of essential variables and parameter to be introduced in the file to be read. Parameters that can be modified include: Duration as Speaking Rate; Intonation from first word marked to a Reset mark; Silence introduced as Durational value; Emphasis at word level increasing Pitch; Volume from first word marked to a Reset mark, increasing intensity.

These variables include essentially the following information:
- the title
- the first verse of the poem
- the last verse of the poem
- a word is the last word of a sentence and is followed by a period
- a word is the last word of a sentence and is followed by an exclamation mark
- a word is the first word of an interrogative sentence
- a word is a syntactic head (either at constituency or dependency level)
- a word is a SUBJECT head – so marked by the dependency parser
- a word marks the end of a verse and is not followed by punctuation
- a word marks the end of a verse and is followed by punctuation
- a word is the first word of a verse and coincides with a new stanza and is preceded by punctuation
- a sentence is a frozen or a formulaic expression with pragmatic content
- a sentence introducing new Topic, a Change, Foreground Relevance as computed by semantics and discourse relations
- a sentence is dependent in Discourse Structure and its Move is Down or Same Level

Here below we show the output of Sparsar for a poem by Chris Wallace-Crabbe, "At the clothesline"

[[pbas 38.0000; rate 140; volm +0.5]at the clothesline . [[slnc 400; rset 0]]
[[pbas 50.0000; rate 140; volm +0.3]what i'd thought a fallen [[pbas 44.0000; rate 130; volm +0.3]]shirt[[slnc 50; rset 0]] under the line . flat on the [[pbas 44.0000; rate 130; volm +0.3]]grass[[slnc 400; rset 0]] was [[rate 110; volm +0.3]]nothing[[slnc 100; rset 0]] but my shadow [[pbas 44.0000; rate 130; volm +0.3]]there[[slnc 300; rset 0]] , hinting that all [[rate 130; volm +0.5]]things pass : that many we loved [[slnc 100]]or used to [[rate 130; volm +0.5]]know[[slnc 50; rset 0]] are dragged already out of sight[[slnc 200]] [[rset 0]] , vanished fast , though stepping [[pbas 44.0000; rate 130; volm +0.3]]slow[[slnc 300; rset 0]] , folded into remorseless [[pbas 44.0000; rate 130; volm +0.3]]night[[slnc 300; rset 0]] .
[[pbas 50.0000; rate 170; volm +0.3]]my dark trace now has quit the [[pbas 44.0000; rate 130; volm +0.3]]lawn[[slnc 300; rset 0]] .
[[rate 110; volm +0.3]]everything[[slnc 100; rset 0]] slips away too soon[[slnc 200]] [[rset 0]] , yet [[rate 110; volm +0.3]]something[[slnc 100; rset 0]] leaves its mark [[rate 130; volm +0.5]]here like a rainbow ring around the [[pbas 36.0000; rate 110; volm +0.5]]moon[[slnc 300; rset 0]] .

7 Conclusion and Future Work

We still have a final part of the algorithm to implement which is more complicated to do and is concerned with Modeling Poetry Reading by a TTS (Text To Speech) system. It is the intermingling of syntactic structure and rhetoric and prosodic structure into phonological structure. What remains to be done is to use syntactic information in order to “demote” stressed syllables of words included in a “Phonological Group” and preceding the Head of the group. This part of the work will have to match tokens with possible multiwords and modify consequently word level stress markers from primary “1” to secondary “2”. A first prototype has been presented in[8], but more work is needed to tune prosodic parameters for expressivity rendering both at intonational and rhythmic level. The most complex element to control seems to be variations at discourse structure which are responsible for continuation intonational patterns vs. beginning of a new contour. Also emphasis is difficult to implement due to lack of appropriate semantic information.

References


---

4 which can be found at the following link: http://www.poetrylibrary.edu.au/poets/wallace-crabbe-chris/the-domestic-sublime-0779005
3. Delmonte R. (1979), Piercing into the Psyche: the Poetry of Francis Webb - with a Select Concordance of Keywords and Keyroots, CETID, Venezia.