

© Mannon & Schockey (eds.) *To Homer*  
Juse Lehiste (Fests, 1987).

## Patterns First, Exceptions Later

Arnold M. Zwicky, *Ohio State University, Columbus, OH., U.S.A.*

†Ann D. Zwicky

### 0. INTRODUCTION

The first moments of a piece of music are usually enough for the listener to identify its key and meter and to establish expectations about its melodic, harmonic, and rhythmic structure. The beginning of a piece normally gives the pattern against which later material will be interpreted; if it does not, the listener's task is greatly increased. We have speculated that the same is true of poetry, and for the same reason.

Perhaps consciously crafted poetry, especially poetry intended for the eye as much as for the ear, can afford to make great demands on its audience, but most poetry will be more considerate: we expect that, in Fussell's words (1965, 38) 'the poet establishes regularity only to depart from it expressively', and we expect that a tendency towards regularity at the beginning of a poem will be strong in verse forms that are usually transmitted orally – in folksongs, nursery rhymes, limericks, nonsense verse, pop music lyrics, and the like. Whatever the practices of 'art poets' might be (we believe that on the whole they are no different from those of folk poets), it is clear that folk verse is created, and is altered in performances and in transmission, according to unformulated canons of what 'sounds good' or 'sounds right'; as rock guitarist Jimmy Page says, 'The way I see it, rock and roll is folk music. Street music. It isn't taught in school. It has to be picked up.' (*Rolling Stone Interviews*, 319). Our proposal is, then, that part of what will 'sound good' in verse and will get picked up on the street is a relatively regular, pattern-setting, opening.

We are **not**, of course, claiming that first lines or first verses of folk poetry are always, or nearly always, perfectly regular, and that deviations cluster densely in material that follows. Rather, our claim is a statistical one, referring to tendencies in the distribution of deviations from abstract patterns.

Our strategy in pursuing this idea has been to examine the rhythmic and rhyming patterns of English folk verse of several kinds. We report here on four studies, two dealing with rhythm and two with rhyme.

## 1. RHYTHM I: FIRST LINES OF LIMERICKS

For our first study we chose the limerick, a form ideal in one respect – its structure is transparent – but less than perfect in another: It is a very short form, just four lines long. Within this small compass, the only comparison we can make is among the four lines. To make this comparison, we need a measure of deviation from the expected pattern.

Our first problem was the expected pattern itself. The limerick is an anapestic tetrameter form, with rests (R) at the ends of lines 1, 2, and 4; these three short lines rhyme, and line 3 (customarily written as two lines) rhymes internally. Deviations from this rhyme scheme are very rare, but deviations from the ideal metrical pattern, which is given in (1),

- (1)      v v / | v v / | v v / | R  
           v v / | v v / | v v / | R  
           v v / | v v / | v v / | v v /  
           v v / | v v / | v v / | R

are so common that it is hard to find a perfect example. Clearly certain variants do not count as irregular or unmetrical. A survey of the first 411 usable limericks in Legman 1964<sup>1</sup> supported our intuition that two types of variants are perfectly metrical: Any line may begin with an iambic foot rather than an anapestic one, as in the examples in (2) below, and extra trailing syllables may occur freely at the end of any line, so long as they are matched elsewhere as required:

- (2)      a. Thěre wás ř yóung ládř řf Árděn  
                   [#2: 1 trailing syllable]  
           b. Thěre wás ř yóung ládř řf Éxětěř  
                   [#31: 2 trailing syllables]  
           c. Whěle Í, wř mř úsuřl ěnthúřřšřšř  
                   [#159: 3 trailing syllables]

Of the first 80 lines in Legman, 42 have iambic first feet. In these same 20 limericks, 14 have trailing syllables in lines 1-2-4 and 8 have trailing syllables in line 3. Neither variant can possibly be seen as unusual or deviant. What remain as deviations are iambic non-first feet, or spondaic (supershort) feet in any position, or extra syllables within any foot, or an actual foot occurring where a rest is expected.

Our next problem was to scan the limericks. Wherever possible, we tried to press the texts into a nondeviant scansion, preferring to err on the side of metrical regularity, even when that called for somewhat unnatural stressings.

Of the 411 limericks examined, 173 were deviant by our definition (and 238 nondeviant). The distribution of deviant lines was as follows:

(3)	line 1 deviant:	40
	line 2 deviant:	40
	line 3 deviant:	136
	line 4 deviant:	50

(These figures sum to 266, which is larger than 173 because some limericks have more than one deviant line in them.)

It is obvious from (3) that the locus of deviations is line 3; it is not so much that lines 1 or 2 are unusually regular as it is that line 3 is exceptionally irregular. As it turns out, the most prominent peculiarity in line 3 is an iambic third foot, that is, a short foot in a position that is **written** as line-initial, as in

- |     |    |                                                               |
|-----|----|---------------------------------------------------------------|
| (4) | a. | Shě rěplied, “Přn mý sóul,<br>Yřur’re in thě wrřng hóle [#10] |
|     | b. | Břt thě kíd wäs sö tíght,<br>Ănd ít wäs dĕep níght [#96]      |

Our intuition is that short feet in this position are not perfectly regular. Let us nevertheless remove the 81 instances of iambic third feet in line 3 from the distribution in (3) and provisionally count them as nondeviant:

(5)	line:	1	2	3	4
	number deviant:	40	40	55	50
	number nondeviant:	133	133	118	123

The difference in the proportion of deviant and nondeviant examples from line to line is not significant ( $\chi^2 = 4.98$ ,  $p = .17$ ). However, it is clear from the raw numbers in (5) that what is important is not so much the line-to-line comparison (lines 1 and 2 are identical, lines 3 and 4 quite close to one another), but rather the split between the first half of a limerick and the second. When lines 1 and 2 are lumped together, and lines 3 and 4 as well, and the numbers are adjusted for deviations in more than one line within the same limerick, the distribution in (6) results:

(6)	half:	first	second
	number deviant:	61	82
	number nondeviant:	112	91

Our first impression from (5) that lines 1 and 2 taken together are more regular than lines 3 and 4 taken together is confirmed by the statistics for (6):  $\chi^2 = 5.256$ ,  $p = .02$ , which is to say that the distribution in (6) would occur by chance only about 1 time in 50.

We should also point out that throughout our data collection in this study, the ranking of limerick lines with respect to the occurrence of deviations was quite stable:

(7) line 1  $\leq$  line 2 < line 4 < line 3

This pattern – little irregularity at the beginning of a poem, increasing throughout until near the end, then dropping off some at the end – is one we have found throughout our research on the distribution of rhythmic irregularity in verse. It is, of course, the familiar curve of dramatic tension, tension in the case of verse being manifested as deviation from an abstract pattern. What is remarkable is that the curve is detectable even in a form so short as the limerick. The general effect is well known for poetry – ‘Another way in which meter can be exploited to strengthen [poetic] closure is as a return to a norm after a deviation’ (Smith 1968, 44) – and has been recognized to co-occur with high initial regularity in certain musical styles and forms: ‘The stability and clarity of the opening and closing pages of a classical sonata are essential to its form, and they make the increased tension of the middle sections possible’ (Rosen 1971, 70).

## 2. RHYTHM II: FIRST VERSES OF FOLKSONGS

Now we turn to a situation more obviously suited to our scheme of patterns first, exceptions later: the rhythmical patterns of folksongs, with entire first verses compared to later ones. Our data here were taken from Lomax and Lomax (1947). This collection has music (arranged by Charles and Ruth Seeger) associated with the texts, an important aid to us in scansion.

We should emphasize at the outset here that there is no *a priori* reason to expect significant rhythmic organization in these songs. The music supplies foot organization, usually with a very clear and simple beat, so that any number of words, with any sort of accent patterns, **could** run over this beat (subject to the limitations of articulation speed and breath capacity). The four-beat lines of nursery rhymes provide such a strong frame that quite extraordinary deviations from the abstract pattern are possible, as in the celebrated second line of ‘There Was an Old Lady Who Lived in a Shoe’: ‘She had so many children she didn’t know what to do’, with four different foot lengths in a single line:

(8)     v / | v v v / | v v / | v v v v /

But even this nursery rhyme has a discernible dominant foot type (the impression given by the first line, that the verse is anapestic, is borne out by the remaining lines), and the folk songs we have examined similarly exhibit patterns against which particular lines can be judged as more or less regular in their rhythms.

Our data were the first 58 usable<sup>2</sup> songs in the Lomax collection. As before, we scanned lines as generously as possible. In this case our initial problem was to assign some **numerical value** to each verse in a song, some measure of its closeness to or deviation from an abstract rhythmic pattern associated with the song. Then we needed to measure how different a given verse was from other verses in the same song; here our goal was to end up with a way of **comparing** verses of different lengths in songs of different lengths with different overall degrees of variation in them.

We will illustrate these operations with some verses from 'Sourwood Mountain' (#24 in the Lomax collection). Here is the first verse:

(9)     Chickens a-crowin' on Sourwood Mountain,  
           Hey-ho dee-iddle-um-day,  
           Chickens a-crowin' on Sourwood Mountain,  
           Hey-ho, dee-iddle-um day,  
           Call up yore dogs and let's go a-huntin',  
           Hey-ho, dee-iddle-um-day.

This is in fact reducible to a two-line verse, with the first line repeated exactly and with a one-line refrain following each verse line. We first extract the two main verse lines from each verse; here are the main lines from the first, third (a representative middle verse), and eighth (the last):

(10)    a.   Chickens a-crowin' on Sourwood Mountain,  
               Call up yore dogs and let's go a-huntin'.  
           b.   My true love is a blue-eyed daisy,  
               Ef I don't git her I'll go crazy.  
           c.   Ducks in the pond, geese in the ocean,  
               Devil's in the women if they take a notion.

The distribution of poetic accents in these three verses is

(11)    a.       / v v | / v v | / v | / v  
               / v v | / v | / v v | / v  
           b.       / v | / v v | / v | / v  
               v / v | / v | / v | / v

c. / v v | / | / v v | / v  
 / v v v | / v v | / v v | / v

There are four beats per line, and the dominant foot type is clearly one with an initial accent in the foot, but from these passages it is hard to tell whether the right classification is as trochaic or dactylic; the evidence of the other verses speaks for trochaic here. Other songs are clearly iambic, anapestic, or 'hyperanapestic' (fourth paeonic) v v v /, and are in dimeter or trimeter, but with tetrameter by far the most common pattern.

We then look for departures from a perfectly trochaic scheme that are repeated in the same place in every verse, departures that are shared by all verses and hence clearly belong to the rhythmic pattern of the song. These could be rests, extra leading syllables at the beginning of a line (or extra trailing syllables at the end of a line, for end-accented feet like iambs and anapests), missing syllables in a particular foot of a particular line, or extra syllables in a particular foot of a particular line. In this case there are none. We then list the departures from perfect trochaicity, verse by verse and line by line:

(12) a. e1 e2  
       e1 e3  
       b. e2  
           t  
       c. e1 s2 e3  
           ee1 e2 e3

Here 'e1' means an extra syllable in the first foot, 'ee1' two extra syllables in the first foot, 's2' a short second foot, and 't' an initial trailing syllable.

At this point in our analysis of limericks, we had a clear notion of the range of allowable departures from the metrical pattern. But each folk song is, in effect, a new verse form. It seems clear to us that in 'Sourwood Mountain' an extra syllable, particularly in foot 3, is an acceptable departure, but that a short foot is a real rhythmic irregularity. However, we cannot rely on such intuitions (even if they were always as clear as in this case), though we do need a way of discounting some departures from the metrical grid. Our (admittedly crude, but workable) scheme was to discount any departure that appeared, in a particular foot in a particular line, in more than half the verses of a song. In 'Sourwood Mountain,' this means in five or more verses, since the song has eight verses. Only one departure meets this (rather stringent) criterion, e3 in line 1, that is, an extra syllable in the third foot of the first line of a verse; e3 in line 1 appears in verses 2, 4, 5, 7, and (see (12c) above) 8.

We are now ready to assign an index of rhythmic irregularity to each verse. As with the limericks, we mark each syllable deviating from the metrical grid

(but not those exemplifying departures allowed by our criterion); the index for a verse is the number of marked syllables in it. For (12a), the index is 4; for (12b), 2; for (12c), 6 (the ee1 in line 2 represents two extra syllables and counts as 2, while the e3 in line 1 doesn't count at all). For the eight verses, the indices are as follows: 4, 3, 3, 1, 2, 1, 4, 6.

Obviously, neither the first nor the last verse of 'Sourwood Mountain' is particularly regular, given the indices associated with the other six verses. Remember that we are not claiming that first verses are **always** more regular than others, only that first verses are – significantly more often than not – more regular than others; we expect to see our expectations disappointed on occasion, as they are in 'Sourwood Mountain'.

Now we need a measure of **how** irregular the first verse (or any other) of a song is, in comparison to the other verses in its song. There are a number of such metrics. We could, for instance, use rank order: In 'Sourwood Mountain', verse 8 is the most irregular, verses 1 and 7 are tied for second, then follow verses 2, 3, and 5, and finally come the most regular verses, 4 and 6. But rank order gives us no easy way to compare the first verse of this eight-verse song with the first verse of a four-verse or thirteen-verse song, and rank order tells us nothing about the relative degree of irregularity of a verse within its song; clearly, second rank means one thing if your verse has index 4 and the more regular verses have indices 2, 2, 1, 2, 1, and it means something quite different if your verse has index 10 and the more regular verses are all perfectly regular, with index 0 in each case. What we want is a measure of how far a given verse is from the irregularity of the average verse in its song.

This is the z score, the number of standard deviations a given index is from the mean of the indices in its set. The mean of the indices in 'Sourwood Mountain' is 2.750, and the standard deviation (a measure of how far the items vary about the mean) is 1.639. The first verse, with its index of 4, is 0.763 standard deviations above the mean. The verses with an index of 1 are 1.068 standard deviations below the mean. The z score for verses 1 and 7 of 'Sourwood Mountain' is then +0.763, and for verses 4 and 6, -1.068.

Our hypothesis that first verses of folk songs are more regular than later ones now translates into the hypothesis that first verses tend to have negative, indeed significantly large negative, z scores. We have calculated the z scores for the first verses of all 58 folk songs, for their third verses (taking the third verse as a representative middle verse), for their last verses (recalling that we expect some tendency towards regularity at the end of a poem), and (for those 24 songs that had them) for their refrains or choruses (since these follow the verses and are repeated many times, they presumably can be irregular with impunity). The mean z scores for each are reported in the third column of (13) below.

(13)	verse	$n$	$m = \text{mean } z$	$zM = m\sqrt{n}$
	first	58	-0.269	-2.049
	third	58	+0.073	+0.556
	last	58	-0.061	-0.465
	refrain	24	+1.158	+5.673

Each row in (13) represents a sampling of  $z$  scores from a large set of them. The full set of  $z$  scores has a mean of 0, and the mean of a sample of  $n$  scores from the set has a standard deviation of  $\sqrt{1/n}$ . These facts permit us to calculate how far the mean  $z$  scores in the third column of (13) are from the mean of all the  $z$  scores, that is to calculate for each another  $z$  score, the 'z of the mean'  $zM$ , given in the fourth column of (13). The  $zM$  figures tell us how irregular the first verses (third verses, last verses, refrains) are as a set. The score of -2.049 for first verses is significantly negative ( $p < .05$ ), showing that there is indeed a statistical tendency for first verses to be more regular than other verses.

The  $zM$  figures in (13) have third verses comfortably positive (more irregular than the others) and last verses comfortably negative (more regular than the others), though in neither case do the figures (about half a standard deviation off the mean of 0) reach statistical significance. Refrains, however, are astronomically irregular, to judge from the  $zM$  of +5.673 in (13).

We conclude that 'Sourwood Mountain' is in fact not a representative folk song in its rhythmic organization. 'Home on the Range' (#62 in the Lon ax collection), on the other hand, is just about perfect: its seven verses have indices 1, 5, 3, 4, 1, 2, 2, in order, and its chorus has index 5. (The pattern is four lines of anapestic tetrameter, with a rest in place of the fourth foot in lines 2 and 4, and with a discountable short first foot in line 1, as in the very first line, 'O give me a home, where the buffalo roam':  $v / | v v / | v v / | v v /$

### 3. RHYME I: FIRST VERSES OF ROCK LYRICS

The types of verse we have so far considered are on the whole very regular in their rhymes, whatever their rhythmic complexities. Occasionally there is an imperfect rhyme or two, but limericks and folk songs are not very interesting in this respect. In contrast, some rock lyricists use imperfect rhyme extensively. In this section we investigate the distribution of imperfect rhyme in the lyrics of the Beatles (Lennon and McCartney, Starkey, and Harrison on *The White Album*) and Bob Dylan (as published in a 1974 Warner Bros. collection of music and lyrics).

From a previous study (Zwicky 1976), we know that deviations from perfect rhyme in Beatles and Dylan lyrics are almost all analyzable as belonging to a restricted system built on two simple types of half rhyme:



**feature rhyme**, in which segments (consonants or vowels) differing minimally in phonetic features count as rhyming; and **subsequence rhyme**, in which a consonant counts as rhyming with zero. The most frequent feature rhyme, *m* matched with *n*, and the most frequent subsequence rhyme, *d* matched with zero, are both illustrated in (14) with lyrics from Dylan's 'Like a Rolling Stone'.

- (14) a. feature rhyme (*m-n*):  
           Once upon a *time*  
                     You dressed so *fine*
- b. subsequence rhyme (*d-Ø*):  
           You used to be so *amused*  
           At Napoleon in rags and the language that he *used*  
           Go to him now, he calls you, you can't *refuse*  
           When you got nothing, you got nothing to *lose*

Not infrequently, compound types occur as well, as in the frequent half rhyme *time-mind*, which can be analyzed as a compound of the feature rhyme *m-n* and the subsequence rhyme *d-Ø*.

As before, our analysis begins with a determination of the abstract pattern associated with a song and then proceeds with an assignment of indices to each verse and a statistical comparison of first verses with the others. Of the *White Album* lyrics we analyzed 18, and from the Dylan collection we analyzed 59 songs.<sup>3</sup> The Beatles lyrics and most of the Dylan lyrics use ballad rhyming schemes which are unmistakable in their simplicity (as in the first verse of 'Back in the U.S.S.R.' (Lennon-McCartney) in (15) below), but some of Dylan's schemes are startlingly complex (see the first verse of 'Like a Rolling Stone' in (16a) below; the rhyme scheme for the song is given in (16b)).

- (15) Flew in from Miami Beach, B.O.A.C. (A)  
       Didn't get to bed last *night* (B)  
       On the way the paper bag was on my *knee* (A)  
       Man I had a dreadful *flight* (B)
- (16) a. Once upon a *time* you dressed so *fine*,  
       You threw the bums a *dime* in your *prime*,  
           *Didn't you?*  
       People'd *call*, say "beware *doll*  
       You're bound to *fall*," you thought they were *all*  
           *Kiddin' you.*  
       You used to laugh *about* everybody that was hangin' *out*,  
       Now you don't talk so *loud*, now you don't seem so *proud*,  
       About having to be scrounging for your next *meal*.



At this point the statistical methods of the previous section can be taken over straightforwardly. 'Like a Rolling Stone' has verses with indices 3, 11, 7, and 9, giving a mean of 7.500, a standard deviation of 2.958, and a z score for the first verse of -1.521. The z scores for groups of first verses can then be lumped together, and their  $m$  and  $zM$  can be calculated. The table in (17) gives these results for four groups of first verses: those from the five Dylan songs with complex rhyme schemes; those from all 59 Dylan songs; those from all 18 of the *White Album* lyrics; and those for all 77 songs.

(17)	source	$n$	$m$	$zM=m\sqrt{n}$
	a. Dylan complex	5	-1.334	-2.983
	b. all Dylan	59	-0.241	-1.849
	c. all Beatles	18	-0.213	-0.904
	d. all lyrics	77	-0.234	-2.053

These figures indicate a general tendency towards regularity in rhyme in the first verses of the lyrics. The tendency is statistically significant (at the .05 level) for the full set of 77 songs, and it is even more pronounced in the five songs with complex rhyme schemes. That is, the results are entirely in line with our predictions.

We are, incidentally, in a position to compare the results for the first verses with those for last verses, and to compare the earlier results on rhythm in last verses with those on rhyme in last verses.

(18)	source	$n$	$m$	$zM=m\sqrt{n}$
	a. Dylan complex	5	+0.750	+1.677
	b. all Dylan	59	+0.123	+0.945
	c. all Beatles	18	+0.403	+1.708
	d. all lyrics	77	+0.188	+1.653

In (18) we see a consistent general tendency towards **irregularity** in rhyme in the last verses of the lyrics. The tendency approaches, but does not quite reach, statistical significance at the .05 level, and it is **more** pronounced for the complex Dylan lyrics than for Dylan lyrics in general – exactly the opposite of the trend seen in the first verses, and also the opposite of the trend in rhythm shown by the last verses of folk songs. We conclude that rock lyricists do not use increased regularity in rhyme as a device of poetic closure – if anything, the opposite is true – although they do use it for opening pattern-setting.

#### 4. RHYME II: SERIAL ORDER EFFECTS IN ROCK LYRICS

In our final study we consider the order in which half-rhyming segments

occur. So far we have treated the relative order of, say, *m* and *n* as irrelevant, the rhyme *time ... fine* counting just the same as the rhyme *fine ... time*. Now we ask whether our proposal that patterns tend to precede exceptions makes a specific prediction about such cases.

On the one hand, our proposal could be interpreted as suggesting that the **more informative** segment should precede the less informative, *d* before  $\emptyset$ , or in general the more marked before the less marked, *m* before *n*.

On the other hand, our proposal could be understood as predicting that the **perceptually easier** segment should precede the more difficult, in general that the less marked should precede the more marked, *n* before *m*.

Neither of these interpretations is particularly straightforward, and they make opposite predictions. The only empirical evidence we have to bring to the issue concerns child language acquisition and the sequencing of *n* and *m*. A number of children acquiring English have been reported as sporadically or even regularly replacing *n ... m* sequences by *m ... n*; our daughter Elizabeth, who for some months replaced *animal* by *aminal* and *cinnamon* by *cimmanon*, indeed transformed the first of these sequences in novel or non-sense words into the second. It is hard to put much weight on this evidence, however, since it has to do with segment sequencing within words, and we are dealing with segment sequencing within phrases, sentences, or even whole songs.

The upshot of this discussion is that we have no clear prediction to make about serial order effects. And when we examined rock lyrics, we found no such effects.

We extracted all instances of *n-m* and *d-∅* half rhymes in the data collected for the Zwicky 1976 study, using only simple and not compound examples. Recall that these were the two most common half rhymes in that study. The serial order of these segments was as follows:

(19)	<i>n-m</i>	<i>d-∅</i>
	<i>n</i> first: 32	<i>d</i> first: 40
	<i>m</i> first: 29	$\emptyset$ first: 52
	total: 61	total: 92

Neither of these distributions differs significantly from the respective 'expected' distributions, 30.5-30.5 and 46-46 ( $\chi^2 = 0.14$  and 1.56, respectively).

#### ACKNOWLEDGEMENTS

A shorter version of this paper was presented in April 1981, first at Ohio State University and then at the Kentucky Foreign Language Conference. Our thanks to Jane Dewald for her weeks of limerick scanning. This version was

completed while Arnold Zwicky was at the Center for Advanced Study in the Behavioral Sciences; he is indebted to the Spencer Foundation for support and to the Ohio State University for a sabbatical year.

#### NOTES

1. We excluded limericks in languages other than English and those that were deliberately and ostentatiously irregular.
2. We excluded those under four verses in length and a few that resisted easy scansion.
3. We excluded songs with fewer than three verses, those that used perfect repetition rather than rhyme, a few that had no discernible rhyme pattern, and a small number whose rhyme pattern we could not agree on. We did not analyze refrains or choruses in this study.

#### REFERENCES

- Fussell, P. (1965). *Poetic meter and poetic form*. New York: Random House.
- Legman, G. (1964). *The limerick*. New York: Crown.
- Lomax, J. and Lomax, A. (1947). *Folk song U.S.A.* New York: New American Library. Paperback ed. 1975.
- Rolling Stone editors (1981). *The Rolling Stone interviews*. New York: St. Martin's Press.
- Rosen, C. (1971). *The classical style: Haydn, Mozart, Beethoven*. New York: Viking Press.
- Smith, B.H. (1968). *Poetic closure: A study of how poems end*. Chicago: University of Chicago Press.
- Zwicky, A.M. (1976). *Well this rock and roll has got to stop. Junior's head is hard as rock*. Papers from the Twelfth Regional Meeting of the Chicago Linguistics Society, 676-697.