The Role of Phonological Phrasing in Sung and Chanted Verse[1]

Bruce Hayes and Abigail Kaun

This is a preprint version.

Published version (1996): *The Linguistic Review* 13:243-303.

Abstract

This article is a study of the metrics of sung and chanted verse, based on a data corpus of 670 English folksong lines, as well as chanted renditions of the corpus by ten native speaker consultants. Our theoretical focus is on the role of phonological phrasing in metrics. We find that in sung and chanted verse, an especially tight correspondence to the metrical pattern is imposed on linguistic material that is either bounded within a tight phrasal domain or located at the right edge of a high-level domain. These patterns have been observed earlier for spoken verse. But for the phenomenon of metrical inversion, the behavior of sung and chanted verse is quite different from spoken verse. We develop an explanation of the difference, based on the idea that inversion in sung and chanted verse occurs only in those cases where it is the best available metrical option.

A further finding is that sung and chanted verse tends to match the number of beats alloted to a syllable to that syllable's natural linguistic duration. We suggest that the relevant measure of duration is phonetic, not phonological; and that the tendency to phonetic duration matching in sung and chanted verse has a categorial, phonologized analogue in the phenomenon of spoken-verse Resolution.

[1] An anonymous reviewer made a number of very thoughtful suggestions for improving this article. We would like to thank him/her, as well as Patricia Keating, Donka Minkova, Marina Nespor, Gilbert Youmans, and members of the audience at the 1994 Vienna GLOW meeting, who also provided important input. Finally, we would like to thank our native speaker consultants for their patience and thoughtfulness.

1 Introduction

The appreciation of poetry is apparently an invariant property of our species, as Kiparsky (1987) has pointed out. But the poetry found in most cultures is not the kind usually studied by generative metrists. As far as we know, the default state of metrical poetry is to be *sung*, or at least rhythmically chanted, rather than spoken. Homer, the Serbo-Croatian *guslar*s, rap artists, and the singers of Appalachian folk ballads are all exponents of poetry in what is arguably its most natural form.

In singing or chanting, rhythmic beats are realized isochronously, or at least approximately so. Isochronous realization opens up a new world of metrics, with quite different structural possibilities from those found in spoken verse. This article forms part of a general research program (see also Hayes and Kaun, in progress; Hayes and MacEachern, in progress) to employ current ideas in phonology and metrics to achieve a better understanding of sung and chanted verse. The focus of this particular article is on the “correspondence rules” (Halle and Keyser 1971) for English sung verse, and especially the role of phonological phrasing in these rules. In particular, our study will serve as a test case for views set forth in Hayes (1989) concerning the relation of metrics and “p-structure”; that is, phonological phrasing.

The content of this article is follows: we present three background sections, which will cover our basic approach to sung and chanted verse (section 2), the theory of p-structure (section 3), and its relation to metrics (section 5). We also describe our research method (section 4). The remaining sections (6-8) present our empirical results and provide theoretical interpretations of them.

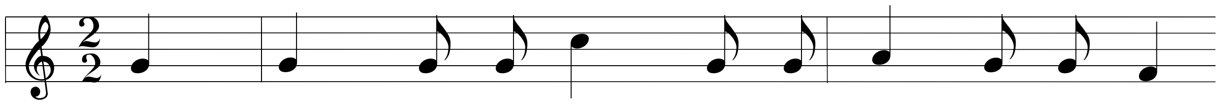
As will be clear below, our work has benefited from earlier research in the area of sung and chanted metrics, notably Liberman (1975), Dell (1975), Stein and Gil (1980), Chen (1983, 1984), Oehrle (1989), and Halle and Lerdahl (1993).

2 Background I: The Study of Sung and Chanted Verse

2.1 Notation

As a formal representation for the rhythm of sung and chanted verse, we adopt the grid notation propounded by Liberman (1975) and by Lerdahl and Jackendoff (1983). In a grid, there are rows of *x*'s or other placemarkers, aligned vertically over one another. Each row of *x*'s is meant to designate a particular interval of beats, which here are intended to be realized at identical temporal intervals. The slower rows are placed over faster ones, and higher columns of *x*'s stand for stronger beats. For example, a folksong line is given in (1a) with its rhythm specified with traditional musical notation. The same rhythm is depicted in grid notation in (1b). In the grid, vertical lines designate the point in time at which each syllable begins.

(1)a.



He róde and he róde till he cáme to the tówn

(Sharp, n.d., p. 9)

```
b.          x               x               x               x
    x       x       x       x       x       x       x       x
    x   x   x   x   x   x   x   x   x   x   x   x   x   x   x   x
    |       |       |   |   |       |   |   |       |   |   |
   He      róde    and he  róde   till he  cáme    to  the tówn
```

Grids are a bit more difficult to read than musical notation, though for our purposes the explicitness they provide is essential. One way to approach reading the grids is first to recite in even rhythm only those syllables linked to a highest-level grid mark ("*rode ... rode ... came ... town*"), then all syllables linked to a mid-level grid mark (*"He ... rode ... and ... rode ... till ... came ... to ... town"*), then the entire line.[2]

Musical notes have durations, so in principle we could designate with the grid where a note ends as well as where it begins. But it appears that for rhythmic purposes it is not particularly crucial where a note ends: note durations mostly determine not rhythm, but rather articulation and musical phrasing. Since these matters are beyond the scope of our article, we adopt the following expedient: we will underscore the grid columns during which a syllable might be expected to be sung, under the assumption that it extends to the beginning of the next syllable. These underscores are meant primarily to make it easier to read the grid; there is no guarantee that the syllable involved actually will be prolonged to the onset of the next syllable. With these underscores, the example in (1) looks like this:

```
(2)          x               x               x               x
     x       x       x       x       x       x       x       x
     x   x   x   x   x   x   x   x   x   x   x   x   x   x   x   x
     -----   -----   -   -   -----   -   -   -----   -   -   -----
     |       |       |   |   |       |   |   |       |   |   |
    He      róde    and he  róde   till he  cáme    to  the tówn
```

2.2 Well-Formedness

Studies of spoken verse (see Hayes 1988 for an overview) have relied heavily on the notion of well-formedness, or **metricality**. A particular linguistic string is said to be a **metrical** instantiation of a rhythmic pattern (that is, a metrical line) if it "sounds right" to the ear of individuals

---

[2] As a further aid to clarity, we have also made a recording containing chanted versions of all of the examples in this paper. Readers who would like a copy of this tape may write to the first author at the address given below, enclosing a blank cassette and a suitable self-addressed envelope. [Note 2000: easier now to visit http://www.linguistics.ucla.edu/people/hayes/metrics.htm#HayesKau.]

experienced in the tradition in which the line was composed. Moreover, different lines are intuitively felt to embody the rhythmic pattern with varying degrees of straightforwardness; these are said to differ in their degree of **metrical complexity** or **tension**. These basic notions are laid out in Halle and Keyser (1971).

The ideas of metricality and complexity carry over fairly straightforwardly to sung and chanted verse. For example, the representation in (2) above may be compared with the one in (3), which involves the very same string of words:

```
(3)           x               x               x                x
      x       x       x       x       x       x        x       x
      x   x   x   x   x   x   x   x   x   x   x    x   x   x   x   x
              |   |           |   |       |   |    |   |   |   |   |
            *He  róde        and he     róde till he cáme to  the tówn
```

The alignment in (3) strikes us as bizarre, and extremely unlikely ever to be used in an actual song. We thus regard (3) as fully unmetrical and (2) as fully metrical. We likewise view the representation in (4) as being possible, but somewhat more complex than (2):

```
(2)          x               x               x               x
     x       x       x       x       x       x       x       x
     x   x   x   x   x   x   x   x   x   x   x   x   x   x   x   x
     |       |       |   |   |       |   |   |   |   |       |
    He      róde    and he  róde   till he  cáme to the     tówn
```

The difference lies in the rhythm assigned to the sequence **cáme to the**. For empirical evidence backing up this judgment, see below.

2.3 Particular Textsettings as the Focus of Analysis

The fact that we used the very same line in (2)-(4) brings up an important point about sung and chanted verse: the metricality of a line is not generally determinable simply from an inspection of its linguistic form. Rather, in order to establish metricality, one must know its actual rhythmic alignment with the grid; what in traditional terms is called its **textsetting**. Strictly speaking, it is only textsettings that may be counted as metrical or unmetrical.

This point deserves some emphasis. It is true, to be sure, that there exist lines linguistically so configured that they could **never** give rise to a well-formed textsetting—for example, any line with a tremendous number of syllables would fit in this category. We could then say that such lines are "unmetrical" in an absolute sense. However, since we have not yet done any research on the question of what would make a sung or chanted verse line "absolutely" unmetrical in this sense, we will limit the term "metrical" here to describe particular textsettings.

In fact, the claim that it is textsettings, not lines, that should be assessed for metricality actually holds true for spoken verse just as much as for chanted verse. The difference is that since spoken verse usually imposes quite rigid restrictions of syllable count, the best "textsetting" (alignment of syllables with the rhythmic grid) will usually be quite obvious.[3]

[3] Here is an exceptionally non-obvious example. Line 6.160 of *Paradise Lost*, "Before thy fellows, ambitious to win," permits two imaginable alignments with the standard grid of iambic pentameter:

2.4 Textsetting Intuitions; Variation

Native speakers of English (and probably other languages) possess fairly well-developed textsetting intuitions, which form the object of our study. We give experimental evidence below that such intuitions are reasonably consistent across individuals.

The use of textsetting intuitions is made more complex when we consider that a given line may often admit more than one textsetting. This is very clear when one examines the recorded rendition of a particular song by different singers. We have also noticed that as individuals, we vary in the details of textsetting when we sing a particular song on different occasions. Our assumption is that such variability usually represents a *spontaneous artistic decision*. To some extent, the singer *improvises* a textsetting, selecting from the well-formed options.

Music publishers often provide the textsetting of only the first verse of a multi-verse song. Presumably, they are relying on the ability of singers to improvise a suitable textsetting for the remaining verses.[4]

```
(i)a.    x        x        x    x       x
       x x    x   x  x     x  x x    x  x
       | |    |   |  |     |  | |    |  |
      Befóre thy féllows, [æmˈbɪʃəs]  to wín

   b.    x        x             x  x       x
       x x    x   x          x  x xx    x  x
       | |    |   |          |  | ||    |  |
      Befóre thy féllows,   [æmˈbɪʃ i əs]  to wín
```

Either alignment is unusual: (i.a) requires a gross mismatch of the stress pattern of *ambitious* to the meter, whereas (i.b) requires the indicated archaic pronunciation of *ambitious*, as well as an extrametrical syllable in the middle of the line. Our guess is that (i.a) represents Milton’s intent, based on which of these odd metrical practices finds clear precedent elsewhere in Milton. Stress mismatches like *ambitious*, though unusual, do occur about 20 other times in *Paradise Lost*. But the hypothesized line-internal extrametrical in (i.b) has no clear precedent at all in *PL*, at least in the view of Sprott (1953, 63) and Evans (1966, 40). Likewise, archaic forms like [æmˈbɪʃiəs] completely lack precedent in Milton’s mature verse, according to Sprott (p. 135).

Irrespective of the outcome in this case, the crucial point is that even iambic pentameters must be analyzed with a particular “textsetting” (or to use the usual term: scansion) in mind.

[4] In light of a comment made by our anonymous reviewer, we wish briefly to address the question of “performance” in textsetting. It is possible that our focus on particular textsettings increases the chance that the data (for example, the data from native speakers described below) will be contaminated by performance errors (see, for example, section 6.9.3). However, on the whole, the reasonableness of most of our consultants’ settings, as well as the systematic patterns found in our data, lead us to think that such contamination does not happen very often.

The ability of the participants in a metrical tradition to obey complex rules “on-line” has been observed elsewhere. In Hausa (Russell Schuh, p.c.), the verse system is quantitative and quite strict, so it is easy for the analyst to detect unmetrical lines. Nevertheless, the participants in this tradition are able to improvise verse lines spontaneously, with remarkably few deviations from correct quantity.

2.5 Textsetting and Singing

Since all of our empirical data consists of sung lines, we wish to address the possibility that textsetting in songs could somehow be made trivial by referring to the musical tune. Note first that we have been able to devise grammars that predict (with a fair level of accuracy) the textsettings of songs given only their linguistic structure, and no information at all about the tune. Second, "pitch-setting" and textsetting very often diverge: multiple pitches per syllable are allowed, and the set of beats on which syllables are initiated can diverge sharply from the set of beats on which pitches are initiated. Very often one finds a rather dull syllable rhythm underneath a relatively interesting pitch rhythm, as in (5).

```
(5)  d       a   b   c   a   d'      c   b   a   b   c   a   d'
     |       |   |   |   |   |       |   |   |   |   |   |   |
     |       |   |   |   |   |       |   |   |   |   |   |   |
     |       x   |   |   |   x       |   |   x   |   |   |   x
     x       x   |   x   |   x       x   |   x   |   x   |   x
     x   x   x   x   x   x   x   x   x   x   x   x   x   x   x   x
     |       |       |       |       |       |       |       |
    And      in    what-     é-     ver     tówn     I      wént
```

(Sharp 1916, #43)

In light of these facts, it seems reasonable to study textsetting as a phenomenon independent of tune, at least as a first approximation. We leave the question of how the tune influences textsetting for further research.

3 Background II: Phonological Phrasing

The main goal of this article is to further our understanding of sung and chanted verse by relating them to formal theories of **phonological phrasing**. As Kiparsky (1975, 1977) and others have shown, it is precisely in their connection with phonological phrasing that the metrical systems for spoken verse usually show their most subtle and interesting aspects, and we should not be surprised if the same turns out to be true for sung and chanted verse.

By "phonological phrasing" we mean here the theory of prosodic domains originated by Selkirk (1978, 1980, 1981) and amplified by, among others, Nespor and Vogel (1982, 1986), Selkirk (1984, 1986), Pierrehumbert and Beckman (1988), Hayes (1989), and various contributors to Zwicky and Kaisse (1987) and Inkelas and Zec (1990).

Here is a brief review. The theory posits that phrasal phonology is governed by a hierarchical constituent structure. This structure has been given different names by different theorists (e.g. "prosodic domains", the "Prosodic Hierarchy"); here, we will use Selkirk's (1986) usefully non-overlapping term "P-structure." Each constituent of P-structure forms a domain of phonological cohesion, and low-ranking constituents form more cohesive domains than high-ranking ones. P-structure is distinct from syntactic structure, although syntax does play a crucial role in determining its form (other possible factors include focus, speaking rate, and constituent length). P-structure is usually assumed to obey severe layering constraints, such that the path of prosodic nodes going from terminal material to the root of the P-structure tree is invariant. (For the particular set of domains assumed here, see (6) below.) Phonological (and metrical) rules may be bounded by particular domains of P-structure, and may refer to the edges of particular domains in their structural descriptions.

A major goal of the present article is to examine various principles that have been proposed for characterizing the relation of metrics to P-structure. We will try to determine to what extent these principles carry over to the domain of sung and chanted verse, and where they don't, we will try to explain why not.

4 Data

Our work is based on a corpus of 670 lines from English folksongs, gathered in fieldwork in the early twentieth century by Cecil Sharp and his colleagues. The material was edited to a fair degree by Sharp, and cannot be regarded as fully authentic. However, for present purposes the editing is all to the good, since it results in the material reflecting rather more uniform textsetting intuitions, namely those of Sharp.[5] The textsettings of the Sharp materials (all cited in the reference section below) will be referred to below as the “original textsettings”.

The songs that we selected all use the rhythmic grid of (1). There are of course other possible grids; two are listed under (48) below. Discussion of the general theory of rhythmic patterns may be found in Lerdahl and Jackendoff (1983) and Hayes and MacEachern (in progress).

4.1 Phonological Coding

We entered the folksong material into a database, translating the musical textsetting from note values into a simple numerical format. The linguistic texts were annotated for stress (using levels from 1 to 4) and for the weight (heavy or light) of stressed syllables. We also coded our data for levels of P-structure, making use of the five-level account of English proposed in Hayes (1989), which is in turn modeled on Nespor and Vogel's (1982) account of Italian P-structure.[6] The five levels of P-structure assumed are as follows:

```
(6)      Utterance
             |
    Intonational Phrase
             |
    Phonological Phrase
             |
        Clitic Group
             |
            Word
```

[5] Although for our purposes a uniform data corpus is more useful, we naturally hope that the methods developed in our project will be useful to those whose interest lies in the study of authentic folk material, as opposed to our own main interest in the theory of metrics.

[6] A couple of details: where the option existed of adjoining a non-branching complement to a head (Hayes 1989, 214-18), this option was taken. Compounds were treated, somewhat arbitrarily, as single Clitic Groups, consisting of two Words.

Because the corpus included very few line-internal Utterance breaks, all Utterance breaks were recorded as Intonational Phrase breaks instead, to simplify computation. All phonological encoding was carried out independently by both authors, whose transcriptions agreed with one another to a fair extent.[7] The final coding used was a compromise, established after examining cases in which the original codings disagreed.

Here is an example of the data for a particular line (Karpeles 1974, #242A):

```
(7)a. Textsetting:

                x               x               x               x
        x       x       x       x       x       x       x       x
        x   x   x   x   x   x   x   x   x   x   x   x   x   x   x   x
        |   |   |       |       |   |   |       |       |       |
     'Twas  in  Ox-    ford    Ci- ty   I      lost    my      way

   b. Weight of stressed syllables:           H      L      H       H
   c. Stress pattern:                   1  1  3  1    4 1 1  3    1  3
                                      'Twas in Oxford City I lost my way.
   d. P-structure:                       |  |    |      |  |  |    |  |
        Words:                           W  W    W      W  W  W    W  W
                                                 |          \/       \/
        Clitic Groups:                           CG         CG       CG
                                                 |            \  /
        Phonological Phrases:                    P              P

        Intonational Phrases:                          I
```

4.2 Native Speaker Judgments

The remaining source of data consisted of the intuitive judgments of ten native speakers of English.[8] These consultants were asked to chant all of the lines of the data corpus in whatever seemed to them to be an appropriate binary rhythm. The textsetting judgments thus rendered were taken down on paper in an ad hoc notation. Consultants who volunteered ternary rhythms (which

---

[7] Differences between transcribers, as a percentage of the total number of values transcribed:

(i)

	Stress (1-4 scale)	P-structure (five degrees of cohesion)
identical:	88.2%	85.0%
differ by 1:	8.8%	13.6%
differ by 2:	2.5%	1.3%
differ by 3:	0.4%	0.1%
differ by 4:	---	0.02%

Large stress differences typically arose from differing judgments of the pragmatic context, which can determine focus stress.

[8] We selected individuals we knew, on the grounds that it was more important to us to have highly motivated consultants who could be asked follow-up questions, than to have a representative sample of English speakers. The consultants were of both sexes, had varying degrees of musical training, and were paid for their participation. We can find no obvious effects of gender or musical education in their textsetting patterns.

are unscannable in the metrical pattern we assume) were gently corrected, by example if necessary, until they produced binary rhythms. Unscannable textsettings, and textsettings in which the consultant had heard the eliciter's example, were excluded from the data set.

To check the reliability of our transcriptions, we transcribed the productions of one of our consultants independently but simultaneously. Our transcriptions agree for 99.6% of the metrical positions, and exhibit complete agreement for all 16 positions in 97.0% of the transcribed lines. This level of agreement suggests that our results are probably not heavily influenced by transcriber error.

Here is how the consultants responded to line (7):

```
(8)     |     | x  |   |    |    |  x |    |    |    | x  |    |    |   | x  |
      x |     | x  |   | x  |    |  x |    | x  |    | x  |    |  x |   | x  |
      x | x   | x  | x | x  | x  |  x |  x | x  | x  | x  | x  |  x | x | x  | x
a. 'Twas| in  | Ox |   |ford|    | Ci-| ty | I  |    |lost|    | my |   |way |
b. 'Twas| in  | Ox |   |ford|    | Ci-|    | ty | I  |lost|    | my |   |way |
c.      |'Twas| in |   | Ox |ford| Ci-| ty |    | I  |lost| my |    |   |way |

    a. = original setting, and consultants AM, CS, CW, DS, EF, JM, KK, and RW
    b. = RH
    c. = JD
```

It will be noted that consultant JD provided a rather unusual textsetting. This was characteristic: in both sessions in which JD participated, he started out with ordinary settings, then, as boredom set in, shifted to what we take to be deliberately unusual ones. JD's behavior is useful in the present context, since we can see (some of) the ways a textsetter attempting to break the rules is able to do so.

The ten consultants together with the original settings yielded an average of 3.41 different settings for each line. On the average, the setting that got the most “votes” received an average of 6.7 votes out of a possible 11.

All of our elicited data were entered, along with the original settings, in a computer data file, and all the generalizations reported below were obtained by searching the file with simple programs.

4.3 Truncated Settings

An important aspect of our data that will be referred to below is a distinction between **full** and **truncated** textsettings. In our typology, we consider only truncations that leave several positions empty at or near the end of the line. This type of truncation has two main forms. One places no syllables between the third and fourth strong positions, so that the syllable in the third position is sung quite long. An example of this truncated textsetting type, which we will call a “Green O” truncation, is given in (9).

```
(9)          x               x               x               x
     x       x       x       x       x       x       x       x
     x   x   x   x   x   x   x   x   x   x   x   x   x   x   x   x
         |   |       |       |       |       |               |
         A- móng    the    leáves   so     gréen,            O!
```

(Sharp 1916, #79)

There is also a more radical type of truncation that leaves unfilled the fourth strong position, so that the line-final syllable is sung very long (shown below), or alternatively, followed by a substantial pause:

```
(10)         x               x               x               x
     x       x       x       x       x       x       x       x
     x   x   x   x   x   x   x   x   x   x   x   x   x   x   x   x
     |       |           |   |       |       |
     I      sówed       the séeds   of      lóve           (Sharp 1916, #33)
```

We will call such settings “three-beat” lines (or just “3”), with a reminder to the reader that the silent fourth beat remains quite detectable in the timing. Any non-truncated line will be called a “four-beat” line (“4”).

The truncations are structurally crucial to folksong, since stanza types are largely defined by the distribution of truncations within them. In general, throughout a song, the verses maintain a consistent pattern of truncation; e.g. 4444, 4343, 43443, G(reen-O)3G3, 4G4G, 444G, and so on. Moreover, unlike for other aspects of textsetting, the choice of truncation type is usually not a matter for variation in performance; rather, the pattern is an inherent part of the song. A general theory of truncation patterns in English folksongs is given in Hayes and MacEachern (in progress).

5 Background III: Metrics

A basic insight found in much of the literature on metrics (beginning with Jespersen 1900) is that the matching of stress contours to meters respects relative, not absolute stress. For example, a stressed syllable will be fairly comfortable filling a weak metrical position provided it is surrounded by syllables with strong stress (cf. the medial syllables of *séas rùn drý́* in (33b) below or *cóme gò báck* in (41)). This idea is expressed in the following rule schema:

(11) Stress Matching

The rises and falls of stress within the line should be matched to the rises and falls of the metrical pattern.

The ill-formed textsetting in (3) sounds bad in large part because its rises and falls of stress have been deliberately placed in consistent disagreement with the metrical pattern. Stress Matching is only a rule schema, not a specific rule. Individual rules under the schema specify *which kinds* of rises and falls (i.e. in which contexts) most strongly require matching to the pattern. We will explore below which kinds of rises and falls are regulated with special strictness in English folksongs. A necessary amplification of the Stress Matching rule schema may be found in the work of Lerdahl and Jackendoff (1983), who note that when a grid position does not correspond to the inception of any auditory event, it counts as having a weaker realization than any grid position that does correspond to the inception of an event. We treat this here as a special provision for "zero", letting the non-inception of a syllable serve in effect as the weakest of all possible stress levels:

(12) <u>Zero Provision</u> (adapted from Lerdahl and Jackendoff 1983, 76; 79)

> When a metrical position is unfilled, or (equivalently) filled by the phonetic continuation of the preceding syllable, treat it as being filled by a stress level that is weaker than that of a stressless overt syllable.

The Zero Provision allows us to explain why (13) is an absurd, almost unperformable textsetting, even though all of its ups and downs of stress are perfectly matched to corresponding ups and downs (boldface) of the metrical pattern:

```
(13)         x               x               x               x
     x       x       x       x       x       x       x       x
     x   x   x   x   x   x   x   x   x   x   x   x   x   x   x   x
         |           |           |   |           |   |           |
      **Lòng        tíme       I've plóughed    the  ó-        cean
```

(construct, after Sharp 1916, #43)

The reason is that all four of the strong positions are occupied by mere syllable-continuations; and are thus grossly mismatched with the weaker positions where syllables are initiated. The Stress Matching requirement, taken in the aggregate, is obeyed rather strictly in the corpus of folksongs we examined. To get a rough idea of this, consider a **stress profile** of the type used in the "Russian" school of metrics (Bailey 1975, 1979; Tarlinskaja 1976, 1987, 1993). The profile below simply averages our conventional stress values (1 weakest, 4 strongest, 0 for null) for all four-beat lines, across the 16 positions of the meter:

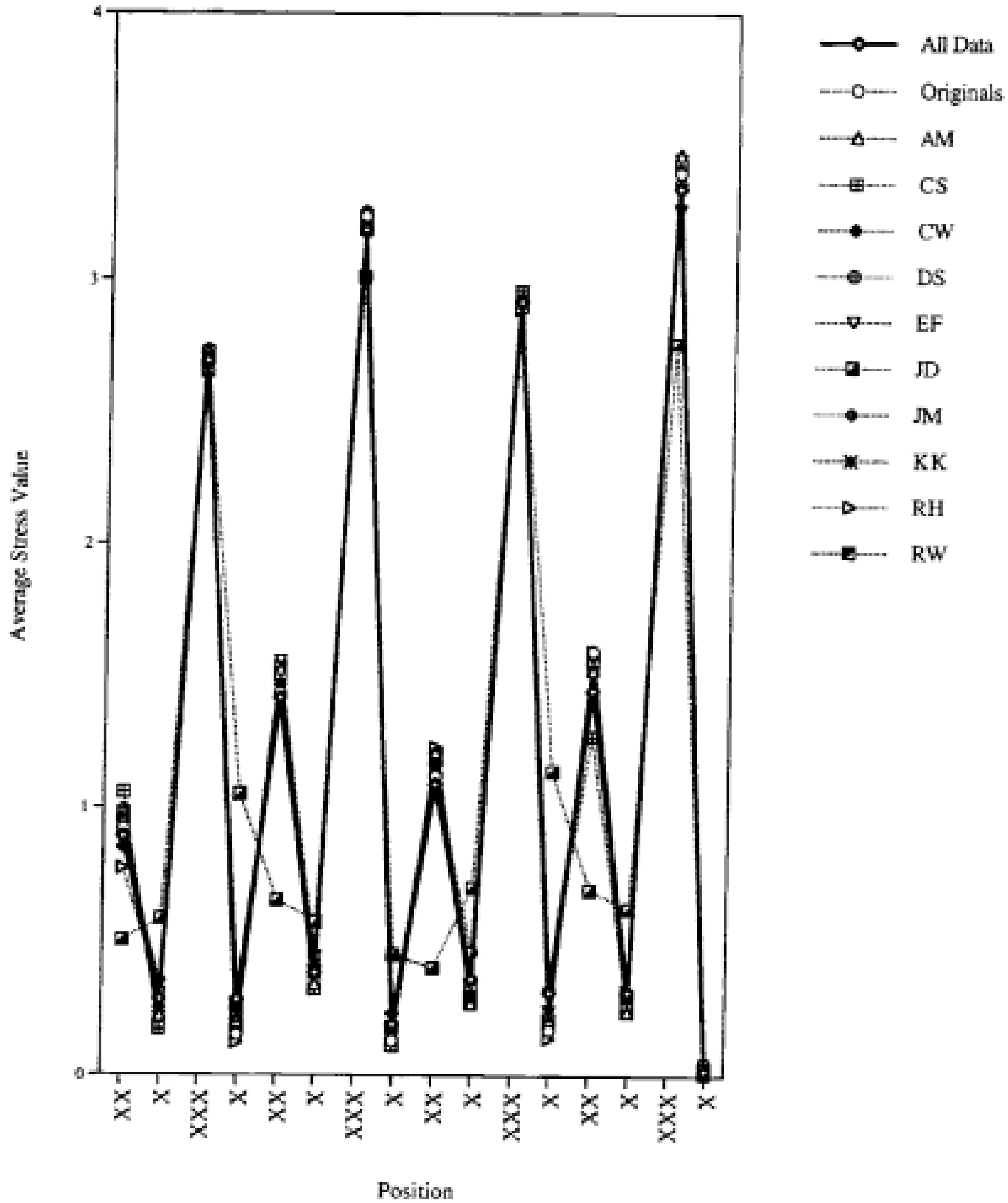


Figure 1. Stress profile of all four-beat lines

It can be seen that (except for consultant JD, on whom more below), stress closely reflects the metrical pattern.

5.1 Metrics and P-Structure I: Matching of Bracketing in the Metrical Pattern

We consider now the influences of P-structure in metrics. Of these, the most obvious concerns the division of verse into lines. In all metrical traditions known to us, there is a requirement that line boundaries generally coincide with high-level breaks in P-structure. The requirement can vary in strictness: for example, the spoken verse of Alexander Pope is characterized by close agreement of line boundaries with prosodic breaks, whereas Milton is famous for his frequent and adventurous disagreements.[9] We judge our folk music corpus to behave more along the lines of Pope than Milton: we seldom transcribed anything less than an Intonational Phrase boundary at line endings, and never anything weaker than a Phonological Phrase boundary. This principle of bracketing agreement can often be taken further: the line has an internal hierarchical structure, and there is a fair degree of agreement of P-structure with this internal hierarchy. For example, Tarlinskaja's (1984) account of juncture in Pope suggests that Pope composed iambic pentameters using the pattern of (14):

```
(14)     x     x     x     x     x
      x  x  x  x  x  x  x  x  x  x
       \/    \/    \/    \/    \/
      Foot  Foot  Foot  Foot  Foot
         \  /        \  |  /
      Hemistich     Hemistich

                Line
```

The tendency for Pope to place large P-structure breaks at line boundaries is duplicated, less strongly, for line-internal constituency. In particular, of the line-internal locations, the one that most attracts P-structure breaks in Pope's verse is the hemistich break, following the fourth position. The hemistich-internal foot boundaries are weakly break-attracting, and the positions that least attract P-structure breaks are the foot-internal ones.

We will later examine what kind of line-internal bracketing would be expected for the grid pattern studied here, and to what extent such bracketing can be verified empirically by its correspondence with P-structure.

5.2 Metrics and P-Structure II: Regulation of Stress Matching

A more subtle case of P-structure effects in metrics concerns the interaction of P-structure with Stress Matching (11). As noted above, this is a rule schema, not a rule. Various locations within linguistic structure differ in the strictness with which stress rises and falls must be matched, and actual rules must govern particular linguistic configurations, specifying the degree of metrical tension (ranging from none through total unmetricality) that a mismatch in that configuration will produce. The actual rules will in fact differ across poets or traditions: to a limited extent, different poets select different locations for strict enforcement of Stress Matching.

An important advance in the metrics of the last 25 years has been an increased understanding of how to characterize these locations. For instance, Magnuson and Ryder (1971) and Kiparsky

[9] Even Milton has his limits, however: see Kiparsky (1975, 606).

(1975) established that rises and falls of stress that occur *within a single word* are often regulated with special strictness. For example, Shakespeare commonly writes lines like (15a), with minor mismatches between stress and meter shown in boldface. But he typically avoids writing quite similar lines in which the mismatch is located within a single word (15b).

```
(15)a.        x        x         x          x       x
         x    x   x    x      x  x    x     x x     x
         |    |   |    |      |  |    |     | |     |            (Shakespeare,
       Plúck the kèen téeth from the fièrce tíger's jáws              Son. 19)

    b.       x        x         x          x     x
         x   x  x     x      x  x  x     x x     x
         |   |  |     |      |  |  |     | |     |
       *Plúck immènse téeth from enràged tígers' jáws       (Kiparsky 1975, 591)
```

Shortly thereafter, Kiparsky (1977) showed how phrasal bracketings as well as word bracketings are crucial to metricality.

Hayes (1989) suggested that these earlier results can be expressed more cogently by making reference to P-structure. His contention was that metrical rules refer to P-structure in just the same ways that rules of phonology do. The relevant typology is the one laid out by Selkirk (1980):

(a) Rules may be bounded by a particular domain; that is, a rule will not apply unless all elements in its structural description occur within the same specified domain.

(b) Rules may make direct reference to edges of a particular rank in their structural descriptions.

(c) Rules may refer (though only rarely) to the juncture of two domains within a larger domain.

The basic hypothesis that metrics works like phonology in this respect allows us to make sense of various metrical rules, as follows. (a) The salience of ups and downs of stress within a single word can be seen as resulting from their occurring within the tightest available phonological domain. (b) Various metrical rules posited by Kiparsky (1977) in terms of the notions of c-command and metrical strength can be shown (Hayes 1983) actually to be linear: they refer to the left and right edges of particular prosodic domains, just as phonological rules do. (c) The fact that no metrical constraints have been found that refer to the juncture of two domains makes sense, given that the phonological analogue of such cases is fairly rare. (d) In those cases in which syntactic structure and prosodic structure are demonstrably distinct, it appears that metrics must refer to prosodic structure, just like phrasal phonology.

The theory of P-structure not only provides a coherent framework for the statement of metrical rules, but also permits the formulation of particular substantive constraints on such rules. Consider a generalization that is frequently observed in English spoken meter: when a stress occurs at the right edge of a domain, then it tends to be matched to the meter (i.e. in strong position) with special strictness. Thus a line like (16), with a mismatched stress at the right edge of a high-ranking domain (plausibly an Intonational Phrase) represents a type found quite rarely in most English poets, and for some, not at all:

```
(16)        x             x           x        x         x      (Shelley, The
        x    x    x        x    x      x   x    x   x     x       Revolt of
        |    |    |        |    |      |   |    |   |     |       Islam 6.19.2)
     [ And they fléd, ]I scáttering.--Ló! with réinless spéed
```

Moreover, when a stress occurs at the left edge of a domain, then it tends to be given special freedom, and frequently occurs mismatched to the meter. This is the well-known phenomenon of "inversion":

```
(17)       x         x               x        x         x       (Shakespeare,
         x  x    x    x            x  x    x   x  x      x          Son. 91)
         |  |    |    |            |  |    |   |  |      |
     [I Rícher than weálth, ] [I próuder than gàrments' cóst ]
```

The account proposed by Hayes (1983, 1989) for these patterns was based on a general principle laid out by Kiparsky (1968):

(18) "Beginnings Free, Endings Strict"

Correspondence of linguistic material to a metrical pattern tends to be free at the beginnings of units, strict at ends.

This can be construed as a strict universal: rules may be stated so as to impose strictness either at right edges or across the board, but not specifically at left edges (and analogously for rules licensing extra freedom).

Previous cases of (18) were based on metrical domains, such as line, hemistich, quatrain, etc. Hayes's suggestion was that the same principle could be made also to hold of linguistic domains, i.e. the domains of P-structure. When we do this, there is an immediate payoff: extensive patterns of metrical complexity follow from the hierarchical nature of phrasing. In particular, a poet will often enforce a right-edge prohibition gradiently; for example, only with moderate rigor at the right edges of Phonological Phrases, more rigorously at the right edges of Intonational Phrases, and extremely rigorously at the right edges of Utterances.

This is entirely expected in a strictly layered hierarchy, since any right edge of an Utterance is necessarily also the right edge of an Intonational Phrase and a Phonological Phrase as well, and similarly with other downward implications. In the same way, the special license to mismatch is awarded most readily to Utterance-initial locations, less often to Intonational Phrase-initial locations, and still less often to Phonological Phrase-initial locations, with the same explanation holding.

Summing up, the theory of accentual-syllabic metrics assumed here takes as its central principle the Stress Matching rule schema of (11). Variation across poets and metrical traditions depends on the particular places in which Stress Matching is enforced by particular rules with greater or lesser strictness. Extra strictness may be imposed in tightly-bounded domains, and at right edges of prosodic or metrical constituents; and extra freedom may be extended at left edges.

5.3 Durational Effects in Textsetting

English is a so-called "stress-timed" language, and its syllables may vary dramatically in their duration. By this we mean phonetic duration, not phonological weight: although it is possible to divide English syllables into the traditional categories of heavy vs. light, this distinction is only one of the factors that determine the phonetic duration of English syllables. Additional factors are discussed below.

Given that sung or chanted lines are rendered in a rough approximation to isochrony, we might expect that preferences would arise in how the natural duration of a syllable is reflected in the textsetting. To see how this might work, consider the following examples. We contrast two syllables, one phonetically quite short and the other quite long, in an analogous metrical position:

```
(19)a. i.           x               x               x               x
            x       x       x       x       x       x       x       x
            x   x   x   x   x   x   x   x   x   x   x   x   x   x   x   x
            |   |   |       |       |   |   |       |       |       |
         'Twas  in  Ox-    ford    Ci- ty   I      lost     my     way

      ii.                           x               x
                                    x       x       x
                                    x   x   x   x   x   x
                                    |       |   |   |
                                  ?Ci-     ty   I  lost

    b. i.           x               x               x               x
            x       x       x       x       x       x       x       x
            x   x   x   x   x   x   x   x   x   x   x   x   x   x   x   x
            |   |   |       |       |       |   |   |       |       |
         'Twas  in  Ox-    ford    town   that  I  lost    my      way

      ii.                           x                x
                                    x        x       x
                                    x    x   x   x   x
                                    |    |   |       |
                                 ??town that I      lost
```

The initial syllable of *city* is phonetically quite short, and seems to be set most comfortably to a single metrical position (19a.i). The syllable *town* is quite long, and is set most comfortably to two positions (19b.i). The distinction perhaps becomes clearer when one attempts to recite (19a.i) to the rhythm of (19b.i) and vice versa, as shown in the examples labeled (ii). The same judgments are given for analogous examples by Oehrle (1989, 109-10).

Such judgments give rise to the following postulated rule schema for textsetting:

(20) Syllable Duration

Reflect the natural phonetic durations of syllables in the number of metrical beats they receive.

The principal factors that determine the "natural phonetic durations" of syllables are as follows: syllable weight, stress level (particularly the stress distinctions that determine vowel reduction), segment identity (lower vowels tend to be longer), and most crucially for our purposes, the location of the syllable within P-structure.

As the work of Ladd and Campbell (1991) and Wightman et al. (1992) has shown, the relationship of P-structure to syllable duration can be characterized rather straightforwardly: the higher-ranking the right edge in P-structure at which a given syllable appears, the more it is lengthened. This holds true in our corpus as well: Bertheau (1994) measured the syllable durations of 100 of our folksong lines, pronounced in prose style by the first author, and found that a fair amount of the variation in syllable durations could be predicted on the basis of our coded stress and P-structure levels; $r^2 = .532$. This value seems reasonably high, since the prediction made no use of segmental information.

The correlation between P-structure and duration gives us a new angle to explore in the effects of P-structure on metrics: a high-level P-structure right edge (such as that of an Intonational Phrase) should induce substantial lengthening on the preceding syllable, and induce a strong preference for that syllable to receive multiple beats. This pattern is demonstrated below.

6 Findings

This section presents the empirical findings of our study, and discusses how they bear on the theories reviewed in the previous section.

6.1 Durational Effects

Section 5.3 laid out a prediction concerning the relation of P-structure to the number of beats assigned to a syllable in textsetting: the higher the right edge in P-structure that the syllable precedes, the more beats it should be assigned, following the Syllable Duration rule schema (20).

To see how we could test this prediction, suppose we are dealing with a stretch of text in which two stressless syllables occur between two stressed syllables. Because of the Stress Matching rule schema (11), the two stressed syllables will usually be placed in strong metrical position (i.e., one of the four strongest positions in the meter). Assuming that this is the case, there are then three logically possible ways to arrange the remaining syllables. Anticipating the result to follow, we have assigned these three textsetting patterns mnemonic names that reflect the duration of the *first* syllable, shown in boldface:

```
(21)a. Long First         b. Short First I          c. Short First II
       x           x         x           x             x           x
       x     x     x         x     x     x             x     x     x
       x  x  x  x  x         x  x  x  x  x             x  x  x  x  x
       |     |  |  |         |  |     |  |             |  |  |     |
      'σ     σ  σ 'σ        'σ  σ     σ 'σ            'σ  σ  σ    'σ
```

Which of the three should be preferred? Short First I is likely to be the least favorite, since it neglects the medium-strength medial position by filling it with a mere syllable-continuation. In terms of our theory, Short First I violates Stress Matching both going in and coming out of this position. Empirically, Short First I is in fact usually avoided; for example, it never occurs in the original versions of our songs; and six of our ten consultants used it only once or not at all.

Long First and Short First II are on a more even footing, since each is guilty only of filling the medial, medium-strength position with a syllable that is merely tied in stress (i.e. stressless) with a neighbor. The distribution of Long First and Short First II is thus likely to be governed by considerations other than stress.

In fact, the controlling factor appears to be the natural linguistic duration of the *first* syllable in the interval: where this syllable is long, Long First is employed, and where this syllable is short, Short First (usually II, occasionally I) is employed. The factors that determine the duration of stressed syllables, as noted above, are primarily weight and location within P- structure.

To see this pattern concretely, the reader may wish to review (19a.i) and (19b.i), which embody Short First II and Long First respectively. In (19b.i), *town* is quite long phonetically, due to its weight and its appearance at a fairly strong phrase ending, probably that of a Phonological Phrase; thus (19b) prefers Long First. In (19a.i), *ci-* is quite short, since it is light and is not at any phrase ending at all; thus (19a) prefers Short First II.

Here are some statistical data backing up our claim. To express the degree of preference for Short First, we will use the percentage in (22):

(22) $$\frac{\text{total Short First cases}}{\text{total Long First cases + total Short First cases}} \times 100$$

Where Long First is preferred, this percentage is close to zero; where Short First is preferred, the value is close to 100. The percentage is plotted in Fig. 2 for various data sets.

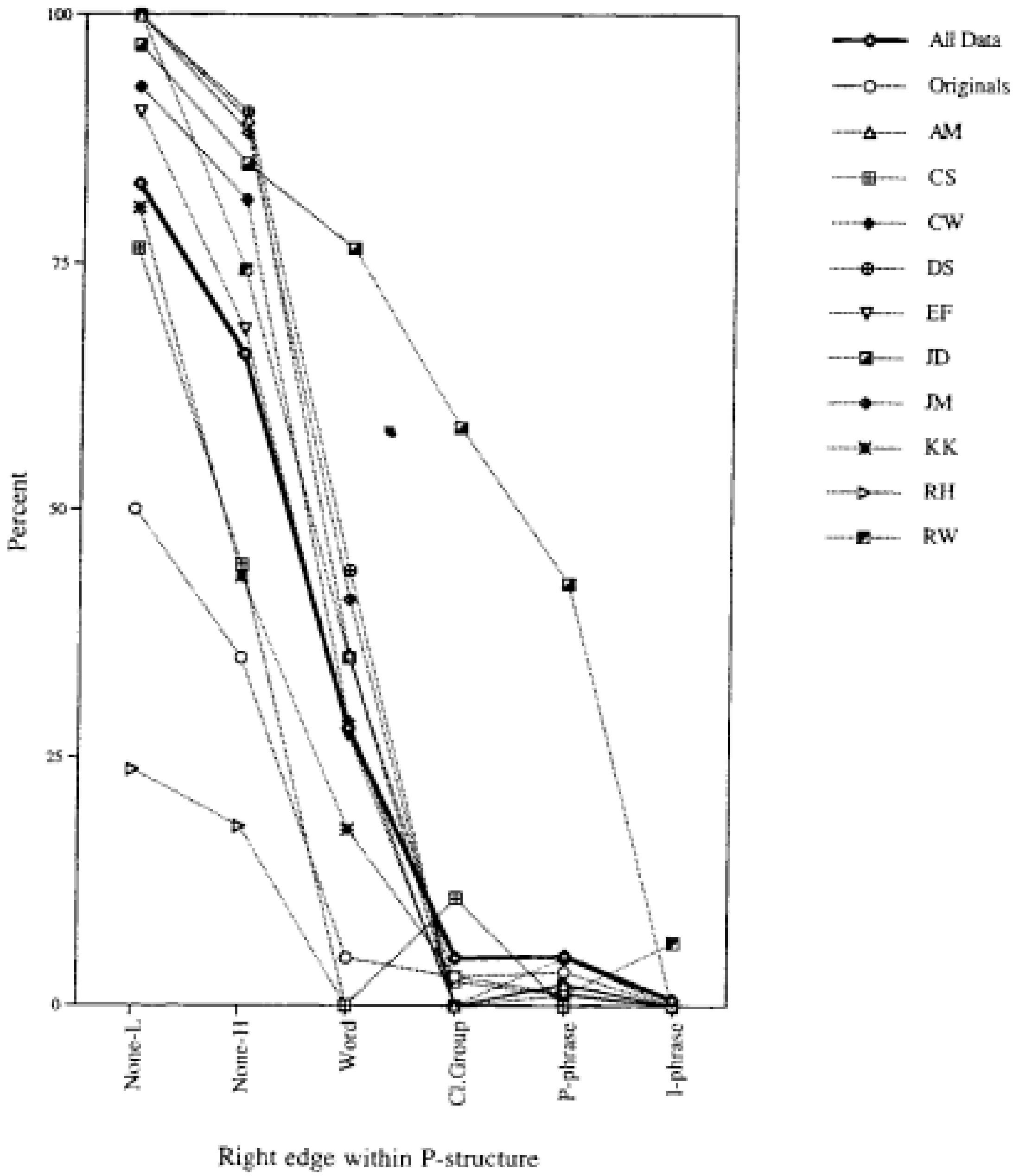


Figure 2. Percentage of Short First textsettings for different syllable durations

The horizontal axis of Fig. 2 is arranged in order of increasing *expected* duration, based on the weight and position within P-structure of the relevant syllables. The syllables that we would expect to be shortest are those that precede no boundary at all in P-structure; i.e. they are word-initial or medial. Moreover, among these, those with light syllables[10] will be shorter. These syllables are

[10] These are, we assume, /C_0ɪ/, /C_0ɛ/, /C_0æ/, /C_0ʌ/, /C_0ʊ/, and in dialects that have it, /C_0ɔ/. The definition presupposes syllabifications like *ci.ty*, *A.dam*; if one wishes to divide such words differently then one would have to

labeled “None-L” or “None-H” on the horizontal axis of the graph, depending on their weight. Beyond these two categories, we consider syllables that are Word-final, Clitic Group-final, Phonological Phrase-final, and Intonational Phrase-final, with each syllable counted as occurring at the highest-ranking available phrase break. Only heavy syllables are relevant here, since all phrase-final syllables are word-final, and in English only heavy syllables may occur stressed in word-final position.

The horizontal axis of Fig. 2, then, represents a continuum of increasing expected duration. By rule (20), Syllable Duration, we anticipate an increasing dispreference for Short First as we move across the graph from left to right. Inspection of the graph shows that this expectation is largely confirmed, both in the aggregate and in the behavior of individual consultants.[11]

Clearly, the consultants exercised a great deal of freedom in displaying an individual bias towards Long First or Short First. The original versions of the songs, as well as the settings of consultant RH, were biased towards Long First, while other consultants tended to favor Short First in varying degrees. But this bias seems always to be expressed in a way that pays attention to the inherent durations of individual syllables.

Much the same conclusion may be drawn if we examine cases in which there is just one stressless syllable between two stressed ones. The three logical possibilities are given labels below (23a,c) taken from traditional musical terminology:

```
(23)a. Scotch Snap       b. Even Rhythm          c. Dotted Rhythm

        x       x           x       x                x       x
        x   x   x           x   x   x                x   x   x
        x x x x x           x x x x x                x x x x x
        | |     |           |   |   |                |     | |
       'σ σ    'σ          'σ   σ  'σ               'σ     σ'σ
```

“Scotch snap” designates rhythms of the form *stressed short + unstressed long*. The term “dotted rhythm” is based on the use of a dot in musical notation to indicate the extra length of the first note. Below are examples of a Scotch snap and a dotted rhythm from the original songs; these examples also include several even rhythms as well:

---

revise the definition of “light syllable” to obtain the same classification. Note that nothing here really depends on assumptions about English syllabification; all we really need is a definition that draws appropriate durational distinctions among syllables.

[11] Two unexpected rises in the graph are based on very small actual numbers, and do not represent statistically significant results:

```
CS, Word (0/17 Short First) vs. Clitic Group (3/28):  p = .162 (chi square)
RW, Phonological Phrase (1/91) vs. Intonational Phrase (2/32):  p = .104
```

The same holds for the unexpected rise for KK in Fig. 4 below (2/41 vs. 3/150; p = .295). Significance figures for all relevant graphs appear in section 8.

```
(24)a.      x           x           x           x
      x     x     x     x     x     x     x     x
      x  x  x  x  x  x  x  x  x  x  x  x  x  x  x  x
      |     |     |     |     |     |  |        |
     Còme  ríse  you    Éng- lish  có- lors,    lòve
```

(Karpeles 1974, #138B)

```
   b.       x           x           x           x
      x     x     x     x     x     x     x     x
      x  x  x  x  x  x  x  x  x  x  x  x  x  x  x  x
      |  |  |     |     |        |  |     |     |
     But the dày  was   wón,    their lìne was  bróke
```

(Sharp 1916, #88)

Naturally, the most common of the three outcomes is even rhythm, since it initiates the syllables in the strongest available metrical positions. Our interest lies in the Scotch snaps and dotted rhythms. If the results obtained above are on the right track, then Scotch snaps should occur most frequently where the initial stressed syllable is short; i.e., non-word-final and light. Further, dotted rhythms should be skewed towards cases in which the first syllable is long; i.e., final within a high-ranking category of P-structure.

The graphs below indicate that this is indeed the case. In the case of Scotch snaps (Fig. 3), we find a general range from prohibition (consultant RH, and nearly so in the original settings) to near-obsession (JD). Across this range, Scotch snaps are favored with the shorter syllables. Dotted rhythms (Fig. 4) show the opposite pattern, favoring longer syllables. Several consultants avoided dotted rhythms almost entirely, while others (JD and the original settings) used them with some frequency. Every consultant who used dotted rhythms with any frequency skewed them towards the longer syllables.

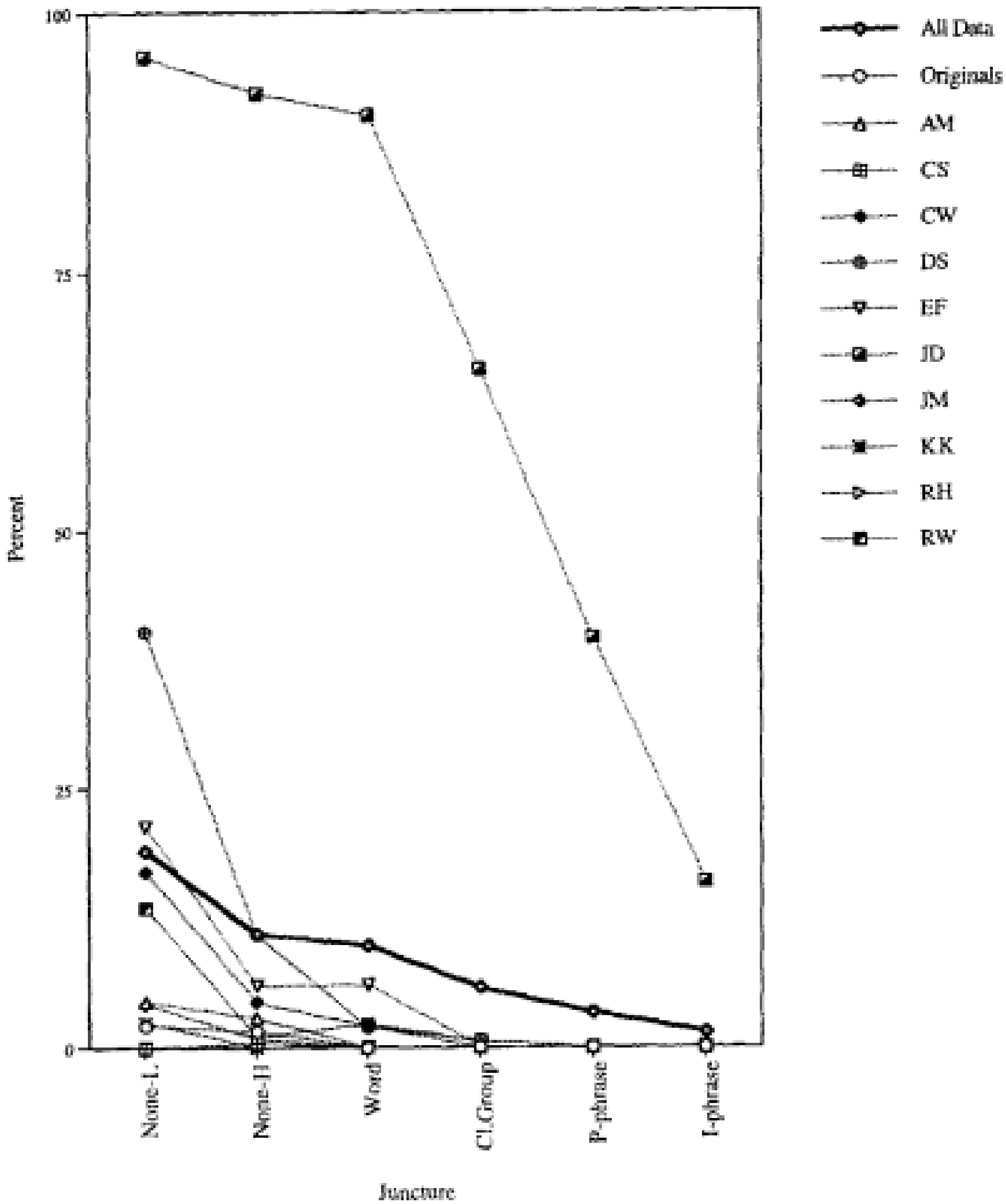


Figure 3. Percentage of Scotch snap textsnappings for different predicted syllable durations

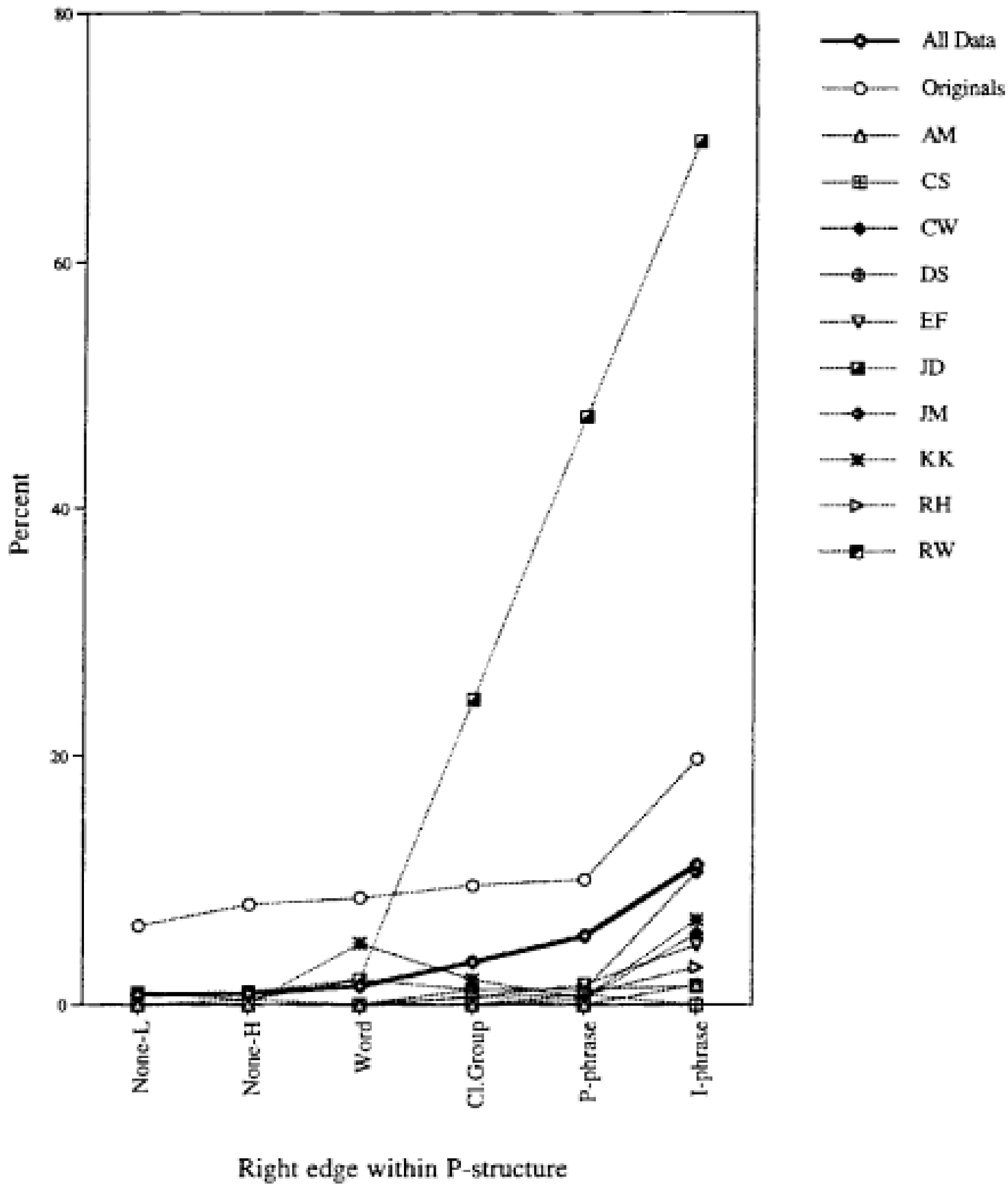


Figure 4. Percentage of dotted rhythm textsnappings for different predicted syllable durations

The core of consultant JD's eccentricity can now be summarized: he largely avoided the normal outcome, namely even rhythm (23b), which is dictated by the Stress Matching rule schema.

Instead, JD used Scotch snaps and dotted rhythms, selecting one or the other largely according to the duration of the relevant stressed syllable.

To summarize this section: P-structure is a principal determinant of phonetic syllable duration. Owing to the Syllable Duration rule (20), syllables in English textsetting are characteristically matched to appropriate numbers of beats. This results in a consistent, gradient effect of P- structure.

6.2 Bracketing Matches

The phenomena discussed in the previous section are only *indirect* effects of P-structure, since they are mediated by the influence of P-structure on syllable duration. But under the theories summarized in section 5, there should be more direct effects as well.

Consider first the most obvious effect, direct correspondence between the hierarchical bracketing of the meter and the bracketing of P-structure (section 5.1). For folk verse, the principles of hierarchical metrical patterns laid out in Piera (1980), Stein and Gil (1980), Lerdahl and Jackendoff (1983), and Hayes (1988) lead us to expect that the full structure of the metrical pattern studied here should be strictly binary branching, as shown below:

```
(25)      x         x         x         x
     x    x    x    x    x    x    x    x
     x  x x  x x  x x  x x  x x  x x  x x  x
      \/   \/   \/   \/   \/   \/   \/   \/
     Foot Foot Foot Foot Foot Foot Foot Foot
        \ /       \ /       \ /       \ /
       Dipod     Dipod     Dipod     Dipod
            \   /               \   /
          Hemistich           Hemistich


                      Line
```

The bracketings posited in (25) can be empirically supported by (among other data) the distribution of Phonological Phrase endings. In the four-beat lines (see section 4.3) of our full data set, these are distributed as shown in Fig. 5.

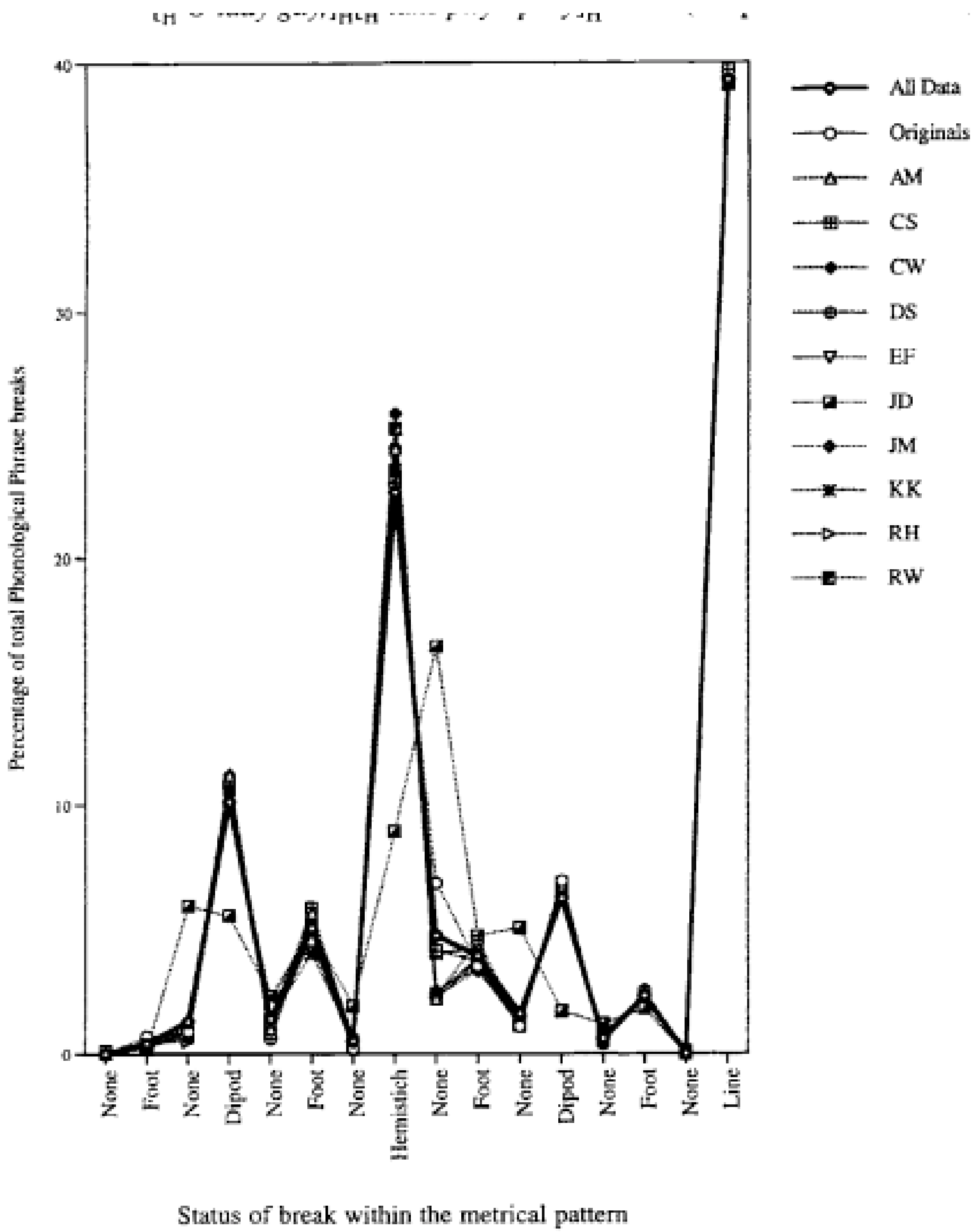


Figure 5. Agreement of Phonological Phrase breaks with breaks in line structure

In this graph, the vertical axis gives the percentage of the total number of Phonological Phrase boundaries that coincide with a particular boundary in the meter. The cases labeled "None" correspond to foot-internal locations in the line, and thus correspond to the edge of no metrical category at all. As expected, the location mostly frequently marked by Phonological Phrase breaks is the line boundary, which results in a high peak at the right edge of the graph. This is followed by the line-medial hemistich boundary, then the hemistich-medial dipod boundaries. Conceivably, this pattern might extend all the way down to the foot level, though here it is hard to tell if we are truly dealing with a junctural effect, as opposed to the syllable duration effects discussed in section 5.3.

At least at the hemistich level, the echoing of metrical bracketing by P- structure emerges rather clearly even at the level of casual inspection. Our corpus abounds in lines such as those of (26) in which the two hemistichs coincide perfectly with Intonational Phrases. The hemistichs are sometimes parallel in structure (26b), or even rhyme with each other (26c).

(26)a. $[_H$ But as she turned out, $]_H[_H$ I'd better been dead $]_H$ (Vaughan Williams and Lloyd 1959, p. 78)
$[_H$ With my hero, $]_H[_H$ Turpin hero $]_H$ (Karpeles 1974, #244B)
$[_H$ O pretty maid, $]_H[_H$ will you wed? $]_H$ (Karpeles 1974, #128B)

b. $[_H$ He's up to the rigs, $]_H[_H$ he's down to the rigs $]_H$ (Karpeles 1974, #242A)
$[_H$ In the month of May, $]_H[_H$ in the month of May $]_H$ (Karpeles 1974, #101A)
$[_H$ But the day was won, $]_H[_H$ their line was broke $]_H$ (Sharp 1916, #88)

c. $[_H$ And the people all in port, $]_H[_H$ they came out to see the sport $]_H$ (Sharp 1951b, p. 24)
$[_H$ He plowed the main, $]_H[_H$ much gold to gain $]_H$ (Vaughan Williams and Lloyd 1959, p. 106)
$[_H$ O lady gay, $]_H[_H$ take pity I pray $]_H$ (Karpeles 1974, #387D)

As noted above, Fig. 5 covers only non-truncated (four-beat) lines. Truncated lines show peaks in their juncture profiles at dipod boundaries, but do not give special prominence to the hemistich break. We conjecture that this is because it is harder to form a complete Phonological Phrase with the limited number of syllables available in the second half of a truncated line.

What emerges from our data is that the binary bracketing hierarchy for folk verse predicted by the theoretical work cited above is confirmed by the “echoing” of the bracketing in the P-structure of the lines that fill the pattern. This is the first, and most obvious, direct effect of P-structure in folksong textsetting.[12]

6.3 P-Structure Bounding Effects

Other effects of P-structure concern the ways in which the rule of Stress Matching (11) is regulated, applying with special strictness in particular locations that P-structure defines. To review: metrical rules are assumed to be subject to bounding domains (typically, words), and may refer to P-structure edges. Following the general metrical principle “Beginnings Free, Endings Strict”, this reference must consist of imparting special metrical freedom at the beginnings of P-structure categories, but special strictness at ends. The higher the level in P-structure, the greater the strictness or freedom should be.

[12] Pilot studies conducted by our undergraduate students suggest this result may generalize quite broadly across languages. In Japanese (Swiger 1994), the effect is very strong indeed; we suspect that non-stress languages like Japanese achieve a perceptible rhythm by means of especially strict matching of P-structure with metrical bracketing.

The strongest bounding effects in spoken verse are those of Words: word- internal sequences are set more strictly than comparable phrasal sequences. One example of this was given in (15); another is (27), which contrasts the treatment of mismatched falling sequences in Shakespeare. Line (27a) is an attested Shakespearian line, whereas (27b) is a constructed line, impossible in Shakespeare's system.

```
(27)a.    x         x         x        x    x
       x   x     x   x      x  x    x    x x  x
       |   |     |   |      |  |    |    | |  |
       Or how  [háps it ]CG I séek not to advánce     (Shakespeare, 1H6 3.1.31)

    b.    x        x        x        x    x
       x   x    x   x     x  x    x    x x  x
       |   |    |   |     |  |    |    | |  |
      *It so [háppens ]W  I séek not to advánce                    (construct)
```

Mismatched falling sequences that follow a major P-structure break have a special status; they are discussed in section 6.6.

Now, in the material examined here, sequences analogous to both (27a) and (27b) are permitted, as shown below (textsettings from the originals):

```
(28)a.       x               x                 x               x
       x     x       x       x       x         x        x      x
       x   x x   x   x   x   x   x   x     x   x   x    x   x  x   x
       |     |   |   |   |   |       |     |   |        |      |
      Have   you a  lìt-tle cráb-   fìsh [ you can     séll    me ]CG ?[13]

                                                          (Sharp 1916, #77)

    b.       x               x               x               x
       x     x       x       x       x       x       x       x
       x   x x   x   x   x   x   x   x   x   x   x   x   x   x   x
       |     |       |       |       |       |       |       |
       As    I      wàlked  oút     one     Mày   [mór-    ning ]W

                                                      (Karpeles 1974, #174)
```

[13] Note that the context of this line is such that *me* cannot be contrastively stressed; any ordinary prose reading would subordinate the stress of *me* to the preceding verb.

But cases like (28a) are considerably more common, relative to their overall textual frequency. The percentage of the time that the two patterns are metrically mismatched is shown in Fig. 6.

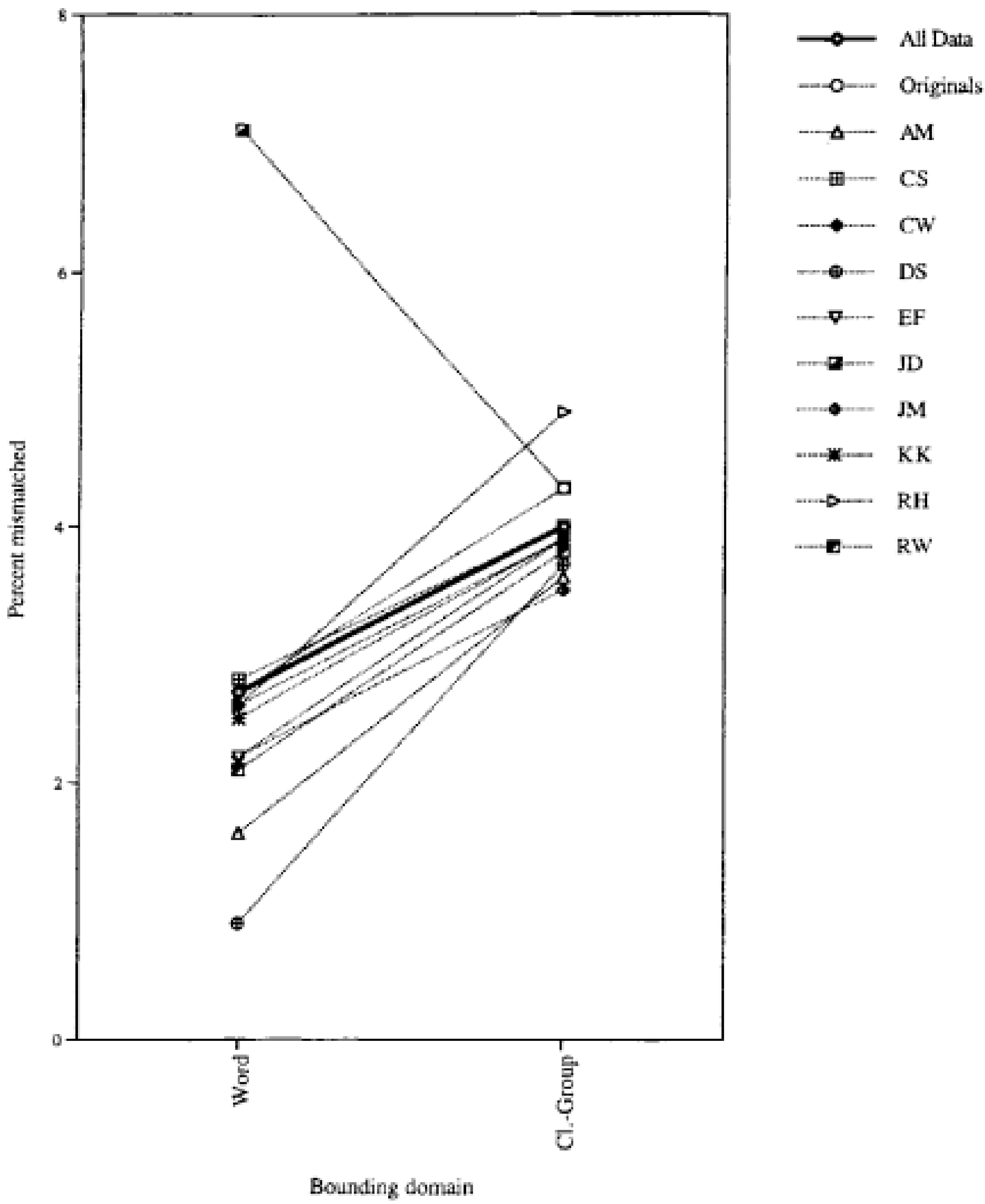


Figure 6. Percentages of sequences mismatched, within Word versus within Clitic Group

As can be seen, the difference between word-internal and cross-word sequences is substantial, and holds good for both the original settings and all but one (predictably, JD) of the ten consultants. For statistical tests, see section 8.

6.4 Strictness at Right Edges

According to the theory assumed here (section 5.2), stresses at the right edges of major domains of P-structure should be text-set with a special strictness corresponding to the rank of the domain. As a test of this hypothesis, we examined two-syllable sequences in which the second syllable has more stress than the first. Such sequences are usually properly matched against the meter, as shown in (29), which includes examples occurring at the right edge of a Clitic Group, a Phonological Phrase, and an Intonational Phrase.

(29)a. Right Edge of Clitic Group

```
                 x               x               x               x
         x       x       x       x       x       x       x       x
         x   x   x   x   x   x   x   x   x   x   x   x   x   x   x   x_
         |       |       |       |       |       |
       That     you      do      lóok ]C so      wán
```

(Sharp 1916, #15)

b. Right Edge of Phonological Phrase

```
                 x               x                x               x
         x       x       x       x        x       x       x       x
         x   x   x   x   x   x   x   x    x   x   x   x   x   x   x   x
             |   |       |       |        |       |       |       |
            And all    Frènch  pów'rs ]P  he      máde     to      sháke
```

(Sharp 1916, #88)

c. Right Edge of Intonational Phrase

```
                 x               x               x               x
         x       x       x       x       x       x       x       x
         x   x   x   x   x   x   x   x   x   x   x   x   x   x   x   x
         |       |       |       |       |       |       |       |
       Stànd   trúe, ]I  my      láds,   like    hèarts   of      oák
```

(Sharp 1916, #88)

Less often, rising sequences at right phrase edges are mismatched against the meter:

(30)a. Right Edge of Clitic Group

```
                 x               x               x               x
         x       x       x       x       x       x       x       x
         x   x   x   x   x   x   x   x   x   x   x   x   x   x   x   x
         |       |           |   |       |       |
        To      méet      with our    pròud ]C fóe        (Karpeles 1974, #287)
```

b. Right Edge of Phonological Phrase

```
             x               x                x               x
     x       x       x       x        x       x       x       x
     x   x   x   x   x   x   x   x    x   x   x   x   x   x   x   x
     |       |       |       |        |       |       |       |
    Her     gòld    ríng ]P off      her     fín-    ger's   gòne
```

(Sharp n.d., p. 10)

c. Right Edge of Intonational Phrase

```
             x                  x               x                 x
     x       x       x          x       x       x         x       x
     x   x   x   x   x      x   x   x   x   x   x    x    x   x   x   x
     |       |       |      |   |           |   |    |    |   |   |
    Còme    ròll    úp, ]I my  láds,       and you shall hàve a príze
```

(Karpeles 1974, #413)

The reader may wish to compare analogous matched and mismatched examples in (29a-c) and (30a-c). Note that the textsetting in (30c), which we would expect to be the worst case, is the original setting; five of the ten consultants found an alternative textsetting placing the syllable *up* in properly matched position.

What should emerge from the data as a whole is that as we examine cases analogous to (29)-(30) with differing P-structure brackets in the designated position, the ratio of mismatched settings like (30) to matched settings like (29) should decrease as we move from lower- to higher- ranking right edges. That this is indeed the case is shown in Fig. 7, which plots as a percentage the fraction

(31) $\dfrac{\text{Mismatched rising sequences}}{\text{All rising sequences}}$

for all such sequences occurring at the right edge of a Clitic Group, of a Phonological Phrase, and of an Intonational Phrase. (Word is not included; such cases are missing almost entirely, due to an independent bounding constraint to be discussed below.) The relevant domain was assumed to constitute a bounding domain; so rising sequences were only counted if both syllables were included in the relevant domain.

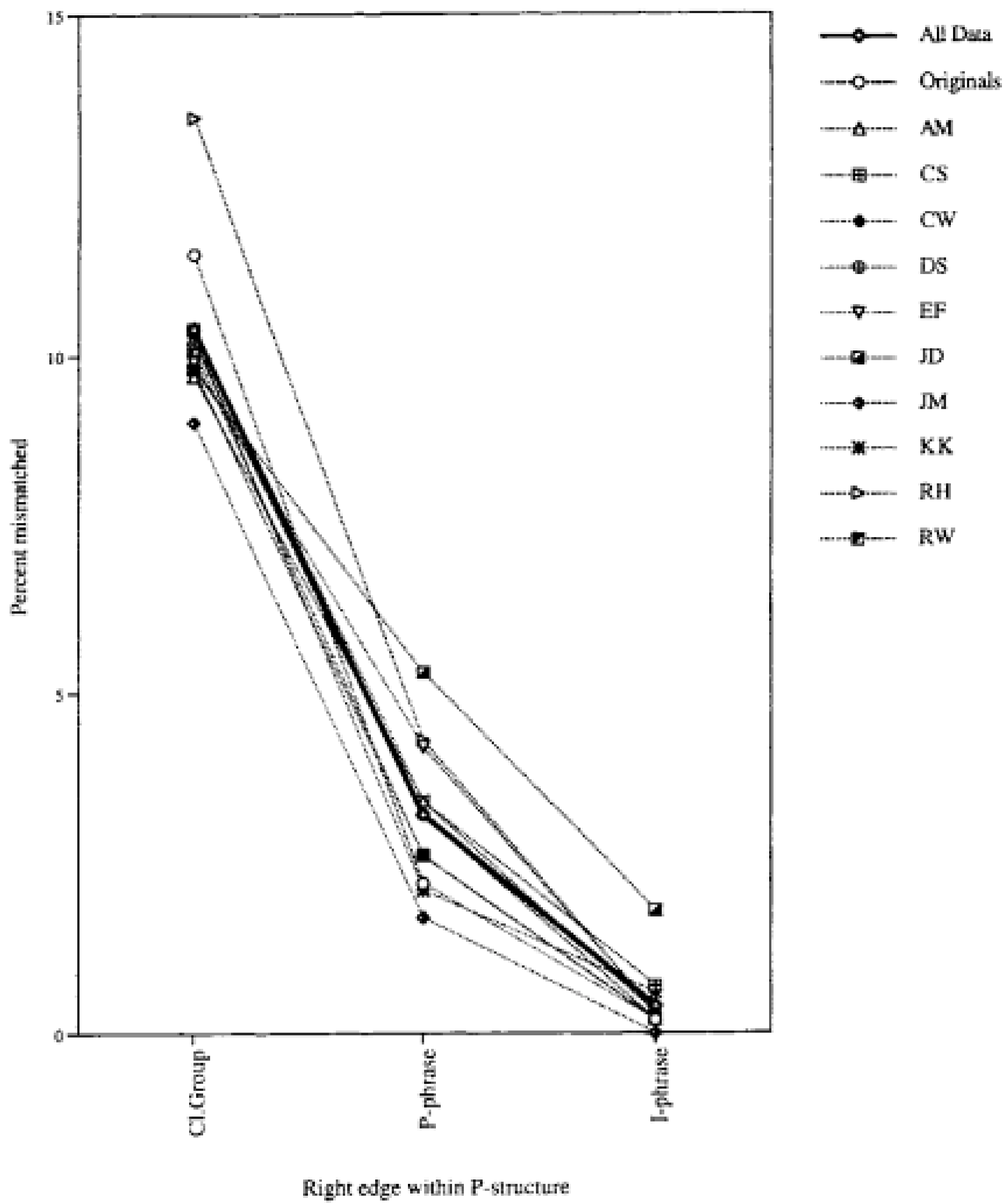


Figure 7. Percentage of rising sequences mismatched, at right edges of high-ranking phrases

As can be seen, there is a sharply falling contour present across the sequence Clitic Group-final, Phonological Phrase-final, Intonational Phrase-final. This falling contour is found both in the original settings and for all ten consultants. The notion of right edges being strict apparently holds quite strongly in sung and chanted verse.[14]

6.5 Bounding and Strictness Together

As Tarlinskaja and Teterina (1974) and Kiparsky (1977) showed, various English poets obey a constraint that simultaneously embodies a right-edge strictness effect as well as a bounding effect. Suppose that a sequence of the form /σ̆ σ́/ is mismatched with the meter, and furthermore suffers from the following double jeopardy: its two syllables are in the same word (the tightest prosodic domain), and moreover the crucial stressed syllable is word-final. Such configurations are rigorously avoided by Milton, Longfellow, and other poets.

They are avoided in sung verse as well. Fig. 8 below compares word-internal sequences of the form / σ̆ σ́ $]_W$ / (i.e., the sequence is both word-bounded and word-final) with analogous sequences that are Clitic Group-bounded and Clitic Group-final.

[14] There two possible confounding factors present. First, it is conceivable that a *skewed distribution* of stress values might produce the observed result. In particular, this would be expected if stronger stresses tend to occur at the ends of larger phrases, and if stronger stresses tend to matched more strictly to the meter. The second possible confound results from the fact that *line-final* syllables are usually strongly stressed, metrically matched, and at the end of a large prosodic domain—the latter because line endings almost always coincide with large P-structure breaks.

It is somewhat more difficult to test our hypothesis while controlling for these factors, since the crucial extreme cases (strongly stressed Clitic Group-final syllables and weakly stressed Intonational Phrase-final syllables) are rare. To get enough data, we did two piecewise comparisons, using non-line-final syllables only:

(i)	Metrically Matched		Mismatched
[3 stress] Clitic Group-final vs.	2639	(cases)	334
[3 stress] Phonological Phrase-final	2437		75
[4 stress] Phonological Phrase-final	1485		60
[4 stress] Intonational Phrase-final	1102		12

Both comparisons came out statistically significant, in the right direction (chi square: $p < .0001$).

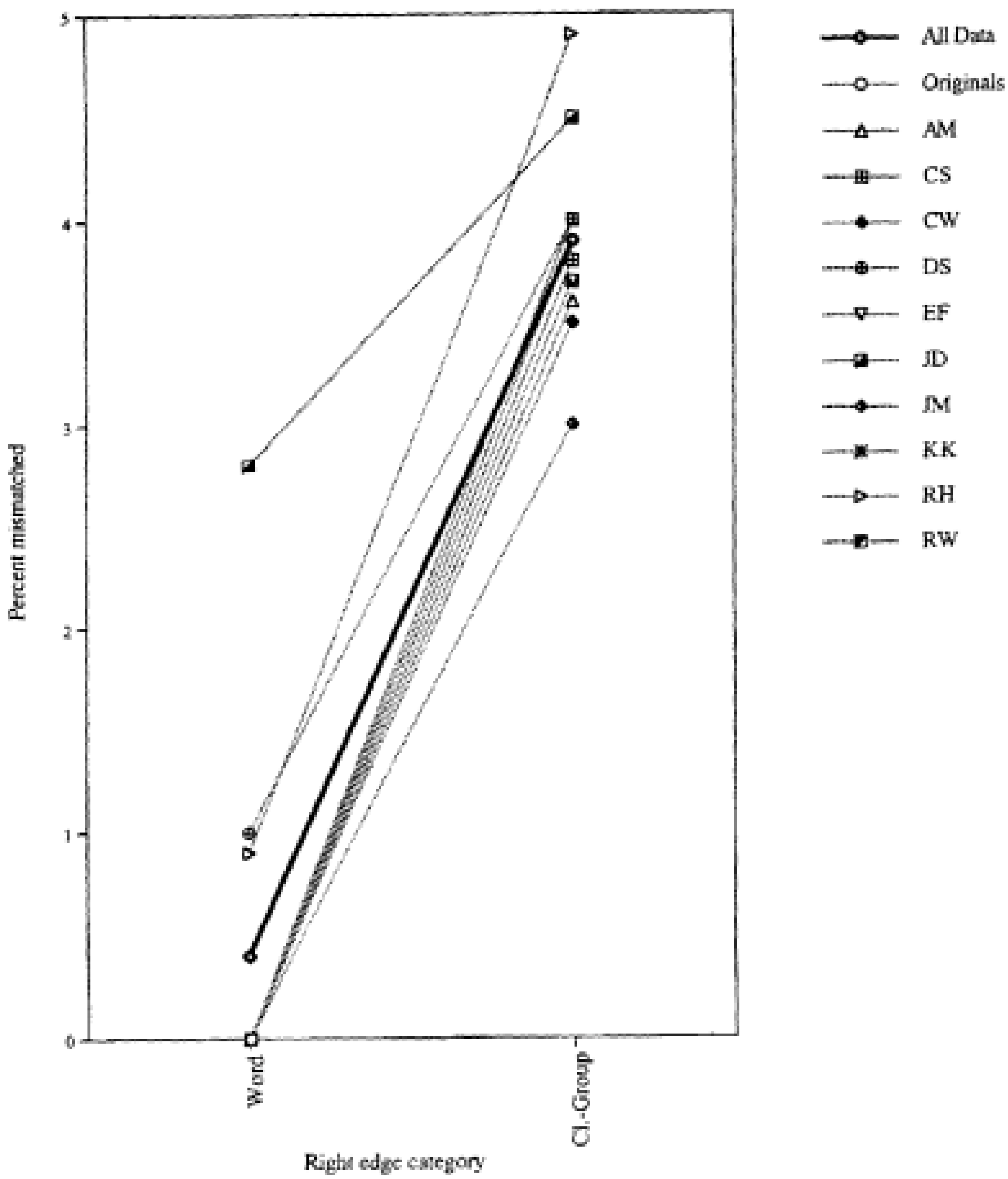


Figure 8. Percentage of stressless + stressed sequences mismatched; right edge of Word versus Clitic Group

Three observations are relevant here: (a) In the original textsettings, as well as for seven of the consultants, the word-final configuration is avoided without exception. (b)úThe effect of word-bounding is evident, in that word- internal sequences are everywhere scanned more strictly than analogous sequences that are only in the same Clitic Group. (c) If one compares Fig. 8 with Fig. 6, the crucial effect of word-final position becomes evident: sequences in final position are everywhere scanned with considerably greater strictness.

Thus far, the general principles assumed here for how metrical rules refer to P-structure have been well supported by the sung verse data: sequences at right edges and word-bounded sequences are indeed set to rhythm with greater strictness. In the following sections, we will examine some rather more problematic data.

6.6 Freedom at Left Edges

Left edges in P-structure have been shown in spoken verse to provide extra metrical freedom, with more freedom obtained at the left edge of higher ranking categories.

In this case, it is striking to see the extent to which the pattern observed in spoken verse largely fails to carry over to sung and chanted verse. Fig. 9 is prepared analogously to Fig. 7; in principle it should show a uniform rise of mismatched cases (i.e. increasing metrical freedom) as one moves from left to right towards higher ranking categories. But this in fact is visible only in the distinction between Phonological Phrase edges and Intonational Phrase edges, and at that only as a very weak tendency, with substantial violations among individual consultants.

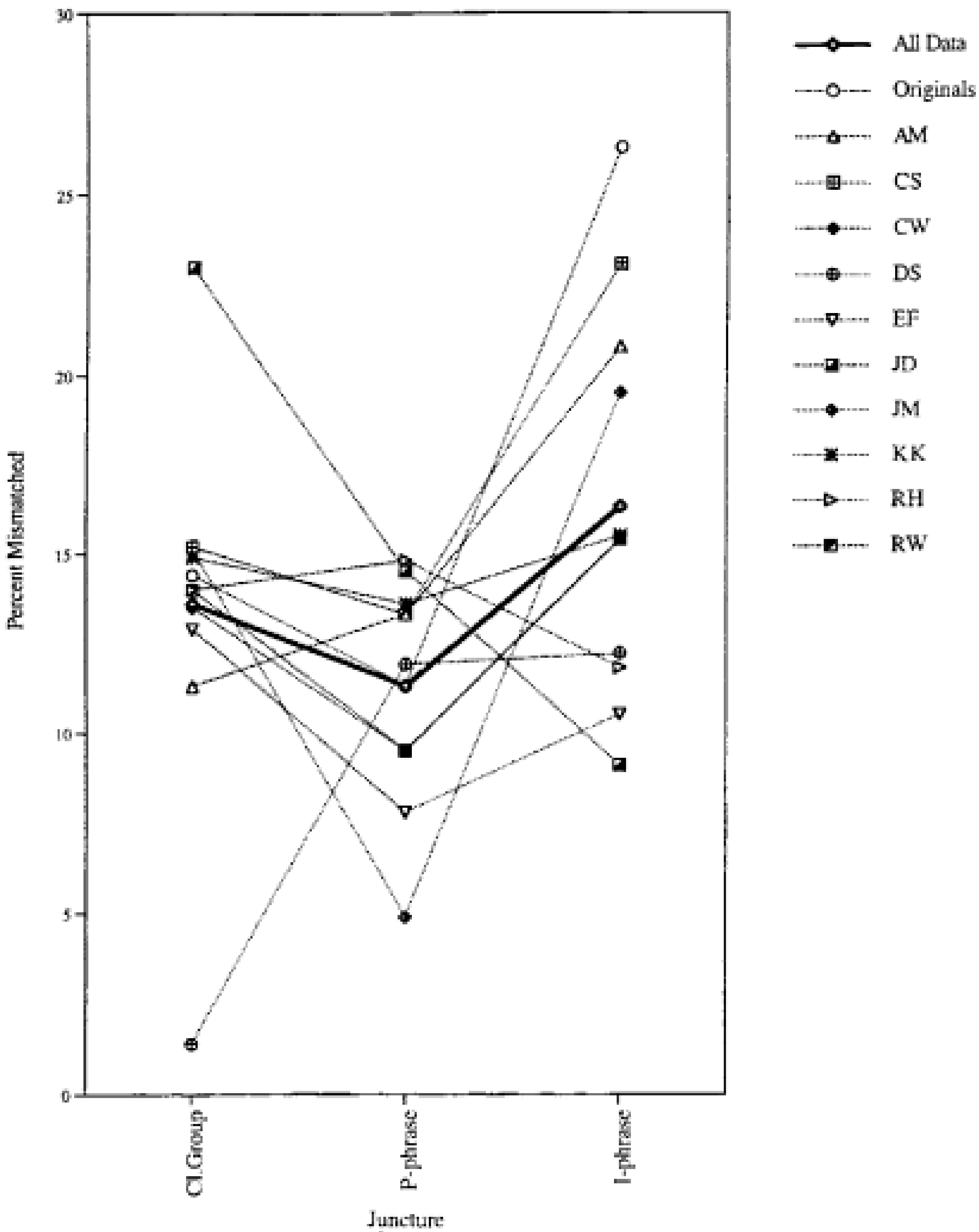


Figure 9. Percentage of stressed + stressless sequences mismatched, at high-ranking left edges of P-structure

Intuition bears out the statistics here. Consider the following constructed pentameters:

```
(32)a.      x       x     x            x          x
        x   x    x  x  x  x        x   x      x   x
        |   |    |  |  |  |        |   |      |   |
      ?So thére he láy a fùll [CG fórtnìght ] of woe

    b.      x       x         x            x       x
        x   x    x  x    x    x       x    x   x   x
        |   |    |  |    |    |       |    |   |   |
      Then lét the séas rùn drý, [I swéetheàrt of míne
```

The mismatched compound *swéetheàrt* in (32b) seems to benefit from its Intonational Phrase-initial position, and seems metrically less awkward than the Clitic Group-initial mismatched compound *fórtnìght* in (32a). Suppose, however, we shift these structures back to the context of the genuine folksong lines from which they were taken:

```
(33)a.          x               x               x               x
        x       x       x       x       x       x       x       x
        x   x   x   x   x   x   x   x   x   x   x   x   x   x   x   x
        |       |       |       |       |   |   |       |       |
       So     thére   he      láy     for  a  fùll    fórt-   níght
```

(Sharp 1916, #84)

```
    b.           x               x               x               x
         x       x       x       x       x       x       x       x
         x   x   x   x   x   x   x   x   x   x   x   x   x   x   x   x
         |       |       |       |       |       |       |       |
       Then     lét     the     séas    rùn     drý,   swéet-  heàrt[15]
```

(Sharp and Vaughan Williams 1951b, p. 73)

The well-formedness difference, if it really exists, is difficult to detect.

In the following sections, we will explore this failure of explanation with further data, and attempt to explain it.

6.7 The Location of Lexical Inversions

A slightly different test of the left-edges-free hypothesis in spoken verse was carried out by Hayes (1989), who examined the frequency of **lexical inversions** at various sites in P-structure. A lexical inversion may be defined as a word-internal /σ̆ σ́/ sequence that is mismatched against the meter. Line (17) above contains two lexical inversions, in the words *rícher* and *próuder*.
In spoken verse, lexical inversions gravitate strongly towards the left edges of major prosodic categories, following the beginnings-free principle. Is this true of sung verse?

Again, the data disappoint this expectation. Fig. 10 shows the strictness with which falling word-internal stress contours are regulated when they occur at the left edges of various prosodic

[15] The sequence *séas rùn drý* forms a dotted rhythm in the original; this is adjusted here to make comparison easier.

categories. Unlike the smoothly rising curves that Hayes (1989) found for pentameter, we get a curious medial peak at the Clitic Group level, with considerable variation across consultants.

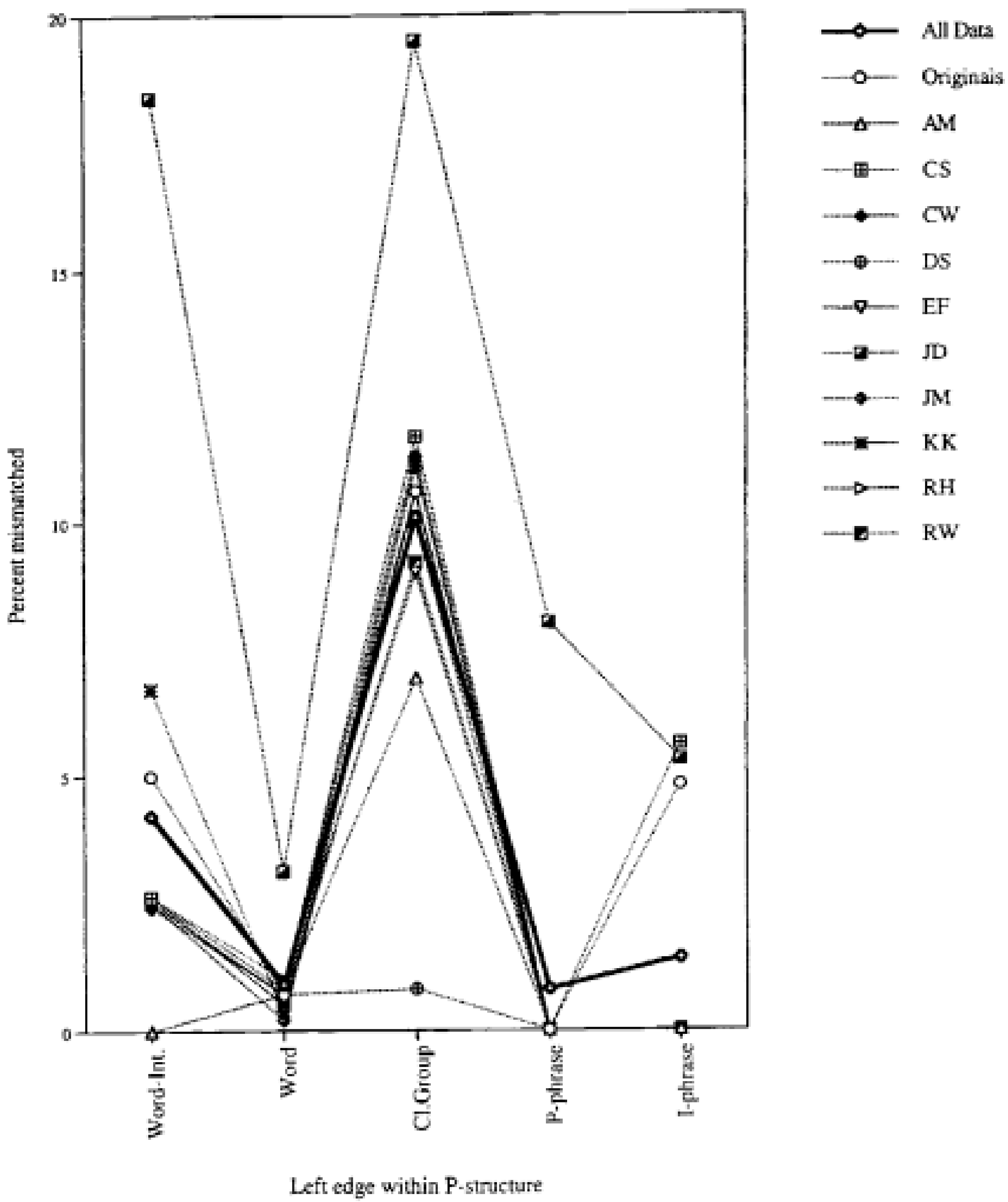


Figure 10. Percentage of falling lexical stress contours mismatched, word-internally and at left edge of all P-structure categories

Before attempting to explain these data, it is worth examining the distribution of lexical inversions from another viewpoint. There is a second way in which lexical inversions in spoken verse are “leftward-biased”: they tend very strongly to occur initially within large-scale *metrical* constituents, of the type discussed in section 5.1 above. We assume, following Kiparsky (1977), Piera (1980), Hayes (1988, 1989), Youmans (1989), and other work, that English iambic pentameters are “quasi-evenly” bracketed: they most often bear the structure of (34a) (see discussion of Pope in 5.1), but in more metrically complex poets pentameters may also have the structure of (34b):

```
(34)a.     x     x     x     x     x     b.    x     x     x     x     x
        x  x  x  x  x  x  x  x  x  x        x  x  x  x  x  x  x  x  x  x
         \/    \/    \/    \/    \/          \/    \/    \/    \/    \/
        Foot  Foot  Foot  Foot  Foot        Foot  Foot  Foot  Foot  Foot
            \/           \ |  /                 \  |  /          \  /
         Hemistich      Hemistich               Hemistich     Hemistich

                  Line                                    Line
```

In other words, the hemistichs of the line must be as equal as they can be, given a five-foot line. The less-marked structure places the longer constituent second, following a general principle of rhythmic preference (Piera 1980, 166-71).

Now, the locations of lexical inversions in iambic pentameter with respect to line structure are well known: the great majority occur line-initially, and of the remainder, most occur in the third or fourth foot. Documentation for this point across the English pentameter tradition may be found in Tarlinskaja (1976, 283). A plausible structural interpretation of this pattern is that lexical inversions prefer to be line-initial, on occasion may be hemistich- initial (under whatever hemistich parsings the poet allows), and only very rarely just foot-initial. As Hayes (1989) notes, this is essentially a Strict Layering effect: the medial hemistich-initial position cannot confer a greater license than line-initial position, since the line-initial position is a hemistich-initial position as well.

If metrical bracketing influences inversion so strongly in spoken verse, clearly we should check out sung verse for the same kind of distribution. We discussed the bracketing structure of folk meter in section 6.2, noting a tendency for bracketing to agree with P-structure at various levels. Given the salience of these levels of metrical bracketing in folk verse, we have every right to think that left edges of metrical units would license lexical inversions.

To test this prediction, consider the four medium-strength locations at which a lexical inversion plausibly would begin (lexical inversions seldom begin in any of the eight fully weak positions):

```
(35) [L                                                                     ]L
     [H                                 ]H[H                                ]H
     [D               ]D[D              ]D[D              ]D[D              ]D
                x                 x                 x                 x
        x       x         x       x         x       x         x       x
        x   x   x   x     x   x   x   x     x   x   x   x     x   x   x   x
        |                 |                 |                 |
        |                 |                 |                Dipod-initial2
        |                 |                Hemistich-initial
        |                Dipod-initial1
       Line-initial
```

If the analogy of pentameter is followed, inversions should occur most often in line-initial position, next most often in hemistich-initial position, and least often in mere dipod-initial position.

As by now we might expect, the analogy with pentameter again fails to go through. In fact, the facts come out almost exactly opposite to the prediction. Fig. 11 gives by percentage the distribution within the line of lexical inversions starting in a medium strength position, for the original songs and for nine of the consultants. (DS is excluded, as he produced no lexical inversions in this context.)

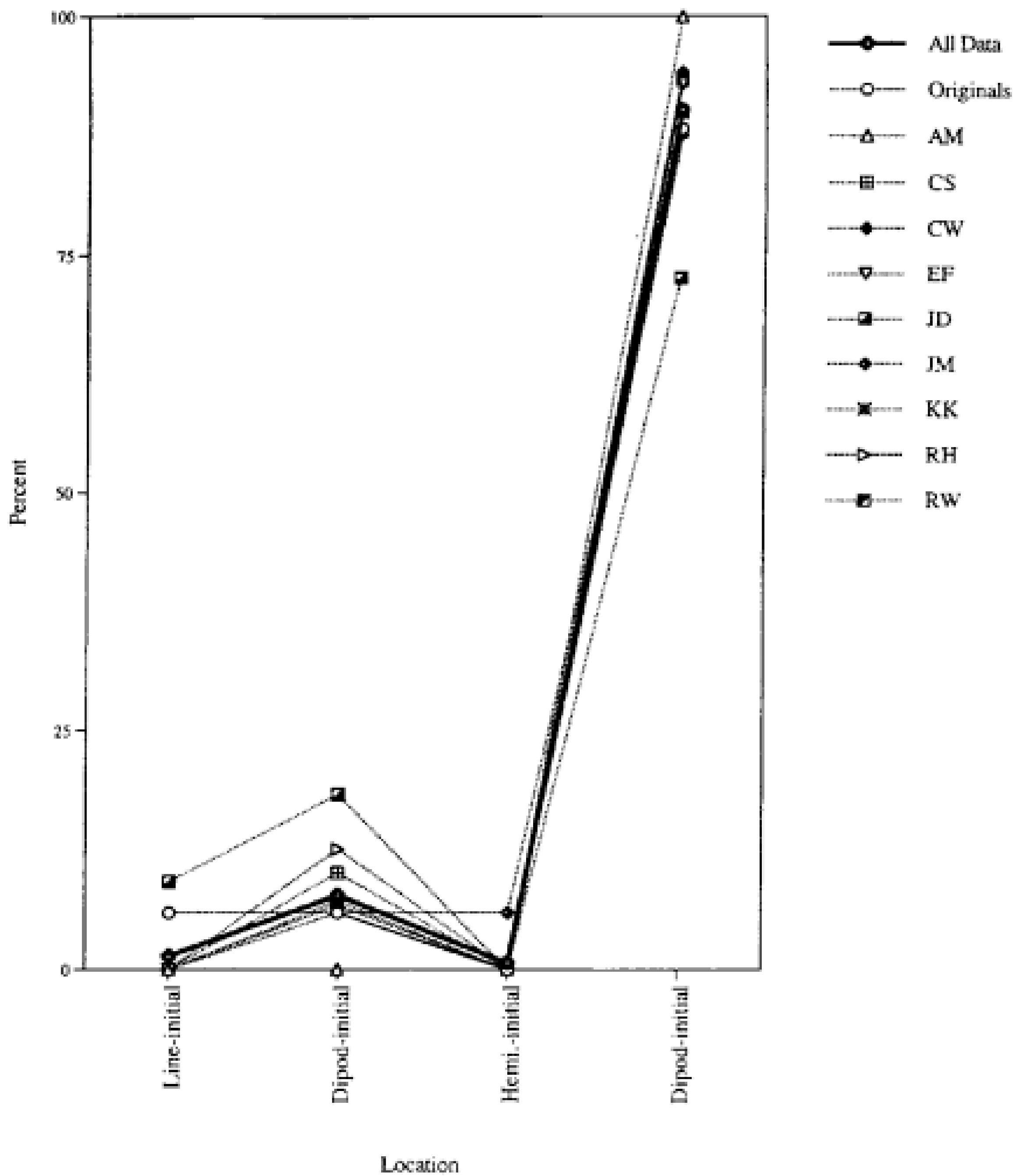


Figure 11. Percentage of total lexical inversions starting in Medium position, by location within the line

As can be seen, there is a large peak in the *final* dipod, corresponding to lines in the corpus like (36).

```
(36)         x                x                x               x
       x     x       x        x       x        x       x       x
       x   x x   x   x   x    x   x   x   x    x   x   x   x   x   x
       |     |       |        |       |        |       |       |
      She   sóon    ràn  through     her      gày    clóth-   ing
```

(Sharp, n.d., p. 10)

The avoidance of line-initial position in Fig. 11 is equally striking: only one line in the whole corpus of original settings involves a "classical" lexical inversion in line-initial position:

```
(37)          x               x                x              x
        x     x       x       x       x        x       x      x
        x   x x   x   x   x   x   x   x   x    x   x   x   x  x   x
        |     |       |       |       |   |    |   |   |   |  |
      Thínk-   ing     to     gét    sómething vé- ry, vè- ry  níce
```

(Sharp and Vaughan Williams 1951a, p. 64)

And this textsetting is unanimously rejected by the group of ten consultants.

These inversion data are potentially deadly to the general approach to metrics taken here, which presupposes as a general principle that beginnings are free and endings strict. Why should lexical inversion in sung verse strongly favor the *end* of the line, counter to spoken verse, and, we claim, a metrical universal?

6.8 An Account of Line-Final Inversion

We can arrive at an answer to this question by considering first some important differences between iambic pentameter and the sung verse we have examined here.

First, in sung verse, metrical positions are plentiful, and unfilled positions are freely tolerated. Indeed, they are ubiquitous: in our data, no line fills more than 14 of the 16 positions, and such cases are rare. The surplus of positions gives the textsetter some freedom not available to the spoken verse poet, a freedom that shows up in the multiple choices made by our consultants. From this perspective, one can ask: is a particular textsetting choice sensible, *given the available options*?

For the most part, it would seem that lexical inversion will almost always compare unfavorably with the alternatives. (The crucial exception is line- final position, on which more below.) Since the use of unfilled metrical positions is rather free, settings like (37) (the sole, highly-exceptional line-initial lexical inversion in the original settings) cannot compete with settings like (38a,b) below, which shows how our consultants treated the word *thinking*:

```
(38)a.        x
        x     x
        x   x x   x   x   x
              |       |   |
           Thínk   -ing to à     (EF, KK, RH, SC)
```

```
b.         x
   x       x       x
   x   x   x   x   x   x
           |   |   |
         Thínking to à        (CW, JM, RW, AM, JD)[16]
```

In these textsettings, the word *thinking* has been "slid over" to the right, vacating the two initial positions and achieving a correct match for the lexical stress contour. Note that vacating the initial two positions is not an especially costly move; fully 10.9% (760/6998) of the settings in our data set are "headless" in this sense.

To put things more generally, in sung verse most candidate lexical inversions will be beaten out by rival textsettings that slide the lexical inversion over into a matching position. Thus lexical inversions usually lose out in the competition for the best textsetting. Iambic pentameter and most other spoken verse, having a completely filled rhythmic template, will not permit the kind of sliding of syllables seen in (38).

To complete the explanation, we need to find the reason why those lexical inversions that do occur in sung verse are concentrated at the end of the line.[17] Again, the basis of our explanation is a fundamental difference between sung and spoken verse, namely that sung verse is performed in isochronous rhythm. Isochrony has the following consequences:

(a) As we have seen, substantial constraints are imposed in sung verse requiring a match-up of natural syllable durations with beat count. Spoken verse, being non-isochronous, permits (as far as we can tell) great flexibility of syllable durations.[18]

(b) The metrical pattern for sung verse is in principle continuous: while we have presented the pattern as a 16-position unit, it is in fact only one part of a larger-scale pattern, in which lines are grouped into couplets, quatrains, and even octets. It is more realistic to conceive the sung meter discussed here as a massive structure containing 64 or 128 positions, not just 16. Crucially, the isochrony of recitation holds for the entire structure, not just for the individual line.

(c) It follows that there is no such thing as an extrametrical syllable (in the sense of Kiparsky 1977, 230-35) in sung verse. Given the continuous pattern and the isochrony of recitation, there is simply no place to put it. In other words, any interpolated syllable that could not be assigned an intrametrical grid interpretation (perhaps by grid expansion; see section 7 below) would interrupt the isochronous recitation, and thus be ill-formed.

---

[16] DS provided a textsetting in ternary rhythm, which likewise placed the syllable *think* in strong position.

[17] One explanation that apparently can be ruled out at the start is that line-final inversions are being used to create rhymes. In fact, the bulk of line-final inversions (12/17, for the original settings) occur in *unrhymed* lines. For further discussion of this point see section 6.9.2.

[18] See however Reichard (1981) for a very interesting attempt to explain aspects of the iambic pentameter on the basis of syllable duration constraints.

Consider now the possibilities for scanning a line that ends in a falling lexical stress contour; this is what traditional scholarship calls a feminine ending. We assume that this is a common enough phrase type in English that there is a fair degree of pressure on the system to be able to accommodate it. We will proceed as we did in (37)-(38), considering what alternatives are available for setting a feminine ending.

In spoken verse, the standard treatment is to make the final syllable extrametrical, as in (39):

```
(39)    x        x     x        x     x
       x x  x    x    x x  x    x    x x
       | |  |    |    | |  |    |    | |
      Tomórrow, and tomórrow, and tomór<row>         (Shakespeare, Mac. 5.5.18)
```

This option is a normal, frequently invoked one throughout the English spoken verse tradition. As we have just seen, this treatment is unavailable for sung verse, which has no place to put extrametricals.

Given this impossibility, we must consider alternatives. One possible location is provided by the "Green O" truncation type, discussed in section 4.3:

```
(40)            x                               x
                x               x               x               x
        x       x       x       x       x       x       x       x
        x   x   x   x   x   x   x   x   x   x   x   x   x   x   x   x
        |       |       |       |       |       |               |
       Lòng    tíme   I've    plóughed the      ó-              cean[19]

                                                                 (Sharp 1916, #43)
```

Another possibility is a three-beat truncated line, which makes use of the medium-strength position between the third and fourth strong beats to accommodate the feminine ending:

```
(41)         x               x               x               x
     x       x       x       x       x       x       x       x
     x   x   x   x   x   x   x   x   x   x   x   x   x   x   x   x
             |       |       |       |       |       |
            Cóme    gò      b ck,   my      hón-     ey
                                                              (Sharp n.d. p. 10)
```

As was noted earlier, the appearance of Green-O and three-beat truncations is stanzaically regulated; i.e. they may only occur in certain positions in certain stanza types. There are many stanza types that exclude both of these truncations, so for many songs the problem of feminine endings is not solved by these options.

Another possibility for a feminine ending is to "steal" a position from a following line:

[19] In (40), some higher-level grid structure is shown. There is good evidence that at least in "Green-O" lines like this one, the third strong beat of the metrical pattern is stronger than the fourth, since in the original songs it is always filled with a syllable bearing stronger stress.

```
(42)           x               x               x               x
       x       x       x       x       x       x       x       x
       x   x   x   x   x   x   x   x   x   x   x   x   x   x   x   x
       |   |   |       |   |   |       |       |       |       |
     Wàke you  úp,   wàke you  úp,    you     sèv-     en    sléep-

                 x               x               x               x
       x         x       x       x       x       x       x       x
       x    x    x   x   x   x   x   x   x   x   x   x   x   x   x   x
       |    |    |       |       |       |   |   |
     -ers / And dó      tàke    wárn-    ing of me
```

(Karpeles 1932, #4E)

This option is seldom taken, since it imposes a gross mismatch of P-structure and line bracketing.

Finally, one might imagine that a feminine ending could be accommodated in the final two positions of the line:

```
(43)          x               x               x               x
      x       x       x       x       x       x       x       x
      x   x   x   x   x   x   x   x   x   x   x   x   x   x   x   x
      |       |       |       |       |       |       |       |   |
    ?Gèorge  Cól-    lins    wàlked  oút     one     Mày     mór-ning
```

(construct, after Vaughan Williams and Lloyd 1959, p. 42)

However, this strategy runs afoul of the rule of Syllable Duration, stated in (20): the phrase-final syllables of morning, which would receive considerable phonetic lengthening in natural speech, are crammed in this setting into the smallest possible space. Indeed, actual cases are quite rare; there are only four in the 670 lines of our corpus in their original settings. They seem to sound best (and occur most often) in lines with many syllables, in which the expected number of beats per syllable is low in any event.

By process of elimination, the only plausible remaining location for a feminine ending is in mismatched position, as in (36), repeated below:

```
(44)          x               x               x               x
      x       x       x       x       x       x       x       x
      x   x   x   x   x   x   x   x   x   x   x   x   x   x   x   x
      |       |       |       |       |       |       |       |
     She     sóon    ràn  through    her     gày    clóth-   ing
```

Despite its lexical stress mismatch, this kind of textsetting frequently wins, because all the rival candidates are either flawed, or limited to certain stanzaic positions.

To summarize our account, we give in (45) another textsetting involving a line-final lexical inversion, along with the rivals it beats out. In context, this line is the third of its stanza, in a song whose stanzaic pattern is 43443; i.e. truncated lines are forbidden in the first, third, and fourth positions.

```
(45)a.          x                x               x               x
        x       x       x        x       x       x       x       x
        x   x   x   x   x   x    x   x   x   x   x   x   x   x   x   x
        |       |       |   |    |       |       |       |       |
       But      I     have been wóoed   by      yòung   Wíl-    liam
```

(Sharp 1916, #15)

```
    b.          x                                 x
                x                x                x               x
        x       x       x        x       x        x       x       x
        x   x   x   x   x   x    x   x   x   x    x   x   x   x   x   x
        |       |       |   |    |       |   |    |               |
       But      I     have been wóoed   by yòung Wíl-            liam
```

(prohibited by stanza structure)

```
    c.          x                x                x               x
        x       x       x        x       x        x       x       x
        x   x   x   x   x   x    x   x   x   x    x   x   x   x   x   x
        |       |       |   |    |       |   |    |       |
       But      I     have been wóoed   by yòung Wíl-    liam
```

(prohibited by stanza structure)

```
    d.          x                x               x               x
        x       x       x        x       x       x       x       x
        x   x   x   x   x   x    x   x   x   x   x   x   x   x   x   x
        |       |       |   |    |       |       |       |       |   |
       But      I     have been cóurt-   ed     by      yòung   Wílliam[20]
```

(phonetically lengthened syllables of William get too few beats)

```
    e.          x                x               x               x
        x       x       x        x       x       x       x       x         x
        x   x   x   x   x   x    x   x   x   x   x   x   x   x   x   x ]L[ x
        |       |       |   |    |       |       |       |       |         |
       But      I     have been cóurt-   ed     by      yòung   Wíl-     liam
```

(line boundary is mismatched)

This interpretation of the line-final bias in lexical inversions is supported by a number of facts.

(a) With a few exceptions noted below, our data largely exclude lexical inversions in “three-beat” lines, as in (46):

```
(46)           x               x               x               x
       x       x       x       x       x       x       x       x
       x   x   x   x   x   x   x   x   x   x   x   x   x   x   x   x
       |       |       |   |   |       |       |
     *My      náme     it  is yòung   Jóhn-   son.
```

[20] We have substituted *courted* for *wooed* to avoid a long rhythmic lapse within the line, which would act as a independent source of ill-formedness.

The reason is that better textsettings are available that use the following beats to accommodate the feminine ending:

```
(47)a.          x               x               x               x
        x       x       x       x       x       x       x       x
        x   x   x   x   x   x   x   x   x   x   x   x   x   x   x   x
        |       |       |       |       |       |       |
       My      náme     it      is     yòung   Jóhn-   son

    b.          x                               x
                x               x               x               x
        x       x       x       x       x       x       x       x
        x   x   x   x   x   x   x   x   x   x   x   x   x   x   x   x
        |       |       |       |       |       |               |
       My      náme     it      is     yòung   Jóhn-           son
```

(= Sharp 1916, #43)

Analogous settings cannot be produced in a four-beat line without exceeding the line boundary.

(b) By the same token, any grid pattern (e.g. (48a,b)) that freely allows feminine endings without overcompression of the final syllables should essentially avoid lexical inversions, since the boldface locations provide a much better location for a feminine ending. Inspection of a number of songs with these grid patterns suggests this is true, though we have not conducted a rigorous search.

```
(48)a. x           x           x           x
       x     x     x     x     x     x     x     x
       x  x  x  x  x  x  x  x  x  x  x  x  x  x  x  x

    b.       x                 x                 x                 x
       x     x     x     x     x     x     x     x     x     x     x     x
       x  x  x  x  x  x  x  x  x  x  x  x  x  x  x  x  x  x  x  x  x  x  x  x
```

(c) One of our consultants, DS, apparently placed very low value on the principle of Syllable Duration, and frequently set phrase-final syllables in highly compressed positions, much as in (45d). As a result, he was able to avoid line-final lexical inversions entirely. In DS, we see evidence that it is precisely the ban on overcompression of syllables near the end of the line that (for other textsetters) forces the use of final lexical inversions.

(d) Finally, Karpeles (1973, p. 25) observes that feminine endings are few in English folksongs. If feminine endings give rise to a set of possibilities that are all undesirable in some respect, plausibly this would give rise to a reluctance to use them in general.

To sum up, the goal of this section has been to explain the line-final bias for lexical inversions in sung verse. We have suggested that initial inversions are not especially favored, as they are in spoken verse, since there is almost always a better alternative available that “slides” the lexical stress into strong position. The inversions that do occur are biased toward final position because of a conspiracy of factors that block the sliding effect line-finally.

Under this view, the metrics of sung and chanted verse takes a fairly ordinary place in the typology of metrical systems. It follows the principle of extra strictness at major prosodic endings, is extra strict within tightly bounded prosodic domains, and fails to show more than the mildest "beginnings free" effects owing to a rather specific combination of circumstances.

6.9 Further Details of Inversion

We have not yet explained why lexical inversions should occasionally be found in positions other than in the final section of a four-beat line. Most of the exceptions fall into three categories: ersatz stressings, mismatches forced by other metrical constraints, and parsing errors.

6.9.1 Ersatz Stressings

The majority of the exceptional cases we have found can be explained on the basis of the following observation. A particular word sometimes appears first in a song inverted "appropriately," in line-final position. Then, later, it appears again, not line-finally, but with the very same textsetting as before:

```
(49)a. Line 1

         x               x               x               x
 x       x       x       x       x       x       x       x
 x   x   x   x   x   x   x   x   x   x   x   x   x   x   x   x
 |       |       |       |       |       |       |       |
 As      I      wàlked  oút     one     Mày     mór-    ning

    b. Line 2

         x               x               x               x
 x       x       x       x       x       x       x       x
 x   x   x   x   x   x   x   x   x   x   x   x   x   x   x   x
 |       |       |       |       |       |               |
One     Mày     mór-    ning    so      eár-            ly

                                        (Karpeles 1974, #174)
```

Here is a list of similar cases; the syllable placed in strong position is shown in boldface:

(50) a. (May) mor**ning**, Karpeles (1974, #94A), Sharp (1916, #61)
b. <u>(stout) mil**lard**</u>, Sharp (1916, #80)
c. <u>(Henry) Mar**tin**</u>, Sharp (1916, #1)
d. <u>(merry) Scot**land**</u>, Sharp (1916, #1)
e. <u>(May) Col**vin**</u>, Ritchie (1965, p. 8)
f. <u>Mus**grave**</u>, Ritchie (1965, p. 36)
e. <u>Wil**liam**</u>, Sharp (1916, #3)

It happens that two such configurations (i.e. (49b) and (50b)) were found in our data corpus, with the second "imitative" lexical inversion starting in the fifth grid position. These settings were adopted fairly generally by our consultants; hence the minor secondary peak shown in the graph of Fig. 11.

The phenomenon of "imitative" inversion suggests that in sung verse, inversion creates a novel stressing for the affected word; e.g. morníng. This "ersatz stressing" forms, as it were, a new lexical representation, which is then available to be used in any part of a subsequent line. We assume that ersatz stressings arise in sung verse only, because unlike in normal spoken verse recitation, sung verse actually places the rhythmic cues for stress on the "wrong" syllable.

The cost of an ersatz stressing apparently resides entirely, or almost so, in its first use. After that, the ersatz stressing is freely available.

That this account is correct is suggested by cases of imitative inversion in which the ordinarily stressed syllable falls in the weakest possible grid position, forming what would ordinarily count as a wild mismatch of stress. For instance, consultant JM did this with line (49b), setting it as follows:

```
(51)           x               x               x               x
       x       x       x       x       x       x       x       x
       x   x   x   x   x   x   x   x   x   x   x   x   x   x   x   x
       |       |           |   |       |       |       |
      One     Mày         mór-ning    so      eár-     ly
```

The mismatch in (51), although it initially appears outrageous, is merely an ordinary dotted rhythm, not a mismatch at all, if we assume that JD has the ersatz stressing *Mày morníng* in mind, taken from the preceding line. Consultant EF treated the phrase "*stòut millárd*" similarly.

A similar, initially very puzzling mismatch from the original settings likewise finds an explanation in these terms:

```
(52)         x               x               x               x
     x       x       x       x       x       x       x       x
     x   x   x   x   x   x   x   x   x   x   x   x   x   x   x   x
     |       |       |       |       |       |       |   |   |
     O      húsh,    O      húsh,   my     prèt-    ty  Pól-ly

                                                   (Karpeles 1974, #2J)
```

Cecil Sharp collected many versions of "The Outlandish Knight", the folksong from which this line is taken. In four of the six for which the "Polly" line settings are printed in the sources, *(pretty) Polly* appears as an ordinary line-final lexical inversion. It is plausible to suppose that the ersatz stressing *Pollý*, which arose in the other versions, was present in the mind of the singer who devised the textsetting in (52), much as JM did for (51).

6.9.2 Lexical Inversions Forced by Other Constraints

Once we have factored out the ersatz stressings, relatively few lexical inversions remain as counterexamples to the analysis of section 6.8. Some are explainable on the grounds that they make it possible to avoid otherwise-undesirable configurations.

For example, lexical inversions at the end of three-beat lines can occasionally be attributed to constraints on stanza structure. A commonplace stanzaic truncation pattern forbids the placement of any syllables following the third strong position in even-numbered lines of the stanza. When such a constraint is in effect, the third beat of a three-beat line becomes entirely analogous in its possibilities to the fourth beat of a four-beat line, so by the same reasoning deployed in section 6.8, we should get line-final inversions. There is in fact one such line in our corpus, which occurs as the fourth line of its stanza:

```
(53)           x               x               x               x
      x        x       x       x       x       x       x       x
      x   x    x   x   x   x   x   x   x   x   x   x   x   x   x   x
      |   |    |       |   |   |       |   |   |
     And your pérch    of the bèst     í- vo- ry
```

(Karpeles 1974, #2J)

Most of our consultants also inverted *ivory*. We conjecture that (on this occasion) they were able to recognize the stanza type of (53) from the context, or perhaps they were just seeking a rhyme for the word me in the second line.

We are puzzled that there were not more cases of this type, though we have found four others (Hendren (1936, 140); Ritchie (1965, 22); Sharp (1916, #3); Karpeles (1932, #29A)) outside our data corpus. Perhaps they are disfavored because most three-beat lines rhyme, and lexical inversions tend in general to avoid rhyming position.

There are a handful of other cases in our data in which the need to avoid violating other constraints induces idiosyncratic inversion. For example, two consultants inverted the very short three-beat line *To plèad for my Geórdie* (Karpeles 1974, #36A) so as to avoid the extreme lengthening of a line-medial syllable that is otherwise necessary. Line (37) above, with initial inversion, plausibly reflects the need to provide a syllable on which two pitches of the tune may be sung.

6.9.3 Temporary Parser Failure

The remaining explanation we have for lexical inversions outside the fourth beat concerns five aberrant settings produced by consultant CS. These settings were lexical inversions at the end of a three-beat line, but not forced (as in the previous section) by rhyme or by stanza type. Here is one example:

```
(54)           x               x               x               x
      x        x       x       x       x       x       x       x
      x   x    x   x   x   x   x   x   x   x   x   x   x   x   x   x
      |        |       |   |   |       |       |
     My       náme     it  is yòung   Jóhn-    son
```

(line text from Sharp 1916, #43)

CS produced these settings for lines occurring quite early in the script we had prepared for our consultants. Specifically, they occurred in the first few lines of the first two songs that had feminine endings. After producing a few aberrant settings like (54) in each song, CS then settled down to a consistent use of ordinary, metrically matched settings like (55) for the rest of each song (and for the great majority of the rest of the script).

```
(55)          x               x               x             x
      x       x       x       x       x       x       x     x
      x   x   x   x   x   x   x   x   x   x   x   x   x   x   x   x
      |       |       |       |       |       |       |
     To     coúrt   yoùng   gírls    and    hánd-   some     (Sharp 1916, #43)
```

A year later, we presented CS with the lines she had earlier scanned as in (54), this time giving her a choice of settings. CS consistently preferred settings like (55).

It is tempting to view CS's behavior in terms of a grammar, which evaluates the relative merits of textsettings; and a parser, which attempts to locate the textsetting that is most preferred by the grammar.[21] CS's internalized grammar for chanted verse unfailingly selects settings like (55) over settings like (54). But her parser, as it happened, was briefly unable to locate the structure (55) under time constraints. Within a short period of time, CS found this structure, and then preferred it consistently. Moreover, when a year later she was given the structure as an explicit option, she preferred it from the beginning.

Once all of the above cases are taken into account, the number of lexical inversions that go against our theory is very small (about 4-8, depending on further assumptions); and given the size of our data corpus we feel reasonably confident that the factual basis of our theory of lexical inversions is correct. In general, the kind of explanation employed here—that lexical inversions occur in the absence of a palatable alternative—seems quite successful in characterizing their distribution.

7 The Origins of Resolution

Sung and chanted verse in a sense have epistemological priority over spoken verse, since they form part of the childhood experience of every poet. (There is little doubt that verse of the type analyzed here is at least as old as the iambic pentameter tradition; Gerould 1932, Ch. 8.) We suspect that childhood experience forms an important base for the poet creating a spoken-verse metrical idiolect. Below, we present a possible case, namely metrical Resolution, in which a particular practice of spoken verse poets might be seen to have a sung-verse analogue on which it is based.

Recall the phenomena of Short First and Scotch snap textsettings, discussed in section 6.1. In these cases, syllables that are phonetically short are assigned textsettings that give them just one metrical beat. Such cases are reminiscent of the spoken-verse phenomenon of metrical Resolution, which is discussed by Young (1923), Kiparsky (1977, 1989), Prince (1989), and Hanson (1991, 1993). The rule of Resolution permits two overt syllables to occupy just one metrical position, as in the following iambic pentameter example:

---

[21] This is in fact the approach taken in Hayes and Kaun (in progress), in which textsettings are generated by the combined action of an Optimality-theoretic grammar (Prince and Smolensky 1993) and a very primitive (i.e., exhaustive-search) parser.

```
(56)       x         x       x       x         x
      x    x    x    x   x   x   x   x   x     x
      |    |    |    |\  |   |   |   |   |     |              (Shakespeare,
     And spénd his pródigal wíts in bóotless rhýmes           L.L.L. 5.2.64)
```

According to Kiparsky, the Resolution rule requires that the first of the two resolved syllables be light. We will also adopt the suggestions of Prince (1989, 53) for further restricting the rule: the first syllable must be stressed, the second syllable must be stressless, and (at least in iambic pentameter) the resolution must occur in strong metrical position.

Actually, these two proposals appear initially to be incompatible. A fair number of lines in Shakespeare appear to require either that we resolve a heavy syllable (as in (57a)), or that we resolve in weak position (57b):

```
(57)a.      x         x          x     x         x
        x   x     x   x      x   x  x  x    x    x
        |   |     |   |      |   |  |  |\   |    |
       His wife, his babes, and all unfortunate souls

    b.      x         x          x     x        x
        x   x     x   x      x   x  x  x  x     x
        |   |     |   |      |   |  |  |  |\    |
       His wife, his babes, and all unfortunate souls           (Mac. 4.1.168)
```

Similarly:

(58)	The doors are open, and the <u>surfeit</u>ed grooms	(Mac. 2.2.5)
	Excite the <u>morti</u>fied man. / Near Birnam Wood	(Mac. 5.2.5)
	You are welcome, <u>gentle</u>men. Come, musicians, play	(R.J. 1.5.25)
	The <u>fugi</u>tive Parthians follow. Spur through Media	(A.C. 3.1.7)
	For an abuser of the world, a <u>practi</u>ser	(O. 1.2.79)
	With any strong or vehement impor<u>tuni</u>ty	(O. 3.3.256)
	Un<u>merci</u>ful lady as you are, I'm none	(Lear 3.7.31)
	Let me be counted <u>service</u>able. How look I	(Cym. 3.2.15)
	So meet for this great errand. Please your <u>lady</u>ship	(W.T. 2.2.49)
	To make a shambles of the <u>Parlia</u>ment House	(3H6 1.1.71)

Similar lines are given by Kiparsky (1977, exs. (18a); (137)). In all of these examples, the stressed syllable of the problematic sequence is antepenultimate or earlier in the word. The relevance of this will be become clear shortly.

At this point we will attempt to establish a connection between spoken-verse Resolution and sung-verse practice. This rests on the fact that for many songs that use a rich grid (e.g. the one described here, with 16 terminal positions), the weakest positions are used quite sparingly—see Fig. 1, as well as Tarlinskaja (1993, 38). To give a concrete example, in the following line, only one of the eight weakest positions is filled by a syllable (-dle).

```
(59)          x               x               x               x
      x       x       x       x       x       x       x       x
      x   x   x   x   x   x   x   x   x   x   x   x   x   x   x   x
      |       |   |   |       |       |       |       |       |
      O      sád-dle to      me      my      mílk-  whìte   stéed
```

(Sharp 1916, #5)

In such lines, it seems plausible to suppose that the lowest row is not characteristically present as a level of rhythmic analysis, but only makes an occasional appearance when needed. We might depict this kind of "occasional row" as follows:

```
(60)        x               x               x               x
      x     x       x       x       x       x       x       x
      |     | \     |       |       |       |       |       |
      |     x   x   |       |       |       |       |       |
      |     |   |   |       |       |       |       |       |
      O    sád-dle to      me      my      mílk-   whìte   stéed
```

Such an analysis is based on the notion of "beat splitting" proposed by Prince (1989), as well as the general idea of local grid expansion proposed by Lerdahl and Jackendoff 1983, 70-74).

Under this scheme of things, "Short First" textsettings in sung verse become a rough equivalent of spoken-verse Resolution.22 Given this equivalence, it is worth considering the putative differences between Resolution as it has been studied in spoken verse and sung verse.

The line taken here is that sung-verse Resolution is sensitive to rather concrete phonetic factors, in particular to the duration of the first syllable in the interval. Sung-verse Resolution is not strictly limited to cases in which the stressed syllable is light, but rather follows a gradient continuum, becoming gradually less possible with longer first syllables. Resolution in spoken verse, by contrast, is a more categorical affair, referring to specific phonological configurations such as syllable weight and word boundaries.

We believe that art verse Resolution is to be viewed as a more tightly crystallized and grammaticized analogue of the phenomena we have found in sung verse. However, spoken-verse Resolution appears to be closer to its phonetic roots that is commonly realized.

Consider in particular the observation made above: that the exceptional cases of heavy-syllable Resolution in Shakespeare are confined to heavy syllables in antepenultimate position or earlier. Now, phoneticians have established that in English words, stressed antepenults tend to be phonetically shorter than stressed penults (Lehiste 1971; Harris and Umeda 1974; van Santen 1992, 534). For example, the *hap-* of *happening* is typically shorter than the *hap-* of *happen*. In our data,

---

[22] A further parallel: the limitation of Resolution to strong position is not universal in spoken verse; Kiparsky (1989) observes weak position Resolution in Hopkins's Sprung Rhythm, and Hanson (1991) notes it as a crucial ingredient of the "iambic-anapestic" verse of Tennyson. Our data corpus also includes many cases of what would count as the sung-verse analogue of weak-position Resolution:

```
(i)           x                   x                   x                   x
     x        x         x         x         x         x         x         x
    _x    x   x    x    x    x    x    x    x    x    x    x    x    x    x    x
     |        |         |    |    |         |         |         |         |
    Two     swéet     lit-tle bábes        to        her       were      bórn.
```

(Sharp n.d., p. 4)

we find that this durational distinction leads to a large difference in the use of sung-verse Resolution (a.k.a. Short First/Scotch Snap). In the full set of folksong data studied here (original settings plus ten consultants), heavy syllables are resolved only 59.4% of the time when in penultimate position of the word (cf. (38b)), but fully 86.5% of the time when antepenultimate. A chi square test indicates that this disparity is significant; p = .0002.

We conjecture, then, that the structural description for Shakespeare's Resolution is as follows: on the metrical side, we follow Prince in assuming that Resolution is limited to strong positions; on the phonological side it requires either a *light* or a *pre-penultimate* stressed syllable. Plainly, this rule is inelegant in purely structural terms. It has a coherent basis, however, in phonetics: the disjunction of the terms “light” and “pre- penultimate” isolates a set of syllables in English that are likely to be phonetically short.[23]

8 Conclusions

We have reached the following, often tentative, conclusions.

First, our study has provided a fair amount of replication, from sung- verse evidence, of earlier work on spoken verse. (a) The idea of a binary- branching hierarchical pattern for our folk verse lines is supported by the echoing of this pattern in P-structure (section 6.2). (b) The typology proposed by Hayes (1989) for how rules interact with P-structure has been partially confirmed: the data examined here show clear bounding effects at the word level, as well as the expected increases of metrical strictness at the right edges of words and of higher-ranking prosodic categories; the latter following the predicted gradient pattern. While the typology's predictions about left-edge effects largely failed to materialize in the data, we have located reasons that would account for this failure, based on fundamental differences between spoken and sung verse: left-edge inversion in sung verse, but not spoken verse, is virtually always a poor choice given the alternatives.

Second, we have obtained indirect empirical support for the theory of P- structure and for the specific rules proposed for English in Hayes (1989), insofar as these proposals can illuminate patterns of textsetting. Note that in a number of the figures in this article, a given textsetting pattern is favored in a gradient way, referring to the various layers of P-structure in a monotonic pattern. The chart below indicates significance levels (from a chi- square test, on the full data set) for adjacent levels on the hierarchy, for various figures presented in this article; “n.s.” = “not significant”.

[23] We suspect that structural descriptions in phonology likewise display this property: where they are structurally inelegant, this indicates an “effort” of the system to reflect idiosyncrasies of phonetics.

(61)	light vs. heavy syllables, word-internal	word-internal vs. word right edge	Word vs. Clitic Group	Clitic Group vs. Phonological Phrase	Phonological Phrase vs. Intonational Phrase
Fig. 2, Short First	<.0001	<.0001	<.0001	n.s.	.0006
Fig. 3, Scotch snaps	<.0001	n.s.	.001	<.0001	.0089
Fig. 4, Dotted rhythms	n.s.	n.s.	.020	.0017	<.0001
Fig. 6, bounding domain of stress mismatch			<.0001		
Fig. 7, mismatches at right edges				<.0001	<.0001
Fig. 8, bounding domain of stress mismatch (right edges only)			<.0001		

While there are unexplained gaps in the significance levels, reassuringly, they occur at different locations in hierarchy. The data together strongly suggest that all five layers of the hierarchy postulated here play a crucial explanatory role.

In this connection, we wish to mention the Clitic Group, a layer of P- structure whose existence has been controversial—cf. Zec (1988) and Inkelas (1989), who propose elegant accounts that use the theory of Lexical Phonology to subsume the Clitic Group under the Word. We caution here that our data should not be taken as conclusive proof of the existence of Clitic Groups. An alternative possibility is that function words are phrased in a flexible manner, sometimes incorporated into a neighboring Word, sometimes not, giving rise to an overall statistical pattern intermediate between that shown by single Words and separate Words within the same Phonological Phrase.

A third result concerns the apparent relevance of low-level phonetic detail in textsetting. The Syllable Duration rule (20) appears to rely crucially on aspects of sound structure that are phonetic, not phonological, namely the gradient amounts of duration added at various right edges of P-structure, and the durational difference between antepenultimate and penultimate syllables (section 7). We suspect that Syllable Duration respects even the subphonemic durational difference between low and high vowels, although the few relevant data we have fall just short of statistical significance.[24] We see no obvious way to restate Syllable Duration in terms of categorial structure

[24] As an intuitive test, the reader may wish to re-enact the comparison in (19) ("'Twas in Oxford City I lost my way"), juxtaposing the original with a revised line in which the very short high vowel [ɪ] has been replaced by the phonetically longer (but still categorially short) low vowel [æ]: "'Twas an Oxford laddie that lost his way."

without massive loss of generality: a rule that referred only to phonological information (P-structure, weight, and vowel height) would have to recapitulate the structural descriptions of many of the rules of the phonetic component, namely, those which compute phonetic syllable durations. Our results thus suggest that the limitation of metrical rules to phonemic material, proposed by Jakobson (1933) and echoed enthusiastically by Hayes (1988), is probably mistaken. We think that an approach to metrics that invokes the full phonetic richness of spoken language, rather than just phonemic representations, may well be fruitful in future research.[25]

Finally, we offer our account of line-final lexical inversion as evidence that some things in metrics happen *faute de mieux*: that certain linguistically-common sequences get metrically arranged in the way that does the least harm, there being no arrangement for them that is ideal from all points of view. Such a view is taken directly from the general philosophy of Optimality Theory (Prince and Smolensky 1993); and we plan to show in future work (Hayes and Kaun, in progress) that Optimality Theory also provides a workable basis for explicit and effective formal grammars of textsetting.

[25] For similar conclusions in phonology, see Steriade (1993), Jun (1995), Flemming (1995), Kaun (1995), Silverman (1995), Kirchner (in progress), and Hayes (in progress).

References

Bailey, James (1975). *Toward a Statistical Analysis of English Verse*. Lisse: Peter de Ridder.

Bailey, James (1979). The Russian linguistic-statistical method for studying poetic rhythm. *Slavic and East European Journal* 23: 251-261.

Bertheau, Florence (1994). A study of the effects of syllable duration in metrical text setting. Ms., Dept. of Linguistics, UCLA, Los Angeles, CA.

Chen, Matthew (1983). Toward a grammar of singing: tune-text association in the Gregorian chant. *Music Perception* 1: 84-122.

Chen, Matthew (1984). The interfacing of language and music. *Language and Communication* 4: 159-194.

Dauer, Rebecca (1983). Stress-timing and syllable-timing reanalyzed. *Journal of Phonetics* 11: 51-62.

Dell, François (1975). Concordances rythmiques entre la musique et les paroles dans le chant: l'accent de l'*e* muet dans la chanson fran‡aise. In *Le soucie des apparences,* Marc Dominicy (ed.), 121-136. Brussels: Editions de l'Université de Bruxelles.

Evans, Robert O. (1966). *Milton's Elisions*. University of Florida Monographs, Humanities, No. 21. Gainesville: University of Florida Press.

Flemming, Edward (1995). *Auditory Representations in Phonology*. Unpublished Ph.D. dissertation, Dept.of Linguistics, UCLA, Los Angeles, CA.

Gerould, Gordon H. (1932). *The Ballad of Tradition*. Clarendon Press, Oxford.

Halle, John and Fred Lerdahl (1993). A generative textsetting model. *Music Perception* 55: 3-23.

Halle, Morris and S. Jay Keyser (1971). *English Stress: Its Form, Its Growth, and its Role in Verse*. New York: Harper and Row.

Hanson, Kristin (1991). *Resolution in Modern Meters*. Unpublished Ph.D. dissertation, Stanford University, Stanford, CA.

Hanson, Kristin (1993). Resolution: Evidence from modern English metrics. *North Eastern Linguistic Society* 23: 159-173.

Harris, M. S. and N. Umeda (1974). Effect of speaking mode on temporal factors in speech: vowel duration. *Journal of the Acoustical Society of America* 56: 1016-1018.

Hayes, Bruce (1983). A grid-based theory of English meter. *Linguistic Inquiry* 14: 357-393.

Hayes, Bruce (1988). Metrics and phonological theory. In *Linguistics: The Cambridge Survey, Vol. 2*, Frederick Newmeyer (ed.), 220-249. Cambridge: Cambridge University Press.

Hayes, Bruce (1989). The Prosodic Hierarchy in meter. In Kiparsky and Youmans (1989), pp. 201-260.

Hayes, Bruce (in progress). A phonetically-driven, optimality-theoretic account of post-nasal voicing. Ms., Dept. of Linguistics, UCLA, Los Angeles, CA.

Hayes, Bruce and Abigail Kaun (in progress). An Optimality-theoretic grammar for textsetting. Ms., UCLA, Los Angeles, CA, and Yale University, New Haven, CT.

Hayes, Bruce and Margaret MacEachern (in progress). Folk verse form in English. Ms., UCLA, Los Angeles, CA.

Hendren, J. W. (1936). *A Study of Ballad Rhythm*. Princeton Studies in English 14, Princeton: Princeton University Press.

Inkelas, Sharon (1989). *Prosodic Constituency in the Lexicon*. Unpublished Ph.D. dissertation, Stanford University, Stanford, CA.

Inkelas, Sharon and Draga Zec, eds., (1990). *The Syntax-Phonology Connection*. Chicago: University of Chicago Press.

Jakobson, Roman (1933). Über den Versbau der serbokroatischen Volksepen. *Archives néerlandaises de phonétique expérimentale* 7-9, 44-53. [Reprinted 1966 in *Selected Writings IV: Slavic Epic Studies*, 51-60. The Hague: Mouton.]

Jespersen, Otto (1900, English translation 1933). Notes on meter. In *Linguistica*, 249-274. Copenhagen: Levin and Munksgaard.

Jun, Jongho (1995). *Perceptual and Articulatory Factors in Place Assimilation: An Optimality-Theoretic Approach*. Unpublished Ph.D. dissertation, UCLA, Los Angeles, CA.

Karpeles, Maud, ed., (1932). *English Folk Songs from the Southern Appalachians*. London: Oxford University Press.

Karpeles, Maud (1973). *An Introduction to English Folk Song*. London: Oxford University Press.

Karpeles, Maud, ed. (1974). *Cecil Sharp's Collection of English Folk Songs*. London: Oxford University Press.

Kaun, Abigail (1995). *The Typology of Rounding Harmony: An Optimality- Theoretic Approach*. Unpublished Ph.D. dissertation, UCLA, Los Angeles, CA.

Kiparsky, Paul (1968). Metrics and morphophonemics in the *Kalevala*. In *Studies Presented to Roman Jakobson by his Students*, Charles Gribble (ed.), 137-148. Cambridge: Slavica.

Kiparsky, Paul (1975). Stress, syntax, and meter. *Language* 51: 576-616.

Kiparsky, Paul (1977). The rhythmic structure of English verse, *Linguistic Inquiry* 8: 189-248.

Kiparsky, Paul (1987). On theory and interpretation. In *The Linguistics of Writing: Arguments Between Language and Literature*, Nigel Fabb, Derek Attridge, Alan Durant and Colin MacCabe (eds.), 185-198. Manchester: Manchester University Press.

Kiparsky, Paul (1989). Sprung rhythm. In Kiparsky and Youmans (1989), pp. 305-340.

Kiparsky, Paul and Gilbert Youmans (1989). *Rhythm and Meter*. San Diego, CA: Academic Press.

Kirchner, Robert (in progress). *Lenition in Optimality-Theoretic Phonology*. Unpublished Ph.D. dissertation, UCLA, Los Angeles, CA.

Ladd, D. Robert and W. N. Campbell (1991). Theories of prosodic structure: evidence from syllable duration. *Actes du XIIème congrès international des sciences phonétiques*, Vol. 2, 290-293. Aix-en-Provence: Service des Publications, Université de Provence.

Lehiste, Ilse (1971). The timing of utterances and linguistic boundaries. *Journal of the Acoustical Society of America* 51: 2018-2024.

Lerdahl, Fred and Ray Jackendoff (1983). *A Generative Theory of Tonal Music*. Cambridge, MA: MIT Press.

Liberman, Mark (1975). *The Intonational System of English*. Doctoral dissertation, MIT, Cambridge, MA. [Distributed 1978 by Indiana University Linguistics Club, Bloomington, IN.]

Magnuson, Karl and Frank G. Ryder (1971). Second thoughts on English prosody. *College English* 33: 198-216.

Nespor, Marina and Irene Vogel (1982). Prosodic domains of external sandhi rules. In *The Structure of Phonological Representations, Part I,* Harry van der Hulst and Norval Smith (eds.), 225-255. Dordrecht: Foris.

Nespor, Marina and Irene Vogel (1986). *Prosodic Phonology*. Dordrecht: Foris.

Oehrle, Richard (1989). Temporal structures in verse design. In Kiparsky and Youmans (1989), pp. 87-119.

Piera, Carlos (1980). *Spanish Verse and the Theory of Meter*. Unpublished Ph.D. dissertation, Program in Romance Linguistics and Literature, UCLA, Los Angeles, CA.

Pierrehumbert, Janet and Mary Beckman (1988). *Japanese Tone Structure*. Cambridge, MA: MIT Press.

Prince, Alan (1989). Metrical forms. In Kiparsky and Youmans (1989), pp. 45- 80.

Prince, Alan and Paul Smolensky (1993). *Optimality Theory: Constraint Interaction in Generative Grammar*. Ms., Dept. of Linguistics, Rutgers University, New Brunswick, NJ, and Dept. of Computer Science, University of Colorado at Boulder.

Reichard, Claude M. (1981). *Meter and Rhythm in English Verse: Towards a Musical Synthesis*. Unpublished Ph.D. dissertation, Stanford, CA.

Ritchie, Jean (1965). *Folk Songs of the Southern Appalachians as Sung by Jean Ritchie*. New York: Oak Publications.

Selkirk, Elisabeth O. (1978). On prosodic structure and its relation to syntactic structure. In *Nordic Prosody II*, T. Fretheim (ed.). Trondheim: TAPIR.

Selkirk, Elisabeth O. (1980). Prosodic domains in phonology: Sanskrit revisited. In *Juncture*, Mark Aronoff and Mary-Louise Kean (eds.), 107- 129. Saratoga, CA: Anma Libri.

Selkirk, Elisabeth O. (1981). On the nature of phonological representation. In *The Cognitive Representation of Speech*, Terry Myers, John Anderson, and John Laver (eds.), 379-388. Amsterdam: North Holland.

Selkirk, Elisabeth O. (1984). *Phonology and Syntax: The Relation between Sound and Structure*. Cambridge, MA: MIT Press.

Selkirk, Elisabeth O. (1986). On derived domains in sentence phonology. *Phonology Yearbook* 3: 371-405.

Sharp, Cecil J. and Ralph Vaughan Williams, eds. (1951a). *A Selection of Collected Folk-Songs*.

Sharp, Cecil J. and Ralph Vaughan Williams, eds. (1951b). *A Selection of Some Less Well Known Folk-Songs*. London: Novello.

Sharp, Cecil J., ed. (1916). *One Hundred English Folksongs*. Boston: Oliver Ditson.

Sharp, Cecil J., ed. (n.d.). *Folk Songs of English Origin Collected in the Appalachian Mountains*. London: Novello.

Silverman, Daniel (1995). *Phasing and Recoverability*. Unpublished Ph.D. dissertation, UCLA, Los Angeles, CA.

Sprott, Ernest S. (1953). *Milton's Art of Prosody*. Oxford: Blackwell.

Stein, David and David Gil (1980). Prosodic structures and prosodic markers. *Theoretical Linguistics* 7, 173-240.

Steriade, Donca (1993). Positional neutralization. Paper given at the 1993 meeting of the North Eastern Linguistic Society, Amherst, MA.

Swiger, Tami (1994). Japanese folk and children's songs: a generative textsetting model. Ms., Dept. of Linguistics, UCLA, Los Angeles, CA.

Tarlinskaja, Marina (1976). *English Verse: Theory and History*. The Hague: Mouton.

Tarlinskaja, Marina (1984). Rhythm - morphology - syntax - rhythm. *Style* 18: 1-26.

Tarlinskaja, Marina (1987). *Shakespeare's Verse: Iambic Pentameter and the Poet's Idiosyncrasies*. New York: Peter Lang.

Tarlinskaja, Marina (1993). *Strict Stress-Meter in English Poetry Compared with German and Russian*. Calgary: University of Calgary Press.

Tarlinskaja, Marina and L. M. Teterina (1974). Verse - prose - meter. *Linguistics* 129: 63-86.
van Santen, J. P. H. (1992). Contextual effects on vowel duration. *Speech Communication* 11: 513-546.
Vaughan Williams, Ralph and A. L. Lloyd (1959). *The Penguin Book of English Folk Song*. Harmondworth, Middlesex: Penguin.
Wightman, Colin W., Stefanie Shattuck-Hufnagel, Mari Ostendorf, and Patti J. Price (1992). Segmental durations in the vicinity of prosodic phrase boundaries. *Journal of the Acoustical Society of America* 91: 1707-1717. Youmans, Gilbert (1989). Milton's meter. In Kiparsky and Youmans (1989).
Young, George (1923). *An English Prosody on Inductive Lines*. Cambridge: Cambridge University Press.
Zec, Draga (1988). *Sonority Constraints on Prosodic Structure*. Unpublished Ph.D. dissertation, Stanford University, Stanford, CA.
Zwicky, Arnold and Ellen Kaisse, eds. (1987). *Syntactic Conditions On Phonological Rules*. Special topics issue, *Phonology* 4.

(Hayes)
Department of Linguistics
UCLA
Los Angeles, CA 90095-1543
e-mail: bhayes@humnet.ucla.edu

(Kaun)
Department of Linguistics
P.O. Box 1504A
Yale Station
New Haven, CT 06520