

## Re-Weaving the Word-Web: Graph Theory and Rhymes

James Joyce

University of California, Berkeley

### 0 Introduction<sup>1</sup>

My title, "Re-Weaving the Word-Web," refers to the ongoing effort to reconstruct earlier states of language [Joyce, 1977]. The word-web in question is that of sound and semantic associations in either prose or poetry. Today we have a much better idea how Chaucer may have pronounced his words than Dryden did, for example. However, we do not have recordings of the actual speech — such as those made by Tennyson, Browning, and others in the nineteenth century [Tennyson, 1956] — with which to guide our re-weaving of marks on the manuscript page that are the suggestions of sounds into knowledge of how to produce those sounds.

Our knowledge of how an earlier state of a language may have sounded is developed from a series of hypotheses formed from written records or extrapolated from present oral practices [McIntosh, 1956, 1963]. A primary source for evidence of pronunciation is that of end-rhymes in poetry having recognizable rhyme schemes. The traditional technique for using rhyme data has been to gather rhymes into an alphabetized list (a rhyming dictionary) and, with other language data, determine what a group of rhyming words may have sounded like. A rhyming dictionary, such as the recent computer-produced dictionary to Pushkin's rhymes [Shaw, 1974], does not preserve an important aspect of rhyme: rhymes may be linked by one or more words intermediate — that is, by a *chain* of rhyming words. The addition of such features to a rhyming dictionary as a frequency count for pairs of rhymes is helpful, but does not address the issue of preserving the chain that allows one to associate word *a* and word *c* via word *b*.

Drawing from mathematical graph theory and applications of that theory in other fields, we may represent rhymes as nodes in a *directed graph*, or, more simply, *graph*. A line connecting two words indicates that they are linked as rhymes *according to the rhyme scheme of the poem*. Visually, the graph is a web with words embedded in it. This graph may be explored and interpreted using results from graph theory to provide both phonological and semantic insights. An algorithm for guiding collocation of rhyme graph data and other language data will be discussed, with examples given from the Middle English *Pearl*.

### 1 A Brief Summary of *Pearl*<sup>2</sup>

For those who may not know *Pearl* or whose recollection is somewhat hazy, it is necessary to review the salient points of the poem because semantic aspects of the word-web require at least a cursory familiarity with the poem. The North-west Midlands *Pearl* is a late medieval allegorical poem in the form of a debate within a vision. This debate genre began with Boethius' *The Consolation of Philosophy*, and was a popular

<sup>1</sup> The research reported here was aided in part by Sigma Xi, the scientific research society of North America.

<sup>2</sup> This summary is an abridged form of one done by my wife and colleague, Tanya Joyce. She is not responsible, however, for errors introduced by me in summarizing her summary.

literary form throughout the Middle Ages. Just as Boethius is deeply troubled at having been made an exile from his homeland at the beginning of *The Consolation*, the narrator is also grieving at a loss — that of his Pearl — and though he has been taught that he may find comfort through the nature of Christ, his will, unhappy, continues to grieve. The Pearl slipped away from the narrator in an enclosed garden, and the narrator, falling among the sweet spice flowers of the garden, experiences a vision of the Pearl.

In the vision, the Pearl is a radiant maiden who instructs the narrator. Just as Lady Philosophy in Boethius' vision tells the author it is now time for healing and not lamentation, the Pearl tells the narrator (also identified as the Jeweller) that he is mistaken in thinking he has lost his Pearl. The Jeweller and the Pearl engage in a debate, and the concepts of the debate confuse and upset the Jeweller. Distinctions between worldly and spiritual comfort appear to the Jeweller to doom him to sorrow. He is upset also to learn that the Pearl is a bride of Christ, because she died at a very young age and had not (to the Jeweller/Narrator) earned the high rank she now possesses. The Pearl uses the Parable of the Vineyard to instruct the Jeweller that God's grace makes it possible for grace to be shared equally by those who deserve it — no matter how long they have lived according to Christian doctrine, or whether or not they died young. Christ died for the love of man in the worldly city of Jerusalem, and now, in the Heavenly city of New Jerusalem, Christ comes each day to celebrate with music and dancing.

The narrator is shown New Jerusalem, and it is so beautiful that he wants to cross the stream separating him and the Pearl maiden, whether or not this crossing means his death. He struggles to cross, and wakes up in the attempt. Having awakened from the vision, the narrator mournfully accepts the lesson of The Parable of the Vineyard; he will become a servant of God, assuming his correct role in the lowest station in God's spiritual court.

## 2 Recognition of End-Rhymes

The poem *Pearl* is 101 stanzas in length, rhyming a b a b a b a b c b c. This rhyme scheme is easy to recognize from orthographic data in the unique manuscript, Cotton Nero A.x, Article 3, in the British Library and in the various editions of the poem, as can be seen from the first stanza of the poem, in Figure 1.

Perle, plesaunte to prynces paye  
 To clanly clos in golde so clere,  
 Out of oryent, I hardyly saye,  
 Ne proued I neuer her precios pere.  
 So rounde, so reken in vche araye,  
 So smal, so smothe her syde<sub>3</sub> were,  
 Quere-so-euer I jugged gemme<sub>3</sub> gaye,  
 I sette hyr sengeley in synglere.  
 Allas! I leste hyr in on erbere;  
 Thur<sub>3</sub> gresse to grounde hit from me yot.  
 I dewyne, fordolked of luf-daungere  
 Of that pryuy perle wythouten spot.

Figure 1: Stanza 1 of *Pearl* [Gordon, 1953]

It is a basic function of rhyme to make use of a word to recall the final cluster of sounds, typically one syllable's worth, in a word that occurred previously. When such a

relation exists between words both of which occur at the end of a line, we say the words are end-rhymes. Other kinds of rhyme, such as internal rhyme, which abound plentifully in *Pearl*, are beyond the scope of the present investigation, although a complete study of rhyme evidence would need to address itself to other rhymes in addition to end-rhymes.

When we see the poem on the page there is the temptation to perceive the end-rhymes as inter-relating back and forth; that is, the sequence of *a*-rhymes in the first stanza — *paye, saye, araye, gaye* — appear visually to exist simultaneously in the stanzaic unit, and they do *as text*. However, in performing the poem — either by reading or hearing it — we encounter the words and their sounds one at a time, ordered by time itself; that is, we read from the first word to the last, in sequence. The progression of the poem is forward, but the end-rhymes reverse our attention: when we encounter *saye* we recall *paye*, for example. One might argue that after reading or hearing one stanza of *Pearl* we would, upon reading *saye*, anticipate *araye*. However, given the number of possible words that one might anticipate as being rhymed the effect of anticipation must necessarily be a weaker relation than that of recalling a particular, already-encountered word. Indeed, were we able to predict a rhyme word we would feel the poet's craft was lacking, as in some twentieth-century popular song lyrics in which *moon* is all-too-predictably rhymed with *June*.

The *recalls* relation among end-rhymes is one that reverses the normal flow of time in a poem. It is also true that though one does not expect to be able to anticipate a rhyme word, when the rhyme is arrived at in a performance of the poem there is an emphasis on the rhyme because its sound recalls one encountered earlier. That is, recall of a sound from one word has the effect of emphasizing the later-occurring word. Rhyme as an ornament for emphasis is a commonplace of prosody, yet the emphasis takes its strength from the relation of a second word to a first. That is, any emphasis depends upon the relation *recalls* to give it force. The relation *recalls* is the fundamental relation among rhymes, and is thus a reasonable relation to model.

### 3 Rationale for the Rhyme-Pair Model

The rhyme scheme of *Pearl* is self-evidently complex. A model of the rhymes should reflect that complexity if it is to model the relations among the rhymes adequately. The *a*-rhymes of the poem, for example, set up a multiplicity of recalling relations as one perceives each additional rhyme. This multiplicity might be modelled as an *n*-tuple, where *n* is the number of *a*-rhymes; the *a*-rhymes in the first stanza would be the 4-tuple: (*paye, saye, araye, gaye*). Other occurrences of, for example, *paye* might be in other 4-tuples (*a*-rhymes), 6-tuples (*b*-rhymes) or pairs (*c*-rhymes, such as the couplet of the last stanza of the poem). Using such a model would prove awkward mathematically, and it is rejected for this reason without further discussion.

To construct another model we might take each occurrence of rhyme as a relation among two words: that is, the word recalled and the word recalling it. The ordered pairs for the *a*-rhymes of stanza 1 are, then: (*saye, paye*); (*araye, paye*); (*gaye, paye*); (*araye, saye*); (*gaye, saye*); and (*gaye, araye*). This model has the advantage that all rhymes are represented as ordered pairs, and these ordered pairs have a convenient representation, as we shall see shortly, as nodes in a directed graph. The model has what appears to be a disadvantage in that each rhyme pair is represented in the same manner whether the rhymes are in consecutive lines, as are the *c*-rhymes, or are separated by a distance of 8 lines, as is the case of the first and last of the *b*-rhymes. The strength of recollection, intuitively, is greater when the rhymes are couplets rather than when they are separated by a long distance. However, it is not the strength of an individual *recalls* relation but the relation's very existence that is being modelled. The

model captures the phenomenon that one word recalls another in a corresponding rhyming position in the stanza.

There is, unfortunately, at least one knotty problem that is true in *Pearl* which, for most rhymed verse, usually is not the case: the stanza both is and is not the contained unit for rhymes in the poem. There are nineteen sections of five stanzas and one section containing six stanzas, each stanza within a section being linked by the same concluding *c*-rhyme. This relationship complicates matters in that one might venture to say that all the *c*-rhymes in a section thus recall all previous *c*-rhymes in that section. From my own experience in reading the poem both silently and aloud it appears unlikely that *yot* (line 10), which is recalled by *spot* (line 12), is also recalled by *clot* (stanza 2, line 10), *not* (stanza 3, line 10), *flot* (stanza 4, line 10), and *schot* (stanza 5, line 10). What does appear likely is that the last word for each stanza recalls the last word for previous stanzas in the same section. As the word each concluding *c*-rhyme recalls is the same word (admittedly sometimes in different senses of the word) in all but two instances (lines 203 and 912), the amount of rhyme data we risk losing at this stage does not require compromising a cautious approach to the data. Within each section, the concluding *c*-rhyme of the stanza — or the root of the *c*-rhyme — is used in the first line of the next stanza. It establishes the alliterative sound for that first line in most cases and is also carried from the final *c*-rhyme in a section to the first line of the following section. The pattern is consistent with one exception, the first line of section thirteen, line 721, which does not use the final *c*-rhyme (or its root) from the previous stanza (though the word, *ryght*, appears later in the stanza).

#### 4 Rhyme pairs and Adjacency Matrices

One may think of a matrix as a table of data arranged into rows (vertically) and columns (horizontally). There is one row and one column for each distinct rhyme word. Thus an adjacency matrix that corresponds to the *a*-rhymes in stanza 1 of *Pearl* will contain four rows and four columns because there are four distinct rhyme words among the *a*-rhymes: *araye*, *gaye*, *paye*, and *saye*. It is customary to take the left part of the ordered pair as the row indicator, and the right part as the column indicator. Thus for the pair (*saye*, *paye*) we find the row for *saye* and the column for *paye*, and indicate the pair is in the matrix by entering the value 1 in the location down-to-the-row-for-*saye* and across-to-the-column-for-*paye*:

	araye	gaye	paye	saye
araye				
gaye				
paye				
saye			1	

The pairs are entered into the matrix until there are no more to enter. Those entries not having a 1 in them are set to 0, which we interpret to mean that there is not a rhyme pair in the poem corresponding to this combination. The adjacency matrix for *paye* in stanza 1 of *Pearl* is Figure 2.

	araye	gaye	paye	saye
araye	0	0	1	1
gaye	1	0	1	1
paye	0	0	0	0
saye	0	0	1	0

Figure 2: Adjacency Matrix for *paye* in Stanza 1

In the matrix a zero entry for the pair (*paye*, *gaye*) means *only* that *paye* does not recall *gaye* in stanza one, not that the two words do not rhyme! The delicate step from saying one word recalls another to asserting the two words rhyme may, as shall be shown later, be justified by the properties of the graph. These properties, based on the poem, constitute primary evidence for pronunciation and semantic association, thus becoming important aids in a coherent reconstruction of language. The adjacency matrix is an important aggregate of the rhymes preparatory to systematic exploration by transformations of the adjacency matrix and also augmentation of rhyme data by other traditional resources.<sup>3</sup>

The adjacency matrix may be used to assess interesting and useful properties of the rhyme words, though such a small matrix should not be the basis of grand interpretations, but rather serve simply as an illustration that is easy to read, examine, and follow. The rows of the matrix represent recalling words, and columns represent recalled words. One indication of sound associations is the word that is most often recalled. It is, intuitively, the word the poet wishes the reader/listener to recall, to remember. The word most often recalled is the most *popular* rhyme. To find it from the adjacency matrix we note the word having the largest column total. The adjacency matrix for *paye* in stanza 1 of *Pearl* with column totals is Figure 3.

	araye	gaye	paye	saye
araye	0	0	1	1
gaye	1	0	1	1
paye	0	0	0	0
saye	0	0	1	0
Totals	1	0	3	2

Figure 3: Adjacency Matrix with Column Totals

The most popular word is *paye*, a situation deriving from *paye*'s position as the first of the *a*-rhymes in the stanza. Although it is tempting to note that one may view *Pearl* as a poem about the price of salvation and thus *paye* (which has been translated "price" more than once) is significantly placed in the stanza, it is wiser to wait until we have the rhyme graph for *paye* in the entire poem before launching into interpretations or justifications.

## 5 Adjacency Matrices and Graphs

In translating an adjacency matrix into a graph to gain a visual sense of the interconnectedness of the rhyme graph there are several basic guidelines one may follow in drawing the graph:

- 1) The fewest lines should cross;  
a graph with no crossed lines is *planar*.
- 2) The most popular rhyme should be visually near the center of the graph, subject to considerations in point 1.

<sup>3</sup> It is not a difficult task to write a computer program that, using the rhyme scheme as the rule for pairing words, will find the pairs of end-rhymes for the entire of *Pearl*. These pairs may be associated through construction of an adjacency matrix [Harary, 1965; Christofides, 1975], a convenient representation that facilitates analysis of the rhyme graphs and their properties. Construction of the adjacency matrix may (and, to avoid tedium, *should*) be done by a computer program that uses linked lists to implement a sparse array and prints the matrix for all words linked as rhymes with each other using Warshall's algorithm [Pfaltz, 1977; Tremblay, 1976].

Following these guidelines gives the graph for *paye* in stanza 1 of *Pearl*, Figure 4.

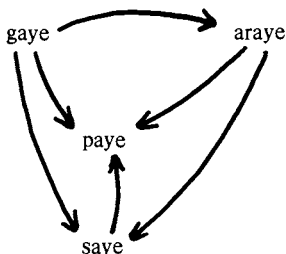


Figure 4: Graph for *paye* in Stanza 1

Having shown in some detail the exact relationships among how one derives ordered pairs from a poem using its rhyme scheme, how the ordered pairs may be entered into an adjacency matrix, and how one may draw a graph from the adjacency matrix, it is appropriate to examine a graph for which the results are more significant than in the simple example of Figure 4. The graph for all rhymes with *paye*, however distantly they are connected, is given in Figure 5.

It is apparent why simpler examples were given earlier; the interconnections here are so great in number that a planar graph for all the rhymes could not be drawn. The principle that the graph should have the fewest crossed lines has, as far as can be determined, been adhered to, and yet there are a great many crossed lines. The most popular rhyme, however, is near the center of the graph, and it is *say(e)*. The spelling *say(e)* means that *say* and *saye* are both represented by the single node labelled *say(e)*. That is, the two forms have been identified as being spelling variations rather than different words or different forms of the same word. Some of the arrows connecting two words have a number associated with them; this indicates the number of times the word *aray(e)*, for example, recalls the word *say(e)*. The adjacency matrix, strictly speaking, does not record multiple occurrences. The reason for this is a matter of definition [Harary, 1965; Christofides, 1975].

## 6 More on the "Most Popular" Rhyme

The wisdom of refraining from a justification for *paye* being the most popular rhyme in stanza 1 is clearer now that the graph for all words rhyming with *paye* in the poem is at hand: the word *say(e)*, rather than the word *pay(e)*, is the most popular word in the graph. Yet the word *say(e)* seems too common a word to occupy a position designated as most popular. Can reasons be found, without stretching for an interpretation, to justify the word in that position? Until now the rhyme graph has been treated as a vehicle for phonological data; it is now explored as a semantic network.

The point made earlier about rhyme as a form of emphasis pertains to the semantic linking of words just as it does to the phonological linking. Throughout the history of end-rhyme it is a test of the skill of the poet to place appropriate words in rhyming positions. This test of appropriateness does not stop at the level of sound, as the earlier example of *moon/June* illustrates: a rhyme may be true yet still clunk. The excellent craftsmanship of *Pearl* is apparent enough that any doubt about the placement of words in the poem is probably more a reflection of our own lack of ability to perceive than the poet's to create — a principle that may be generalized to other works of art readily. This is not to say that the poet could not have put *say(e)* in a position to be recalled often because of the need for a rhyme, or that the poem is a narrative and thus there is

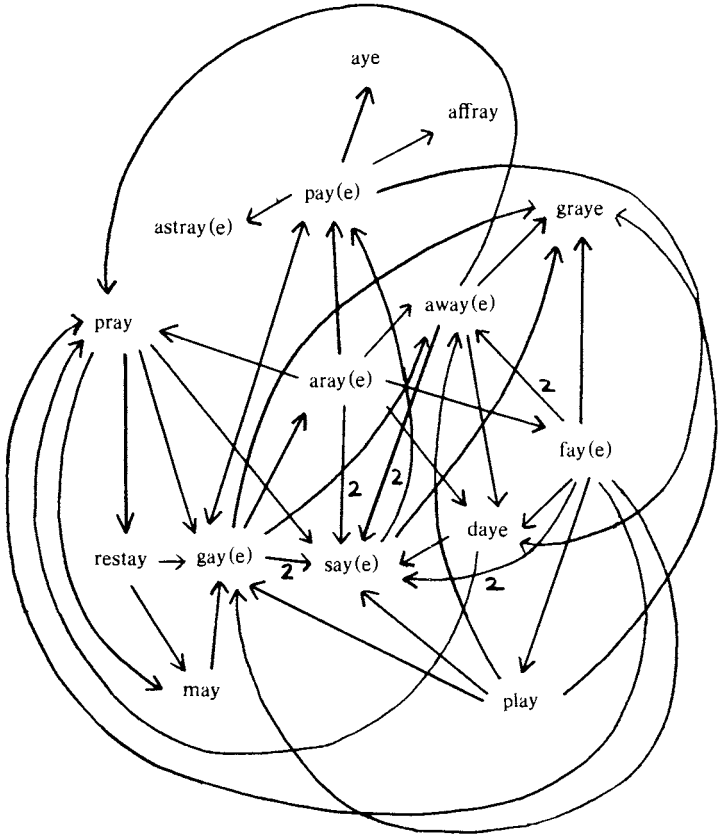


Figure 5

quite a bit of saying; however, such reasons should be given as a last resort rather than as a first impulse.

In another large graph in the poem, that for the word *jueleliueleliuelle* ("jewel" in its variant spellings), the most popular word is *mel(l)e* ("speak, tell"); see Figure 6 for this graph. I have argued elsewhere [Joyce, 1977] that evidence from various rhyme graphs and *The Middle English Dictionary* lead me to consider *mele* and *melle* as spelling variants of the same word.<sup>4</sup> The graph for "jewel" is important for several reasons: its size, as it is among the largest for the poem; also, the jewel is a direct reference to the Pearl-maiden, the lost loved one who is a bride of Christ and upon whom the narrator of the poem focusses his attention. The most popular word is again another word for speaking; this might be just a coincidence, but there are reasons in the tradition from which the *Pearl* comes that provide a unifying explanation.

As indicated in the summary of *Pearl*, above, the narrator and the Pearl-maiden engage in a debate. The debate form in medieval literature was an important vehicle through which mankind could understand God's truth [Piehler, 1971]. That is, through the saying and speaking that is debate God's truth is knowable to mankind. Thus the early emphasis on *Pearl* as an elegy in late nineteenth century criticism is misplaced, as is the later emphasis on the poem as an allegory [Schofield, 1904]. The emphasis on synonyms for speaking, as shown in the rhyme graphs, points to the poet's central concern with the debate process of the poem.

But again the question of the need for a rhyme or the need for words such as "he said" and "she said" in a narrative with dialogue needs to be examined. If we study rhyme graphs for the *Pearl*-poet's contemporary, Geoffrey Chaucer, we find that in *The Knight's Tale* (which contains approximately the same proportion of dialogue as *Pearl*) the large graphs do not have as the most popular rhymes synonyms for speaking, but rather the names of characters — Arcite, Emily, and Palamon, and so on. Chaucer's concern for character is well-known, and is reflected in the rhyme graphs. The *Pearl*-poet does not ignore character, but the rhyme graphs emphasize the process for knowing God's will, the debate, rather than the individuals engaged in the debate. This attitude on the part of the *Pearl*-poet may strike us as properly medieval, just as Chaucer's emphasis on character may strike us as Renaissance-like — a term used to describe Chaucer's art more than once.

There is not space here to answer all objections and counter-proposals to my interpretations of the reason certain kinds of words appear in rhyme graphs for the *Pearl*-poet and in Chaucer. The purpose of giving these examples is to introduce the concept that rhyme graphs as semantic networks function as plausible structures for interpretation. The lines that connect any two words in a graph, we must remember, were put there by the poet; the associations, therefore, are surely his.

The most popular rhyme in a rhyme graph, as has been mentioned several times, is the word most often recalled by words in that graph. The rhyme sound in any recalled word is, for the pair of rhyme words, the *true* rhyme, as it occurs first and establishes the expectation for what the sound will be; the second word is or is not a true rhyme, depending upon the aesthetic for rhyming operative for the verse of that particular time. The most popular rhyme in a rhyme graph, then, is the *true rhyme* for the rhyme graph; another way of putting it might be that the most popular word is the *best embodiment* of the rhyme sound the graph represents. All words in the rhyme graph that, given the poet's aesthetic of rhyme, share the most popular rhyme's sound

<sup>4</sup> Of the two editions of the poems that have been published since my 1977 article was written, one does not address the question [Moorman, 1977], but the other appears to agree with me in that the glossary gives "mel(l)e" rather than a separate entry for each [Andrew, 1978].

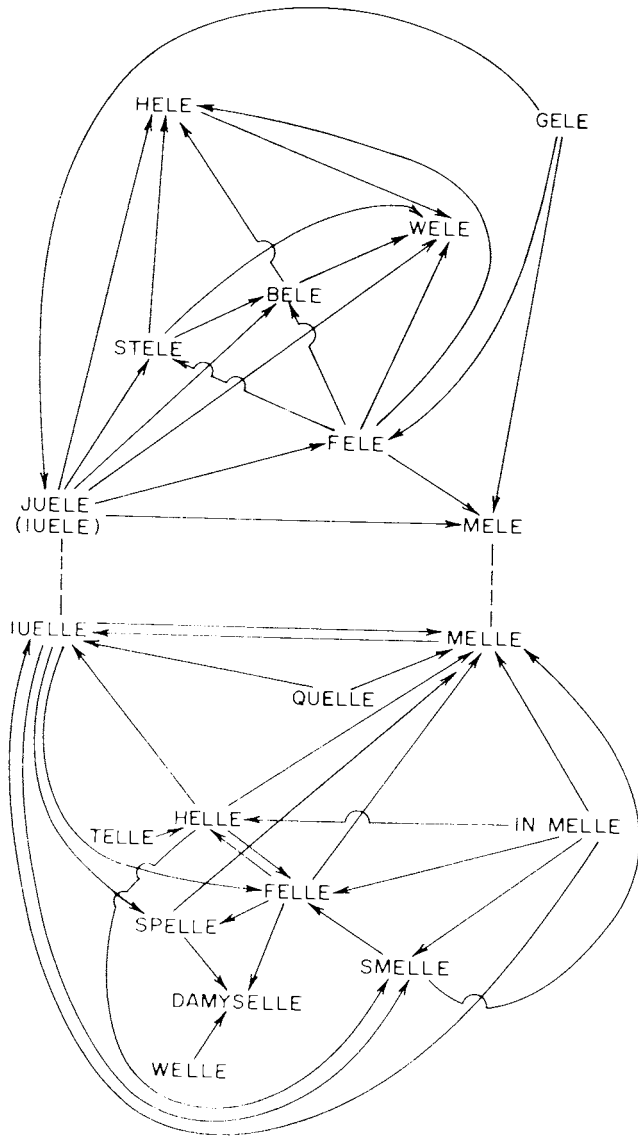


Figure 6

are members of the set of *true rhymes* for the graph. These definitions arise from the graph model of rhyme relationships, and provide precise definitions for literary phenomena.

Such definitions clearly acknowledge that words may appear in a rhyme graph because of their association by a rhyme scheme and yet not be true rhymes. One technique of rhyming is to use a rhyme scheme established earlier in the poem to link two words that do not sound alike and thus to achieve slant rhyme or eye rhyme. Were the rhyme graph model to assert that all words in any rhyme graph are always true rhymes the model, though succeeding often, would be seriously flawed and its usefulness limited. There is a line of reasoning, however, that does allow a hypothesis about which rhymes are likely to be true rhymes, based upon a particular property of the rhyme graph: its *strong components*.

## 7 Strong Components in Rhyme Graphs

A *strong component* is a path (obeying the arrows as though they were one-way streets) from a word  $w$  through a series of words returning at last to  $w$ . In the graph for *pay(e)* in all of *Pearl*, the pair (pay(e), gay(e)) and also (gay(e), pay(e)) exist, and thus there is a path from pay(e) to gay(e), and a path from gay(e) to pay(e); this path is one *strong component*. Remembering that the pairs are made up of words associated by the rhyme scheme in the relation *recalls*, we interpret the strong component as *pay(e)* recalls *gay(e)* and *gay(e)* recalls *pay(e)*. Each recalls a sound in the other, a situation that strongly implies that the sound recalled in each pair is the same sound. *Strongly implies* is not the same as *proves*, however. Indeed, a graph (or any other) model of a relation does not *prove* facts about the relation, but rather *implies* the facts exist — if the model is a valid one. In graph theory it is common to treat strong components as if they condensed to a single point. It would be absurd to say that the condensation of *gay(e)* and *pay(e)* means that the two words are the same; rather, it is meaningful to say that they *rhyme*; thus the relation *recalls* provides a means of predicting the relation *rhymes*.

More generally, all words in a strong component are rhymes with each other. The graph for *pay(e)* is so interconnected that only the words *astraye*, *aye*, *affray*, and *graye* are not part of the largest strong component that includes *pay(e)*. All words in a strong component are also *true rhymes* with each other in that no matter which word in the strong component is the most popular word within that strong component, the other words in the strong component rhyme with it. Thus they are all *true rhymes* with each other. If the strong component includes the most popular rhyme for the graph, as defined above, the *true rhyme* for the graph, all words in the strong component are *true rhymes* for the graph.

The concept of what is a *true rhyme* and what is not depends upon the aesthetics of the writer and the time in which the verse is written. Thus if the language of a particular time is known, its aesthetic of rhyme may be established by determining the degree of conformity of sound between words in the same strong component. The aesthetic is thus expressible precisely. Comparative studies of a particular aesthetic of rhyme may be done either synchronically to establish a region or nation's aesthetic (and language) at a particular time, or diachronically to study the change over time.

Words not in a strong component are simply words connected in the graph; inference about their being true rhymes or other kinds of rhyme (slant, eye, etc.) requires further study of the graph's structure and, of course, external evidence as well. Exploring the graph's structure requires some means of exploration, and that means is obtained through interpretation of mathematical properties of the graph, as has been done in several examples earlier. The next aspect of the rhyme graph model is a

method for determining words to study as possible rhymes using *external* evidence that will realize the greatest contribution to further interpretation using the rhyme graph model.

## 8 An Algorithm for Scholarship

Strong components in a graph are important in that the model predicts that all words in a strong component will be true rhymes. Thus we want to find the largest number of words that are in the same strong component.<sup>5</sup> It would make matters even easier if the rhyme graph were one large strong component. The graph for *pay(e)*, Figure 5, comes close in that only *astraye*, *aye*, *affray*, and *graye* are not part of the large strong component in the graph. The graph for *juele*, Figure 6, is a different matter, however, in that the words in the strong component are *juele*, *fele*, *mel(l)e*, *helle*, *felle*, and *smelle*. The great majority of words in the graph for *juele* are not in a strong component, and will require evidence external to the rhyme graph to identify them as true rhymes.

However, there is a way of reducing the number of words that need to be examined using evidence external to the poem to identify them as true rhymes. Since words in a strong component are considered true rhymes, there should be a method for finding the two words that, if identified as true rhymes, and thus connected with a link joining them in the rhyme graph, will add the largest number of nodes to the largest strong component in the rhyme graph. This problem involves calculating the *marginal complexity* [Rhodes, 1973; Kahn, 1977] of the graph. As Kahn succinctly puts it, "intuitively, marginal complexity measures the greatest change in complexity [in a graph] when an arrow is added from a given vertex [word] to any other vertex [word] in the graph." The concept of a graph's complexity, as defined by Rhodes, is  $CX(G) = n - 2$ , where  $n$  is the number of words in the graph's largest strong component.<sup>6</sup>

In Figure 6, the greatest change in complexity may be realized in four different ways. That is, there is a tie among four pairs of words as to the number of words they will add to the strong component. These pairs are (wele, gele), (wele, in melle), (wele, telle), and (wele, quelle). To take one such pair, (wele, gele), the words added to the strong component are *stele*, *bele*, *wele*, *hele*, and *gele*. The other pairs that tie for the number of words they will add to the strong component add the same first four words as *gele*, adding themselves just as the pair (wele, gele) adds *gele* to the strong component. It is apparent that *wele* is the crucial word here, as it is the first word in every pair.<sup>7</sup>

Knowing the rhyme pairs for which marginal complexity in the associated rhyme graph is greatest informs us where to direct scholarly activity to realize the greatest return in results. If scholarship provides reasons for considering the words *wele* and *gele* as true rhymes that arrow should be added to the graph and marginal complexity computed again. The rhyme pair that yields greatest marginal complexity again informs the scholar where to direct scholarly activity to realize the greatest return in results. This procedure may be followed until all words in the graph are in a strong component. Thus calculating marginal complexity in a rhyme graph provides an algorithm for scholarship.

<sup>5</sup> A strong component may be composed of several smaller strong components, or two strong components having only one word in common (figure-8 style), and still be a strong component.

<sup>6</sup> For anything but the simplest graphs calculation of marginal complexity can best be done by computer. A program in Pascal is available that computes marginal complexity.

<sup>7</sup> The fact that *wele* is a reference to the Pearl-maiden should not be missed here, though interpretation of the semantic associations must await another writing.

If a particular pair of words yielding greatest marginal complexity prove not to rhyme, based on scholarly use of external evidence, delete the word in the pair most likely to be a slant rhyme from the graph and recalculate marginal complexity. If the most likely candidate of the two cannot be determined, delete one of the words and recalculate; then restore it to the graph, delete the other word, and recalculate. Two graphs result, and two marginal complexity values. Direct scholarly activity toward the graph having largest marginal complexity. Note that because the pairs of words associated by marginal complexity were not in the strong component for the original graph, and because we remove only one word at a time, we lose the least amount of evidence for pronunciation in trying to determine which words are true rhymes through this technique.

## 9 Conclusion

The rhyme graph model is an improvement over the traditional treatment of rhyme evidence as an alphabetized list of words. The model encourages both phonological and semantic interpretations of its properties because it preserves the interrelations among words established by the use of rhyme. It is possible to use the rhyme graph's structure to guide scholarly activity to associate data in the graph with external evidence for pronunciation.

Future uses of the rhyme graph model are many and varied. They include the combination of rhyme graphs for poetry of a particular dialect region to construct a model of the dialect. They also include study of rhyme graphs for a particular word from poetry written at different times to determine changes in pronunciation over time. These phonological studies have their semantic counterparts as well, the interpretation of which may bring new insight into the poetry of a time or place.

The rhyme graph model, like any other model, is not fool-proof in that the model provides the evidence for interpretation, but not the interpretation itself. Fortunately, the model may be studied using a computer to perform the more time-consuming and tiresome calculations.<sup>8</sup> Machine-aided analysis also makes it feasible to duplicate the research reported here using other poets and poetry to test whether the arguments for generality indeed hold. Early efforts along these lines confirm what has been reported here, and there is much room for further refinements in interpretation through further exploration of results in graph theory.

## Bibliography

- [Andrew, 1978] Malcolm Andrew and Ronald Waldron, *The Poems of the Pearl Manuscript*, London: Edward Arnold, 1978.
- [Christofides, 1975] Nicos Christofides, *Graph Theory: An Algorithmic Approach*, NY: Academic Pr., 1975.
- [Gordon, 1953] E. V. Gordon, *Pearl*, Oxford: Oxford Univ. Pr., 1953.
- [Harary, 1965] Frank Harary, Robert Z. Norman, and Dorwin Cartwright, *Structural Models: An Introduction to the Theory of Directed Graphs*, NY: John Wiley & Sons, Inc., 1965.

<sup>8</sup> Credit for the various programs used in this study go to students past and present, whose observations and efforts speeded research time considerably. The individuals are (in alphabetical order): Cayford Burrell, Dennis Caswell, Renée Stoddard, and David Wilkes. A complete list of acknowledgements would exceed the space available here.

- [Joyce, 1977] James Joyce, "Networks of Sound: Graph Theory Applied to Studying Rhymes," in *Computing in the Humanities*, eds. Serge Lusignan and John S. North, Waterloo: University of Waterloo Pr., 1977, pp. 307-316.
- [Kahn, 1977] Edward Kahn, "Semantic Marginal Complexity: A Model for Figurative Meaning in Literature," *Poetics*, 6 (1977), 27-40.
- [McIntosh, 1956] Angus McIntosh, "The Analysis of Written Middle English," *Transactions of the Philological Society*, 1956, 26-55.
- [McIntosh, 1963] Angus McIntosh, "A New Approach to Middle English Dialectology," *ES*, 44 (1963), 1-11.
- [Moorman, 1977] Charles Moorman, *The Works of the Gawain-Poet*, Jackson, Miss.: Univ. of Mississippi Pr., 1977.
- [Pfaltz, 1977] John L. Pfaltz, *Computer Data Structures*, NY: McGraw-Hill, 1977.
- [Rhodes, 1973] John Rhodes, "Axioms of Complexity," *Advances in Mathematics*, 11 (1973), 210-214.
- [Schofield, 1904] W. H. Schofield, "The Nature and Fabric of *The Pearl*," *PMLA*, 19 (1904), 154-215.
- [Shaw, 1974] J. Thomas Shaw, *Pushkin's Rhymes: A Dictionary*. Madison: Univ. of Wisconsin Pr., 1974.
- [Tremblay, 1976] Tremblay, Jean-Paul and Sorenson, P. G. *An Introduction to Data Structures with Applications*, NY: McGraw-Hill, 1976.
- [Tennyson, 1956] Sir Charles Tennyson, "The Tennyson Phonograph Records," *British Institute of Recorded Sound Bulletin*, Winter, 1956, 2-8.