Phonology **28** (2011) 235–282.
doi:10.1017/S0952675711000157

*Gerard Manley Hopkins' sprung rhythm: corpus study and stochastic grammar**

Bruce Hayes
University of California, Los Angeles

Claire Moore-Cantwell
University of Massachusetts, Amherst

Sprung rhythm is a complex poetic metre invented and used by Gerard Manley Hopkins. We re-examine and amplify a seminal analysis of this metre by Kiparsky (1989). We coded the sprung rhythm corpus for stress, weight and phrasing, then used a computer program to locate every scansion compatible with Kiparsky's analysis. The analysis appears to be nearly exceptionless. However, it is incomplete in that it permits dozens or even hundreds of scansions for certain lines. We propose a Parsability Principle for metrics mandating that ambiguity of scansion be minimised, and suggest that under this proposal, the Kiparskyan system is not a possible metre. Our own revised analysis adds ten new constraints and is cast in the form of a stochastic maxent grammar. It produces an acceptably low level of ambiguity in metrical parsing, and is supported by a good match to the diacritics Hopkins employed to mark his intended scansion.

> For that piece of mine is very highly wrought. The long lines are not rhythm run to seed: everything is weighed and timed in them. Wait till they have taken hold of your ear and you will find it so.
>
> (Gerard Manley Hopkins, in a letter to Robert Bridges (18 October 1881); text from Abbott 1935b)

* For helpful input we would like to thank Joan Bresnan, Robert Daland, Dieter Gunkel, Kristin Hanson, Karen Jesney, Patricia Keating, Wendell Kimper, John McCarthy, Donka Minkova, Joe Pater, Colin Wilson, Kie Zuraw and audience members at the UCLA Phonology Seminar, the University of Massachusetts Phonology Group and the 2010 meeting of the Linguistic Society of America, Baltimore. Thanks to Caren Rotello of the University of Massachusetts Department of Psychology for assistance with statistical testing, as well to four anonymous reviewers for *Phonology* and the associate editor for their helpful comments. To Paul Kiparsky our special thanks, not just for his comments on the manuscript, but also for providing the original basis of our project. This research was supported by the Committee on Research of the UCLA Academic Senate.

1 Introduction

The celebrated Victorian poet Gerard Manley Hopkins wrote some of his best-known poems in a metre he invented and named 'sprung rhythm'. This metre differs in many ways from the standard forms of English metre, and constitutes an indisputably hard case for the theory of metrics. Hopkins' own efforts to explain the metre to his friends provide only clues, not an explanation, of how sprung rhythm works.

Our study takes as its starting point the work of Kiparsky (1989), which offers an intricate and original analysis of this metre. Departing from previous accounts, Kiparsky proposed that the key to understanding sprung rhythm is syllable weight, defined in ways partly standard, partly specific to Hopkins. Kiparsky elaborated this idea into an analysis that makes precise predictions about whether or not any given line is a legal instantiation of sprung rhythm.

We think the time has come to return to the sprung rhythm corpus, subjecting it to further scrutiny. Any intricate proposal in linguistic theory such as Kiparsky's should be tested as carefully as possible against the data that motivated it. Furthermore, the resources now available for studying sprung rhythm – textual, theoretical and computational – are richer today than they were in the 1980s, when Kiparsky prepared his study. We use digital technology to explore two main aspects of Kiparsky's system: occurrence of exceptions and number of possible scansions per line. In addition, we apply more recent developments in stochastic (probabilistic) constraint-based grammatical frameworks (Boersma & Hayes 2001, Goldwater & Johnson 2003, Boersma & Pater 2008) to the problems of scansion selection and optionality in Hopkins' verse. We argue that the use of stochastic grammar is essential to a full understanding of sprung rhythm.

Our methodology has three components: corpus preparation, machine scansion, and modelling with stochastic grammar. First, we annotated the corpus of sprung rhythm poems, syllable by syllable, for the relevant phonological information: stress level, phonological phrasing and syllable weight. Second, we devised a computer program that inspects this corpus, discovering for each line the complete set of scansions that are legal under Kiparsky's grammar. Scrutinising the corpus with these tools, we confirm that exceptions to Kiparsky's analysis, that is, lines in the corpus which are defined as unmetrical by his grammar, are indeed very rare – specifically, they are rarer than expected, based on a random corpus of comparable prose lines.

Despite this success, we argue that the Kiparskyan system as it stands is incomplete, and indeed that it does not describe a possible metre. It often permits dozens of different scansions per line, occasionally even hundreds. We consider the typology of metrical systems, and find a general absence of metres with many possible scansions per line. We propose that metrical systems are governed by a PARSABILITY PRINCIPLE, which requires that the scansion be in most cases recoverable from the phonological form of a line. Sprung rhythm as described in the Kiparskyan system is incompatible with this principle.

Lastly, we augment the Kiparskyan grammar in a way that solves this problem. We recast the system as a stochastic grammar using Kiparsky's rules (recast as inviolable constraints) together with a set of violable constraints. Our grammar assigns probabilities to lines rather than making up-or-down decisions, thus allowing multiple scansions per line while also distinguishing marked from unmarked ones and providing a reasonable degree of parsability. This grammar predicts the scansions that Hopkins intuitively preferred, as shown by the location of the diacritics he included in his poems to help readers with scanning, and allocates most of its output probability to just one or a few scansions for each line.

Our work has benefited from the study of Hanson (1992: §5.1), which re-examines the sprung rhythm corpus and proposes specific revisions to Kiparsky's theory discussed below; as well as Hanson & Kiparsky (1996), which integrates both Kiparsky & Hanson's accounts into a general parametric theory of metre.[1]

2 Background

2.1 Gerard Manley Hopkins (1844–1889)

Gerard Manley Hopkins was educated at Oxford and was converted there to Roman Catholicism. His adult life was spent as an ordained Catholic priest, a member of the Jesuit order. He was little known in his lifetime and for the most part circulated his poems only among family and friends. The latter included sympathetic fellow poets, with whom Hopkins carried on a correspondence that survives and helps illuminate his work. Hopkins' surviving poems were first published only in 1918, long after his death, but since then the esteem for his work has gradually risen, so that many now view him as one of the most important poets of his time. His poetry is dense, often difficult, and bold in its technical innovations.

Hopkins could fairly be called a 'phonologist's poet'. His letters show a striking talent for phonetic and phonological observation, as well as a penetrating understanding of metre; notably, he anticipated the core idea of generative metrics (Halle & Keyser 1966) that the metre must be construed as an abstract entity, not directly observable in the phonological form of lines,[2] and he also shows flashes of insight in his struggles to make clear to others the basis of his sprung rhythm.[3]

[1] We have also consulted recent work by literature scholars on sprung rhythm that has addressed Kiparsky's study: Holder (1995), Wimsatt (1998, 2006) and Hurley (2005). We have not found this work helpful for our purpose, since it does not in general address the distributional evidence on which Kiparsky's system rests.

[2] For Hopkins as phonetician/phonologist, see his essay in House (1959: 267–288), and the correspondence in Abbott (1935a: 41, 1935b: 44, 180, 1956: 273). For Hopkins as metrist, see Abbott (1935b: 41, 1956: 109, 328).

[3] Hopkins' writings should *not* be considered an infallible guide. On some occasions he seems to have been quite wide of the mark, for instance when he cites a variety of

2.2 Phonological preliminaries

As will be apparent below, analysis of Hopkins' verse requires an accurate characterisation of the phonemes, stress patterns and lexical representations of Hopkins' native language, Standard British English. For extensive discussion of the contrasts and surface phonology of this dialect, see Wells (1982). To check pronunciation of individual words, we consulted the *Oxford English Dictionary*.

Second, analysis requires an account of syllable division, since it is on this basis that syllables are classified into heavy and light. We assume a syllabification process that (a) respects the Maximum Onset Principle (e.g. Kenstowicz 1994: 256–261), (b) is word-bounded (Nespor & Vogel 2007: 72–81) and (c) at least for purposes of weight computation, involves no ambisyllabic consonants (Kahn 1976, Gussenhoven 1986). The actual syllable weight categories needed are heavy (symbol [¯]), light ([˘]) and ambiguous ([⏓]). The basis of these weight categories is discussed in detail in §3.4 below; until that point we ask the reader to take our weight symbols as given.

Third, we need to describe the stress contours borne by English phrases and sentences. We assume the basic patterns embodied in Chomsky & Halle's Nuclear Stress Rule and Compound Stress Rule (1968: 89–94), as well as rules creating phrasal rhythmic alternation (Beat Addition, the Rhythm Rule); see Selkirk (1984) and Hayes (1995, 2009). We use integers to indicate relative stress level.

Lastly, we need to have a characterisation of phonological phrasing. We assume a version of the Prosodic Hierarchy (see Selkirk 1980 and the large research literature that has developed from it), in which Phonological Words are grouped into Clitic Groups, which are grouped into Phonological Phrases, which are grouped into Intonational Phrases, which are grouped into phonological Utterances. We assume the principle of Strict Layering (Selkirk 1984), whereby each category dominates only members of the immediately lower ranking category. We follow the outline principles for phrasing given in Hayes (1989).

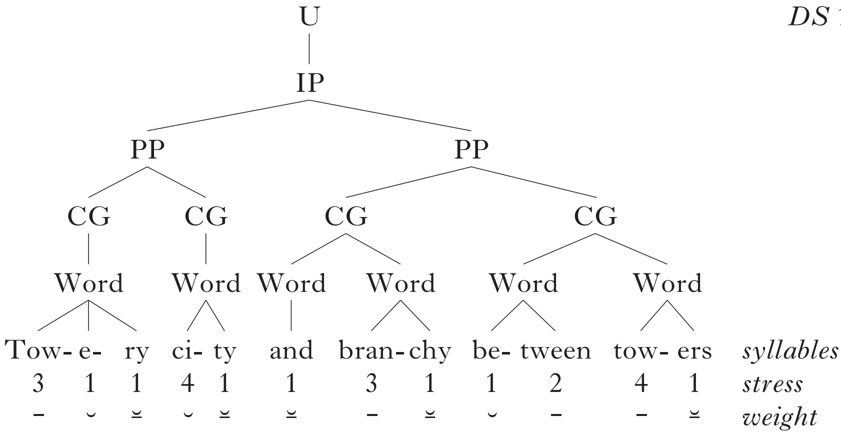
We coded all the lines of the Hopkins sprung rhythm corpus with these structures, using various forms of digital shorthand. (1) gives our representation of an example line (the phrasing levels are abbreviated as follows: U = Utterance; IP = Intonational Phrase; PP = Phonological Phrase; CG = Clitic Group).[4]

other verse systems as instances of sprung rhythm (MacKenzie 1990: 117). Our practice follows the norm of linguistic scholarship, which assumes that the *analysis* of the data cannot be obtained directly from the native speaker. The Hopkins correspondence gives us hints concerning why certain lines felt right to Hopkins as sprung rhythm (and others, such as Robert Bridges' attempts, did not; Abbott 1935b: 81). But in articulating the analysis itself, Hopkins was just one struggling metrist among many.

[4] For abbreviations of poem titles see Appendix C.
Our full set of transcriptions is available in the supplementary online materials at http://journals.cambridge.org/issue_Phonology/Vol28No02.

(1) *A sprung rhythm line coded for phrasing, stress and syllable weight*

DS 1



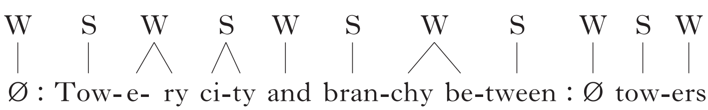
3 The rules for sprung rhythm

In this section, we summarise Kiparsky's original analysis. Since we will ultimately be augmenting the analysis in a constraint-based framework, our presentation re-expresses the Kiparskyan analysis as a set of inviolable constraints.

3.1 Basic rules for metrical positions and what can fill them

A METRE in sprung rhythm consists of an alternating sequence of S (Strong) and W (Weak) positions, beginning and ending with W. The metres are classified according to the number of S positions they contain. For example, the metrical template for a sprung rhythm tetrameter is, approximately, WSWSWSWSW (this is modified in §3.2 below). Following classical views in generative metrics (e.g. Kiparsky 1977: 190), the metre serves as a kind of measuring stick for lines: a line is metrical if all of its syllables can be aligned to the metre according to the rules.

A SCANSION of a line is an alignment of its syllables with the S and W positions of the metre. In sprung rhythm, a position can be aligned with more than one syllable, and sometimes with no syllable, shown here as Ø. Here is the line of (1) shown in legal alignment with the template for sprung rhythm pentameter.[5]

(2) 

[5] The large colon, subscript arcs and accents that appear in Hopkins' lines will be explained in §3.5.

In what follows we will save space by using representations in which labelled slashes separate the metrical positions, as in (3).

(3) $/_{w}$Ø$/_{s}$: Tow-$/_{w}$ery$/_{s}$city$/_{w}$and$/_{s}$bran-$/_{w}$chy be-$/_{s}$tween :$/_{w}$Ø$/_{s}$tow-$/_{w}$ers;

Alignment to the metre follows rules specific to S and W positions. For each, we first give the rules, then discussion and examples.

3.1.1 *W position*

(4) *Legal sequences filling W positions*
  a. a single stressless syllable
  b. a sequence of stressless light syllables
  c. a stressed monosyllable
  d. a resolved sequence (defined below)
  e. null

3.1.1.1 *W as any single stressless syllable*. This is the most common realisation, and instances are too numerous to mention here.

3.1.1.2 *W as any sequence of stressless light syllables*. In practice, the upper limit seems to be four; here is an example with three ([təd] and [ɪn] can be light syllables in the Hopkins weight system).

(5) $/_{w}$As a$/_{s}$dare$/_{w}$gale$/_{s}$sky-$/_{w}$lark$/_{s}$scan-$/_{w}$*ted in a*$/_{s}$dull :$/_{w}$Ø$/_{s}$cage   *CS* 1
[təd ɪn ə]

3.1.1.3 *W as a stressed monosyllable*. As is true for much English verse (Magnuson & Ryder 1970, Kiparsky 1975), a stressed syllable may occur in weak position when it is the only syllable of a monosyllabic word. This is true, for instance, of the second word of (6).

(6) $/_{s}$March,$/_{w}$*kind*$/_{s}$com-$/_{w}$rade, a-$/_{s}$breast$/_{w}$him;   *BC* 30

The line would be unmetrical with the stressed syllable of a polysyllabic word in weak position.

(7) *$/_{s}$March,$/_{w}$*com*-$/_{s}$rade,$/_{w}$a-$/_{s}$breast$/_{w}$him;   (*construct*)

It would likewise be unmetrical if we substituted a polysyllabic word in the same location.

(8) *$/_{s}$March,$/_{w}$*faithful*$/_{s}$com-$/_{w}$rade, a-$/_{s}$breast$/_{w}$him;   (*construct*)

3.1.1.4 *W as a resolved sequence*. A resolved sequence is defined as a stressed light followed by a stressless non-heavy syllable in the same word. Resolved sequences act as units elsewhere in English

(Old English: Russom 1987, Dresher & Lahiri 1991; Chaucer: Young 1928: §112); and in iambic pentameter they occasionally are allowed to fill an S position (Kiparsky 1977: 236, Prince 1989: 53). In sprung rhythm, Hopkins extends the practice by allowing it in W as well. Here is an example; in *yellow* both [ˈjɛ] and stressless final [loʊ] may count as light syllables for Hopkins (see §3.4).

(9) $/_{\mathrm{w}}$Her $/_{\mathrm{s}}$fond $/_{\mathrm{w}}$*yellow* $/_{\mathrm{s}}$horn- $/_{\mathrm{w}}$light $/_{\mathrm{s}}$wound $/_{\mathrm{w}}$to the $/_{\mathrm{s}}$west, *SS* 3, first hemistich
[ˈjɛloʊ]

The ability of resolved sequences to appear in W must be construed as overriding the otherwise general ban on lexical stress in W. The unmetrical example (8) illustrates this; since the first syllable of *faithful* is heavy, it cannot be accommodated in W under resolution. In the view of Hanson (1992: 139) and Hanson & Kiparsky (1996), the special privilege of resolution can be formalised as correspondence between a metrical position and a (phonological) foot, specifically a moraic trochee, in the sense of Hayes (1995) and other work.[6]

3.1.1.5 *Null.* The use of null W positions is what led Hopkins to call his metre 'sprung'. An extravagant example, given in context, is shown in (10).

(10) My aspens dear, whose airy cages quelled
Quélled or quenched in leaves the leaping sun,
$/_{\mathrm{w}}$ Ø $/_{\mathrm{s}}$Áll $/_{\mathrm{w}}$ Ø $/_{\mathrm{s}}$félled, $/_{\mathrm{w}}$ Ø $/_{\mathrm{s}}$félled, $/_{\mathrm{w}}$are $/_{\mathrm{s}}$áll $/_{\mathrm{w}}$ Ø $/_{\mathrm{s}}$félled; *BP* 1–3

In the manuscripts the null W positions are often marked diacritically by Hopkins using what he called a 'great colon' (a colon written larger than normal).

In Kiparsky's description, the ways that W can be filled are stated as licenses, but the pattern can also be translated into constraints.[7] In the statement of (11b), we employ a widely used term of generative metrics: a LEXICAL STRESS (Kiparsky 1977) is a stress in a polysyllabic word; it is

[6] While correspondence to a foot rationalises the appearance of moraic trochees in W position, we caution that a theory in which *all* the scanned entities are phonological feet is unlikely to be tenable. It is inconsistent with the Kiparskyan criteria for weight, since light + heavy words like *damask*, *spirits* and *mammocks* (from *DO* 11, *SS* 9 and *TG* 12) would have to occupy single positions. It is also inconsistent with the (roughly ten) cases in which Hopkins places an outride mark on the second syllable of a word whose first syllable is light, e.g. *river* (*BP* 8), *babble* (*CS* 10), *lashes* (*HF* 4); cf. Kiparsky (1989: 318–319). In these cases, a scansion compatible with the diacritics must have a light syllable in strong position.

[7] We follow the constraint definition schema recommended for constraints in Optimality Theory by McCarthy (2003).

called ‘lexical’ because its greater stress relative to its neighbour(s) is part of the lexical representation.

(11) *Constraints for filling W*

a. *LEXICAL STRESS IN W
   Assign a violation to every lexical stress in a weak position that is not part of a resolved sequence.

b. *HEAVY IN MULTIPLY FILLED POSITION
   Assign a violation to every position containing more than one syllable, one of which is heavy.

Neither of these constraints is unique to sprung rhythm. The exclusion of lexical stresses from W (11a) is found throughout the metrics of English as well as German and Russian (Magnuson & Ryder 1970, Bailey 1975, Kiparsky 1975). The ban on heavy syllables in multiply filled W positions is documented by Hanson (1992) for the ‘binary-ternary’ verse of Hopkins’ contemporaries, Tennyson and Swinburne.

3.1.2 *S position.* S in sprung rhythm may be filled with any of the sequences in (12).

(12) *Legal sequences filling* S *positions*

a. a single stressed syllable
b. a resolved sequence
c. a single stressless syllable, provided it is not light

Here are examples and discussion.

3.1.2.1 *S filled by a single stressed syllable.* This is the normal realisation and examples are abundant.

3.1.2.2 *S filled by a resolved sequence.* S can be filled with a resolved sequence just like W, though examples are less frequent.

(13) $/_{\mathrm{w}}$This $/_{\mathrm{s}}$*very* $/_{\mathrm{w}}$very $/_{\mathrm{s}}$day $/_{\mathrm{w}}$came $/_{\mathrm{s}}$down $/_{\mathrm{w}}$to us $/_{\mathrm{s}}$af- $/_{\mathrm{w}}$ter a $/_{\mathrm{s}}$boon $/_{\mathrm{w}}$he on — *BC* 5
[ˈvɛ̮ɹɪ̮]

3.1.2.3 *S filled by single non-light syllable.* As a less usual option, S may be filled with a single stressless syllable provided it is not light, as in (14).

(14) $/_{\mathrm{w}}$Till a $/_{\mathrm{s}}$life- $/_{\mathrm{w}}$belt $/_{\mathrm{s}}$and $/_{\mathrm{w}}$God’s $/_{\mathrm{s}}$will — *LE* 16.3
[ə̮nd]

Kiparsky's account can again be expressed in constraints, for which we propose (15).

(15) *Constraints for filling* S
  a. *NULL IN S
    Assign a violation to every unfilled S position.
  b. *STRESSLESS LIGHT IN S
    Assign a violation to every S position containing a stressless light syllable.
  c. *MULTIPLY FILLED S
    Assign a violation to every S position containing more than one syllable; resolved sequences are excepted.

Constraint (15a) actually has exceptions, of which there are precisely two; Hopkins wrote these empty Ss with the symbol '…'.[8] We will simplify our treatment slightly by glossing over these two exceptions and treating (15a) as if it were exceptionless. Constraint (15b) would be violated by the hypothetical lines in (16).

(16) a. *$/_{\mathrm{w}}$ Till it $/_{\mathrm{s}}$ streng- $/_{\mathrm{w}}$ then $/_{\mathrm{s}}$ a $/_{\mathrm{w}}$ man's $/_{\mathrm{s}}$ will (*construct*)
[ə̮]

  b. *$/_{\mathrm{w}}$ To An- $/_{\mathrm{s}}$ dro- $/_{\mathrm{w}}$ me- $/_{\mathrm{s}}$ da $/_{\mathrm{w}}$ the $/_{\mathrm{s}}$ will[9] (*construct*)
[də̮]

Both (15a) and (15b) ban highly non-sonorous syllables (of which null is the extreme case) in rhythmically strong position; for phonological analogues of these constraints see de Lacy (2004).

The ban on multiply filled S (15c) seems sensible from the viewpoint of typology: while there are systems that freely allow multiply filled W (e.g. English folk verse), it seems that S is typically constrained to be filled just once.

3.1.3 *Outrides.* Outrides are Hopkins' extension of a common English verse practice, extrametrical syllables. These are syllables that, in most analyses, are not affiliated with a strong or weak metrical position, but float on their own, licensed by specific metrical rules. The conditions for

[8] The lines are *SS* 1 and 'The Leaden Echo and the Golden Echo' 2 (the latter not studied here; see §4 below).

[9] Compare *Now time's Androme<u>da</u> on this rock rude*, from 'Andromeda', a poem written by Hopkins in conventional iambic pentameter; the latter metre imposes no quantitative restrictions on S position (Kiparsky 1989: 319).

extrametricals are fairly consistent across poets; we summarise the main tendencies in (17).

(17) *Conditions on extrametrical syllables in standard metre*

a. An extrametrical syllable must follow an S position (Kiparsky 1977: 231).

b. An extrametrical syllable must be weaker in stress than the syllable occupying the preceding S position. Some poets require that extrametricals be fully stressless.[10]

c. In some poets and genres, extrametricals can occur only at the end of a line (this is true, for instance, in Shakespeare's *Sonnets*). For others, extrametricals can occur line-medially, but only before a strong phonological break; e.g. at the end of an Intonational Phrase.

d. Extrametrical syllables must be followed by a prosodic boundary which is larger than the prosodic boundary between the extrametrical and the preceding S position.

Here are some extrametricals from Shakespeare; it can be seen that each such syllable is more weakly stressed than its left neighbour, and precedes a (usually) major phonological break.

(18) *Extrametrical syllables in Shakespeare*

a. $/_{\mathrm{w}}$Hav- $/_{\mathrm{s}}$ing $/_{\mathrm{w}}$dis- $/_{\mathrm{s}}$pleased $/_{\mathrm{w}}$my $/_{\mathrm{s}}$fa- $/_{\mathrm{ex}}\underline{\text{ther}}$, $/_{\mathrm{w}}$to
$/_{\mathrm{s}}$Law- $/_{\mathrm{w}}$rence's $/_{\mathrm{s}}$cell *Romeo and Juliet* 3.5.234

b. $/_{\mathrm{w}}$E'er $/_{\mathrm{s}}$since $/_{\mathrm{w}}$pur- $/_{\mathrm{s}}$sue $/_{\mathrm{ex}}\underline{\text{me}}$. $/_{\mathrm{w}}$How $/_{\mathrm{s}}$now, $/_{\mathrm{w}}$what $/_{\mathrm{s}}$news
$/_{\mathrm{w}}$from $/_{\mathrm{s}}$her *Twelfth Night* 1.1.22

c. $/_{\mathrm{w}}$The $/_{\mathrm{s}}$thane $/_{\mathrm{w}}$of $/_{\mathrm{s}}$Caw- $/_{\mathrm{ex}}\underline{\text{dor}}$ $/_{\mathrm{w}}$be- $/_{\mathrm{s}}$gan $/_{\mathrm{w}}$a $/_{\mathrm{s}}$dis- $/_{\mathrm{w}}$mal
$/_{\mathrm{s}}$con- $/_{\mathrm{ex}}\underline{\text{flict}}$ *Macbeth* 1.2.53

d. $/_{\mathrm{w}}$That $/_{\mathrm{s}}$is, $/_{\mathrm{w}}$the $/_{\mathrm{s}}$mad- $/_{\mathrm{ex}}\underline{\text{man}}$; $/_{\mathrm{w}}$the $/_{\mathrm{s}}$lov- $/_{\mathrm{w}}$er, $/_{\mathrm{s}}$all $/_{\mathrm{w}}$as
$/_{\mathrm{s}}$fran- $/_{\mathrm{ex}}\underline{\text{tic}}$, *Midsummer Night's Dream* 5.1.10

In his sprung rhythm poetry Hopkins extends the concept of extrametricality and (in his commentaries) names the extra syllables OUTRIDES. Outrides differ from ordinary extrametricals in that they may consist of more than one syllable, under the same conditions given in (4) for

[10] Kiparsky (1989: 309) suggests that on occasion an outride in sprung rhythm can have stronger stress than the syllable in the preceding S. We suggest that the more restrictive account given above ought to be retained here, in light of our examination of the relevant cases. Kiparsky's example *skate's heel* (*WH* 6) could easily have falling stress, not rising, on analogy with similar constructions like ˈ*crow's* ˌ*nest*, ˈ*cat's* ˌ*paw*. His other examples (1989: 323) are instances of what Ladd (1986: 330–331) calls intonational tags: short, intonationally separate elements that continue the intonational contour of the preceding phrase; thus *child* in *FR* 11: *Thy tears that touched my heart, $\underline{\textit{child}}$, Felix, poor Felix Randal*. Ladd argues that such tags are subordinated in stress; if this is correct, then tags would constitute perfectly good outrides.

W positions: they can either consist of all non-heavy, stressless syllables, or can be (rarely) a resolved sequence. Here are some examples of outrides from Hopkins' poetry; we use the poet's own subscript arc notation to mark them.

(19) *Examples of lines with outrides*

a. $/_{\mathrm{s}}$Shares $/_{\mathrm{w}}$their $/_{\mathrm{s}}$best $/_{\mathrm{w}}$gifts $/_{\mathrm{s}}$sure- $/_{\mathrm{o}}$ly, $/_{\mathrm{w}}$fall $/_{\mathrm{s}}$how $/_{\mathrm{w}}$things $/_{\mathrm{s}}$will),[11] *BC* 4

b. $/_{\mathrm{w}}$A- $/_{\mathrm{s}}$round; $/_{\mathrm{w}}$up a- $/_{\mathrm{s}}$bove, $/_{\mathrm{w}}$what $/_{\mathrm{s}}$wind- $/_{\mathrm{o}}$walks!
$/_{\mathrm{w}}$what $/_{\mathrm{s}}$love- $/_{\mathrm{w}}$ly be- $/_{\mathrm{s}}$ha- $/_{\mathrm{w}}$viour *HH* 2

c. $/_{\mathrm{s}}$Breath- $/_{\mathrm{w}}$ing $/_{\mathrm{s}}$bloom $/_{\mathrm{w}}$of a $/_{\mathrm{s}}$chas- $/_{\mathrm{o}}$tity $/_{\mathrm{w}}$in $/_{\mathrm{s}}$man $/_{\mathrm{w}}$sex $/_{\mathrm{s}}$fine). *BC* 16

d. $/_{\mathrm{w}}$Woe, $/_{\mathrm{s}}$wórld- $/_{\mathrm{o}}$sorrow; $/_{\mathrm{w}}$on an $/_{\mathrm{s}}$áge $/_{\mathrm{w}}$old $/_{\mathrm{s}}$án- $/_{\mathrm{w}}$vil $/_{\mathrm{s}}$wínce
$/_{\mathrm{w}}$and $/_{\mathrm{s}}$sing[12] *NW* 6

Example (19a) is a very ordinary outride; (19b) is a stressed outride (although stressed, it is weaker than the stress on the preceding syllable, being part of a compound word). (19c) is a disyllabic outride, and (19d) is a resolved one.

The constraints we assume on outrides are given in (20).

(20) *Constraints on outrides*

a. OUTRIDES MUST FALL
Assign a violation for an outride that fails to have less stress than its left neighbour.

b. OUTRIDES MUST FOLLOW S
Assign a violation for an outride that follows a W position.

c. OUTRIDES MUST COHERE
Assign a violation for an outride that coheres more closely to the following phonological material than to the preceding.

d. OUTRIDES MUST PRECEDE BREAKS
Assign a violation for an outride that fails to occur at the end of a major phonological constituent.[13]

e. *LEXICAL STRESS IN OUTRIDE
Analogous to (11a).

f. *HEAVY IN MULTIPLY FILLED OUTRIDE
Analogous to (11b).

[11] By the principle of overreaving (§3.2), the initial W position of this line is actually occupied by the syllable *he* from the preceding line; we omit it for legibility. The same holds for (19c) and many other lines.

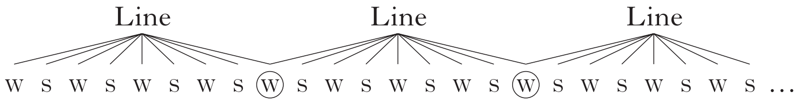
[12] Hopkins here uses a different diacritic, the long superscript arc, which generally implies a sequence of light syllables.

[13] In a total of seven lines, Hopkins includes highly counterintuitive outride marks that violate (20c) and (20d), encompassing syllables that are proclitic (for example

3.2 Overreaving

'Overreaving' is not another metrical license but a restriction, not found (to our knowledge) in standard metres. The metre for a line begins with W and ends with W, but in fact the line-final W and the line-initial one are the *same* position. In effect, this W is 'ambistichic', as suggested in (21).

(21) *Overreaving as ambistichic* W *positions*



Graphically, Hopkins presents his lines as units, placing the syllables where they would belong according to sensible criteria of rhyme and phrasing. But in the metrics, the W is counted as just one position: it cannot bear more material than is allowed in one single W (see (4), (11)). Thus it is easy to construct examples of line pairs that scan individually but could not occur as a concatenation: for example if the first line ends with a W filled by a heavy syllable, and the second begins with a filled W.

3.3 Paraphonology

English poetry is often observed to be based on a count of syllables that is smaller than would be expected from the ordinary phonemic representations of the words. These divergences are systematic and can be predicted using phonological rules[14] – processes of glide formation, vowel deletion and the like. For instance, the phrase *many a* is often treated by poets as two syllables. This plausibly reflects a general glide formation process that converts /ˈmɛ.ni.ə/ to [ˈmɛ.njə]. Kiparsky (1977: 190) proposes to include these rules in a component of the metrical grammar that we will here call the PARAPHONOLOGY. The rules of the paraphonology apply to the normal phonological representations, yielding the representations employed in scansion. For instance, Hopkins often applies a paraphonological rule of desyllabification, converting stressless /i u oʊ ɹ̩ l̩ n̩/ to [j w w ɹ l n] when

BC 12: *Low-latched in leaf-light housel his too huge godhead*) or in one case even word-internal (in *CC* 12). The remaining examples are *BC* 24, *BC* 48, *BP* 8, *CC* 12, *CS* 10 and *HR* 6. Kiparsky (1989: 323) plausibly explains *CC* 12 as a scribal error; the others may conceivably represent substitution of the subscript arc for the superscript, which simply indicates multiply filled W (cf. (19d)). In our transcriptions we have dealt with these by taking Hopkins at his word, placing Clitic Group breaks at the termini of his outrides even where they are not syntactically motivated.

[14] Or their constraint-based equivalents in other theories; the choice matters little here.

they are prevocalic; this reduces the number of syllables and often affects the scansion possibilities. Desyllabification can be fed by a process of context-free [h]-deletion.

We will not treat Hopkins' paraphonology in detail here; it closely resembles the paraphonology of John Milton, whose verse Hopkins admired (for detailed treatments of Miltonic paraphonology, see Bridges 1921 and Sprott 1953). Usually, when Hopkins applied a paraphonological rule, he marked the place with a superscript arc. In the (less common) places where we have had to appeal to paraphonology to get a line to scan, we have indicated this in our web-posted scansions.

3.4 Weight assignment

We complete our summary of the Kiparskyan analysis by stating the principles that classify syllables by weight.

Standard British English vowels can be short ([ɪ ɛ æ ɒ ʌ ʊ ə]) or long (all others, including diphthongs). As in many languages, a syllable counts as light if it is open and short-vowelled and heavy if it is closed or contains a long vowel or diphthong.[15] There is a special provision for final stressless syllables: a single consonant may be ignored, hence final CV̆C may optionally count as short. Unstressed non-low long vowels and diphthongs ([iː uː eɪ əʊ]) also may count as short when final.[16] Hanson (1992: 146) points out parallels to these patterns in English phonology.

Another source of variation is vowel–sonorant merger, which we treat as a paraphonological process. It merges a stressless vowel with a following coda sonorant, creating a short syllabic sonorant. This process can create light syllables all by itself (e.g. stressless *can* can be light; /kən/ → [kn̩]); and it can also apply in conjunction with consonant extrametricality: *and* is optionally light, since /ənd/ can become [n̩d] by paraphonology, with the final /d/ then ignored.[17]

The application of these principles of weight assignment are illustrated in (22).

[15] Syllables closed by [ɹ] are counted as closed, even though coda [ɹ] is dropped in the dialect Hopkins spoke. Hopkins recognised that coda *r* is 'silent', yet judged that 'at long as the *r* is pronounced by anybody, and it is by a good many yet', coda [ɹ] must count as a consonant for purposes of verse (Abbott 1935a: 37). An instance of coda [ɹ] counted as heavy is found in *shoul<u>der</u>* (*HP* 4).

[16] Kiparsky limits this provision to function words, though it is not obvious that this limitation is necessary.

[17] In a small number of cases (Kiparsky 1989: 315, Hanson 1992: 137), we must assign light status to syllables that are not entirely stressless, but are secondarily stressed syllables of content words or of polysyllabic function words: *across* in *HH* 4, *Gala<u>had</u> in BC* 40 and a few others.

(22) *Examples of weight assignment*

		IPA	*weight*	*comment*
a.	*dapple*	[ˈdæ.pəl]	˘ ⏑̄	Final [l] can be ignored; alternatively, /əl/ → [l̩].
b.	*ample*	[ˈæm.pəl]	– ⏑̄	[æm] is closed, hence heavy.
c.	*maple*	[ˈmeɪ.pəl]	– ⏑̄	[eɪ] is long, so [meɪ] is heavy.
d.	*havoc*	[ˈhæ.vək]	˘ ⏑̄	Final [k] optionally ignored.
e.	*dandled*	[ˈdæn.dəld]	– ⏑̄	Light variant: ignore final [d] and apply /əl/ → [l̩].
f.	*the*	[ðə]	˘	Short vowel in open syllable.
g.	*they*	[ðeɪ]	⏑̄	Special provision for final non-low vowels.
h.	*day*	[ˈdeɪ]	–	[eɪ] is stressed, hence heavy.
i.	*by*	[baɪ]	–	[aɪ] is a low diphthong, hence heavy.
j.	*its*	[ɪts]	–	A heavy syllable remains even if [s] is ignored.
k.	*damask*	[ˈdæ.məsk]	˘ –	A heavy syllable remains even if [k] is ignored.

3.4.1 *Correption.* One other paraphonological rule affects weight: CORREPTION shortens a vowel or diphthong when another vowel immediately follows, not necessarily in the same word. For instance, because of Correption, a word like *how* can be treated as light when it precedes a vowel-initial word; or *lion* can be resolved since its first vowel is prevocalic.[18]

Correption relates to the phonology of English in an intriguing way. English words obey a minimum length requirement: no content words may consist of a single light syllable (Pater & Tessier 2003). What happens

[18] Correption is a candidate for a natural phonological process; we have found cases of it in Latin (Mester 1994: 19), Greek (Sihler 1995: 74), Sanskrit (Kessler 1992: 28), Hungarian (Siptár & Törkenczy 2000: 125–128) and Kasem (Callow 1965: 32). For Correption in Philip Sidney's verse, see Hanson (2002: §4.9). Our use of the term follows Kiparsky and ultimately comes from classical philology.

Hopkins himself gave a rough version of a rule of Correption in a letter he wrote to Robert Bridges (Abbott 1935b: 44). He allows Correption not just before vowels, but also before 'semivowels or *r*'. Kiparsky (1989) follows Hopkins' prescription. However, with just one possible exception ((28) below) we think these additional environments are not needed; we find we can scan our corpus using only prevocalic Correption.

Correption sometime appears to be applying before words starting with /h/; thus for example *how he* scanned as /˘ ˘/. This is not a property of Correption *per se*, but rather results from it being fed by /h/-deletion (§3.3).

when Correption applies to a monosyllabic word? As Hanson (1992: 136) observes, the resulting illegal configuration is resolved by treating the word in question as stressless; hence it may occupy a polysyllabic W position.[19]

(23) *Correption with destressing*

a. $/_\mathrm{w}$ Say it is $/_\mathrm{s}$ ásh- $/_\mathrm{w}$ boughs: $/_\mathrm{s}$ whé- $/_\mathrm{w}$ ther on a De- $/_\mathrm{s}$ cém- $/_\mathrm{w}$ ber $/_\mathrm{s}$ day $/_\mathrm{w}$ and $/_\mathrm{s}$ fúrled *AB* 4

b. $/_\mathrm{w}$ Nay in $/_\mathrm{s}$ all $/_\mathrm{w}$ that $/_\mathrm{s}$ toil, $/_\mathrm{w}$ that $/_\mathrm{s}$ coil, $/_\mathrm{w}$ since $/_\mathrm{s}$ (seems) $/_\mathrm{w}$ I $/_\mathrm{s}$ kissed $/_\mathrm{w}$ the $/_\mathrm{s}$ rod, *CC* 10

c. $/_\mathrm{w}$ Now, and $/_\mathrm{s}$ séeing $/_\mathrm{w}$ some- $/_\mathrm{s}$ whére $/_\mathrm{w}$ some $/_\mathrm{s}$ mán | $/_\mathrm{w}$ do $/_\mathrm{s}$ all $/_\mathrm{w}$ that $/_\mathrm{s}$ man $/_\mathrm{w}$ can $/_\mathrm{s}$ do, *SD* 11

It appears that the destressing version of Correption applies only to high-frequency words, which perhaps are already semi-destressed in ordinary speech.

3.5 Hopkins' diacritics

Because Hopkins was frustrated by the inability of his friends to make sense of his new metrical system, he extensively annotated his poems with diacritics. These are interpretable in light of the metrical structures of sprung rhythm as outlined above. In a number of cases, Kiparsky (1989) demonstrated that, given the right analysis, the distribution of Hopkins' diacritics is actually inevitable – even in cases where critics have found them counterintuitive.

The study of the Hopkins diacritics has been facilitated since the publication of Kiparsky (1989) by Mackenzie's (1991) preparation of a complete facsimile edition of Hopkins' manuscripts. We have extracted all the diacritics from this source, guided by Mackenzie's editorial work. In cases where the manuscripts disagreed, our practice was to assume the presence of a diacritic if any one of the manuscripts included it.[20] For overviews of the diacritics, see Mackenzie (1990, 1991) and Moore-Cantwell (2009).

In (24) we list the principal diacritics and their meanings.

[19] There are seven other examples, all with *now*: *LE* 8.2, *LE* 15.2, *LE* 16.1, *LE* 16.2, *LE* 17.1, *LE* 17.2, *SS* 12.

[20] In three lines (*FR* 3, *FR* 8, *BC* 41), the diacritics imply conflicting scansions. For these, we simply included two copies of the lines in question, one for each diacritic marking.

(24) *Hopkins' diacritics*

	diacritic	*symbol*	*meaning*
a.	single and double acute accents	´ ˝	scanned in S
b.	double grave accent	̏	scanned in W
c.	large colon	:	empty W
d.	pipe	|	midline caesura: equal numbers of S positions must precede and follow
e.	superscript arc	⁀	syllables merged through paraphonology
f.	subscript arc	‿	outride

3.6 Sample scansion

We provide below a scansion (following the analysis of §§6.2–6.6 below) of the first eight lines of Hopkins' poem 'The Windhover'. We give IPA transcriptions which, when combined with the stress marks, yield the weights shown. The scansion is the one established under our analysis and is identical to Kiparsky's. Hopkins' diacritics are included.

(25) *Scansion of 'The Windhover'* 1–8

a. /$_w$ I /$_s$ caught /$_w$ this /$_s$ mór- /$_w$ ning /$_s$ mor- /$_w$ ning's /$_s$ mí- /$_w$ nion /$_s$ king-
aɪ ˈkɔːt ðɪs ˈmɔː nɪŋ ˈmɔː nɪŋz ˈmɪ njən ˈkɪŋ
– – ⏒ – ⏒ – ⏒ ⏑ ⏒ –

b. /$_w$ dom of /$_s$ day- /$_w$ light's /$_s$ dau- /$_o$ phin, /$_w$ dap-ple- /$_s$ dáwn- /$_w$ drawn
dəm əv ˈdeɪ ˌlaɪts ˈdɔː fɪn ˈdæ pəl ˈdɔːn ˌdrɔːn
⏒ ⏒ – – – ⏒ ⏑ ⏒ – –

/$_s$ Fal- /$_w$ con, in his /$_s$ ri- /$_w$ ding
ˈfɔː kən ɪn hɪz ˈɹaɪ dɪŋ
– ⏒ ⏒ ⏒ – ⏒

c. Of the /$_s$ ról- /$_o$ ling /$_w$ le-vel /$_s$ ún- /$_w$ der- /$_s$ néath /$_o$ him /$_w$ stea-dy /$_s$ áir,
əv ðə ˈɹəʊ lɪŋ ˈlɛ vəl ˌʌn də ˈniːθ hɪm ˌstɛ dɪ ˈɛː(r)
⏒ ⏑ – ⏒ ⏑ ⏒ – ⏒ – ⏒ ⏑ ⏑ –

/$_w$ and /$_s$ strí- /$_w$ ding
ənd ˈstɹaɪ dɪŋ
⏒ – ⏒

d. /$_s$ High /$_o$ there, /$_w$ how he /$_s$ rung /$_w$ u-pon the /$_s$ rein /$_w$ of a
ˈhaɪ ðɛə haʊ hɪ ˈɹʌŋ ə pɒn ðə ˈɹeɪn əv ə
– – ⏒ ⏒ – ⏑ ⏒ ⏑ – ⏒ ⏑

/$_s$ wim- /$_w$ pling /$_s$ wing
ˈwɪm plɪŋ ˈwɪŋ
– ⏒ –

e. $/_{w}$In his $/_{s}$ec- $/_{w}$sta-sy! then $/_{s}$off, : $/_{w}$Ø $/_{s}$off $/_{w}$Ø $/_{s}$forth $/_{w}$on $/_{s}$swing,
ɪn hɪz ˈɛk stə sɪ ðɛn ˈɒf ˈɒf ˈfɔəθ ɒn ˈswɪŋ
⏓ ⏓ – ⏑ ⏑ ⏓ – – – ⏓ –

f. $/_{w}$As a $/_{s}$skate's $/_{o}$heel $/_{w}$sweeps $/_{s}$smooth $/_{w}$on a $/_{s}$bow- $/_{w}$bend:
əz ə ˈskeɪts ˌhiːl ˌswiːps ˈsmuːð ən ə ˈbəʊ ˌbɛnd
⏓ ⏑ – – – – ⏓ ⏑ – –

$/_{s}$the hurl $/_{w}$and $/_{s}$glid- $/_{w}$ing
ðə ˈhɜːl ənd ˈglaɪ dɪŋ
– ⏓ – ⏓

g. Re- $/_{s}$buffed $/_{w}$the $/_{s}$big : $/_{w}$Ø $/_{s}$wind. $/_{w}$My $/_{s}$heart $/_{w}$in $/_{s}$hi- $/_{w}$ding
rɪ ˈbʌft ðə ˈbɪg ˈwɪnd maɪ ˈhɑːt ɪn ˈhaɪ dɪŋ
⏑ – ⏑ – – ⏓ – ⏓ – ⏓

h. $/_{s}$Stirred $/_{w}$for a $/_{s}$bird, – $/_{w}$the a- $/_{s}$chieve $/_{o}$of, $/_{w}$the $/_{s}$más-
ˈstɜːd fə(ɹ) ə ˈbɜːd ði ə ˈʧiːv ɒv ðə ˈmɑːs
– ⏓ ⏑ – ⏑ ⏑ – ⏓ ⏑ –

$/_{w}$te-ry of the $/_{s}$thing!
tə ɹɪ əv ðə ˈθɪŋ
⏑ ⏑ ⏓ ⏑ –

4 Digital corpus and machine scansion

Our study of sprung rhythm encompassed most of the 27 poems proposed by Kiparsky (1989: 339) as instances of this metre. We omitted 'The Wreck of the Deutschland', since it is not written in normal sprung rhythm but in a non-quantity-sensitive metre that Kiparsky (§8) calls 'semisprung' rhythm. We also excluded 'The Leaden Echo and the Golden Echo', because its lines vary in length according to an unpredictable scheme; it is essential in testing the system to know how many S positions each line has. The remaining corpus of 25 poems contains 583 lines.

We encoded the 6127 syllables of the corpus according to a scheme marking stress and phrasing (see §2.2) and syllable weight (§3.4). Phonological coding necessarily involves a certain application of intuition and native speaker judgment. We believe our codings are sensible, and invite scrutiny of them.[21]

Hopkins sometimes wrote poems that deliberately varied the requirements of sprung rhythm. Generally he specified these deviations in the manuscripts. Thus Hopkins notes that 'Brothers' is written without overreaving (Mackenzie 1991: 211) and that 'The Loss of the Eurydice' employs overreaving within stanzas but not across stanza breaks (Mackenzie 1991: 136). 'Tom's Garland' and 'Ashboughs' are specified

[21] The codings are available in the supplementary online materials at http://journals.cambridge.org/issue_Phonology/Vol28No02.

as not allowing empty W positions.[22] 'Brothers' is specified as having the special licence of allowing the line-initial 'inversion' often seen in iambic pentameter. We respected all of these specifications in our scansions.

For reliability, we scanned using a computer program. The input file to our program specifies for each line its phonological encoding (stress, phrasing, weight), as well as its intended metre (trimeter, tetrameter, etc.). The program outputs a list consisting of every scansion that obeys the inviolable metrical constraints given above. The program implements the principle of overreaving, keeping track of the syllables in the final W position of one line as it scans the next. The program code embodies a precise characterisation of every constraint we have used and is posted on the article website.

4.1 Results and interpretation

The great majority of the lines in the corpus scan straightforwardly, in ways that match Hopkins' diacritic marks and strike us as intuitively sensible. To put the claim differently, we believe that our checking backs up the claims Kiparsky originally made that the Hopkins sprung rhythm corpus obeys the constraints that he posited. However, there are also a few lines that do not scan straightforwardly under the system. The system 'fails gracefully' with these lines, in that we can bring them within the compass of the system with minor adjustments such as paraphonology. These adjustments are enumerated in our phonological codings (see the supplementary online materials). Beyond these, there appear to be two fairly clear outright exceptions to the constraints.

First of all, in one line, we must place a 'lexical stress' (stressed syllable of a polysyllabic word) in W position, violating the requirements laid out in (4) above.

(26) $/_{\mathrm{W}}$Wag $/_{\mathrm{S}}$or $/_{\mathrm{W}}$cross- $/_{\mathrm{S}}$bri- $/_{\mathrm{W}}$dle, in a $/_{\mathrm{S}}$wind $/_{\mathrm{W}}$<u>lift-</u> $/_{\mathrm{S}}$éd $/_{\mathrm{W}}$wind- $/_{\mathrm{S}}$laced – *HP* 14

Hopkins himself intended the deviation here, as we know from the diacritic he placed that specifies S position for the second syllable of *lifted*.[23]

[22] Wimsatt (2006: 52) claims on this basis that 'Ashboughs' is not in sprung rhythm at all. This strikes us as terminological quibbling, since this poem (similarly 'Tom's Garland') shows almost all the essential traits of this metre: resolved sequences, multiply filled but weight-restricted W positions, no stressless light syllables in S – it could not possibly be scanned as ordinary verse, as Wimsatt suggests. Hopkins himself described the metre more accurately (Mackenzie 1991: 318): 'common rhythm, but with hurried feet'. (Wimsatt misleadingly omits the last four words when quoting this specification.)

[23] A reviewer notes that sometimes *X-xx* compounds are treated by poets as if they were simple trisyllabic words with initial stress. This would solve the problem with (26), but seems unlikely to us; Kiparsky (1977: 221) notes that such cases typically involve high-frequency, lexically listed compounds, which is hardly the case with *wind-lifted*.

Kiparsky mentions two other possible such cases (1989: 326), also supported by Hopkins' diacritics, but according to our calculations these two scansions are not actually required under the system of constraints.

There is one other line that is very difficult to scan under the system.

(27) Forward-like, but however, and like favourable heaven heard these
BC 48

Our best response to this line is to assume the phonological representation [ˈfɔːwəd] for *forward* (cf. note 15), which renders the vowel eligible for Correption (§3.4.1); this is the only case of Correption before a consonant, but perhaps the homorganic character of [ɔː] and [w] renders this plausible. We also apply Correption to the [aʊ] of *however*, take Hopkins' outride mark on *and* at face value and obtain (28).

(28) $/_{w}$ Forward- $/_{s}$ like, $/_{w}$ but how- $/_{s}$ ev- $/_{o}$ er, and $/_{w}$ like $/_{s}$ fa- $/_{w}$ vourable $/_{s}$ heaven $/_{w}$ heard $/_{s}$ these

We feel that this is going out on too many limbs at once, and would prefer to regard (28) as an exception. We note that the second token of the syllable *like* is a last-minute interpolation by Hopkins; it seems possible that he did not fully check the consequences for scansion when he added it.

4.2 Assessment using a prose sample

There are only 583 lines in the corpus; hence we ask how meaningful having just two exceptions is. A way to address this question is the prose model method (Tarlinskaja & Teterina 1974, Tarlinskaja 1976, Biggs 1996): one locates lines of prose similar in their phonological properties to the verse in question and tries to scan them as if they were verse. This establishes a baseline that tells us to what extent the regularities observed are merely the result of the poet using the language at hand, with its characteristic word lengths, stress patterns and phrasings.

For our primary prose sample, we took Hopkins' 'Author's Preface' (Mackenzie 1990: 115–117) and a few of his letters, chosen for being available in digital form. We separated these texts into 'pseudo-lines' of sequences separated by punctuation marks. From this material, we randomly selected 155 pseudo-lines in such a way that the distribution of line lengths, counted by syllables, was the same as that found in the corpus of Hopkins' actual verse. For each such pseudo-line, we assigned a specific number of S positions to which it had to be scanned. These assignments were made randomly, but matched the corpus of real lines statistically. To give one example, the 15-syllable pseudo-line in (29) was assigned at random to be scanned as a sprung rhythm pentameter (six of the 21 15-syllable lines in the corpus are pentameters).

(29) The poems in this book are written some in Running Rhythm,
(*Author's Preface*)

Since Hopkins' real lines must be scanned in accordance with the principle of overreaving (§3.2), we broke up our sample into seven-line 'poems', seven lines being the average size of domain (poem or stanza; §4) within which overreaving is enforced in the real corpus. Lastly, we annotated our corpus of pseudo-lines for stress, weight and phrasing, just as we did the original corpus.

We scanned the lines with our program, and for those lines in which no legal scansion was found, we pondered various 'repairs' of the type used earlier to rationalise difficult lines from the corpus. However, even when we carried out these repairs, 19 of the 155 lines in our prose sample (12·3%) remained unscannable. In most cases the lines were simply too long to scan. We give an example in (30); even when we cram as many syllables into each position as possible, we still have leftovers (the syllables 'Falling Rhythms') at the end.

(30) *An unscannable line of prose* (*pentameter assumed*)

$/_{w}$ In which the $/_{s}$ stress $/_{w}$ comes $/_{s}$ first $/_{w}$ are $/_{s}$ called $/_{w}$ Ø $/_{s}$ Fal- $/_{w}$ ling
ɪn wɪʧ ðə stɹɛs kʌmz fɜːst ɑː/ə kɔːld fɔː lɪŋ
⏓ ⏓ ⏑ – – – ⏓ – – ⏓

$/_{s}$ Feet $/_{w}$ and / … *Fal-ling Rhy-thms
fiːt ənd fɔː lɪŋ ɹɪ ðəmz
– ⏓ – ⏓ ⏑ ⏓

Most unscannable lines were too long like this; the line 'Amiability' (scanned as a tetrameter) was too short.

We used the prose sample to assess the possibility that the relative paucity of unmetrical lines in the Hopkins corpus could have arisen by accident. The relevant proportions here are 19/155 unmetrical lines in the prose sample, 2/583 for the real lines. A Fisher's exact test indicates that this difference is highly significant; $p < 10^{-11}$. Our test affirms that the Kiparskyan analysis reflects genuine regularities in the sprung rhythm corpus.

4.3 Compatibility with Hopkins' diacritics

As Kiparsky noted, an adequate account should not just make it possible to scan every line; but should provide a scansion that is compatible with Hopkins' diacritics. Thus we further check whether the grammar is producing such a scansion.

Since by Hopkins' own account, he did not diacritically annotate everything,[24] we adopt a 'positive evidence only' criterion: if there is a diacritic present in the source, an adequate scansion must match it; but less common scanning options not matched by a diacritic are assumed to be possible. Checking by this criterion, we find that there are only three lines that cannot be scanned in a diacritic-compatible way. These are given in (31), which includes the best scansion obtainable in the model developed below (§§6.2–6.6).

(31) *Lines not scannable in agreement with the Hopkins diacritics*

a. $/_{\mathrm{w}}$ Ø $/_{\mathrm{s}}$ Poe- $/_{\mathrm{o}}$ try tó it, $/_{\mathrm{w}}$ as a $/_{\mathrm{s}}$ trée $/_{\mathrm{w}}$ whose $/_{\mathrm{s}}$ bóughs $/_{\mathrm{w}}$ brȅak
$/_{\mathrm{s}}$ ín $/_{\mathrm{w}}$ the $/_{\mathrm{s}}$ ský. *AB* 3

b. $/_{\mathrm{w}}$ But $/_{\mathrm{s}}$ quench $/_{\mathrm{w}}$ her $/_{\mathrm{s}}$ bonni- $/_{\mathrm{w}}$ est, $/_{\mathrm{s}}$ dear- $/_{\mathrm{o}}$ est | $/_{\mathrm{w}}$ to her, her
$/_{\mathrm{s}}$ clear- $/_{\mathrm{w}}$ est $/_{\mathrm{s}}$ sel- $/_{\mathrm{w}}$ vèd $/_{\mathrm{s}}$ spark *HF* 10

c. $/_{\mathrm{w}}$ Ø $/_{\mathrm{s}}$ Ever $/_{\mathrm{w}}$ so $/_{\mathrm{s}}$ black $/_{\mathrm{o}}$ on it. $/_{\mathrm{w}}$ Oúr $/_{\mathrm{s}}$ tale $/_{\mathrm{w}}$ O oúr $/_{\mathrm{s}}$ ora- $/_{\mathrm{w}}$ cle! |
$/_{\mathrm{s}}$ Lét $/_{\mathrm{w}}$ life, $/_{\mathrm{s}}$ wáned, $/_{\mathrm{w}}$ ah $/_{\mathrm{s}}$ lét $/_{\mathrm{w}}$ life $/_{\mathrm{s}}$ wínd *SS* 10

Kiparsky (1989: 314) suggests that (31c) involves a marginal possibility of trisyllabic resolution (on *oracle*), an option later abandoned by Hanson & Kiparsky (1996) and here; we treat the line as scannable but not matchable to the diacritics.

If we add these three lines as exceptions to the two already discussed, the exception rate (5/583) remains less than 1%, and the statistical comparison with the prose sample remains highly significant (Fisher's exact test, $p < 10^{-7}$).

4.4 Two new inviolable constraints for sprung rhythm

There appear to be two additional inviolable constraints detectable in the data.

First, we find that while Hopkins is rather flexible in mismatching stressed and stressless syllables to S and W position (see Kiparsky 1989: §10), there is one place where he is quite strict: the rightmost S position of a line must always have a greater degree of stress than the following W. The only possible exceptions can be defined on grounds of rhyme; see note 29. This condition is not surprising in the context of English metrics, where the stress conditions are often enforced more strictly for the last S position in the line (Tarlinskaja 1976: 143–144).

Second, we find that the few appearances of empty W positions in word-medial position are limited by morphology. The only words that have empty W within them (position shown) are: *un Ø selve* (*BP* 21), *un Ø cumbered* (*CS* 13), *Un Ø Christ* (*LE* 24.4) and *dis Ø membering* (*SS* 7). Plainly, this is not a random collection; in all of them the empty W is

[24] He said (in the manuscript of 'Harry Ploughman'; Mackenzie 1991) that he employed the acute accent 'in doubtful cases only'.

placed after a productive prefix (in terms of Lexical Phonology (Kiparsky 1982), a prefix of Level II). Thus, quite a bit of analytical work can be done by a constraint that forbids empty W within simplex words.

With these two observations, we add to the constraints given so far the following.

(32) a. Final Match
Assign a violation when the rightmost S is filled by a syllable that does not have more stress than the syllable that fills the following W.

b. Word Integrity
Assign a violation to every simplex word defined at Level I which is interrupted by an empty W.

We will call the grammar which uses the Kiparskyan inviolable constraints plus these two the 'amplified Kiparskyan grammar'.

5 Number of scansions

The amplified Kiparskyan grammar allows multiple legal scansions for most lines of sprung rhythm. In this and the following section we address the issue of multiple scansions, arriving at the view that while in some cases multiple scansions are appropriate, overall the system provides too many scansions, violating a Parsability Principle that we propose as a metrical universal.

5.1 Free variation in Hopkins' practice

That more than one scansion should be regarded as possible for certain lines is indicated by Hopkins' own metrical practice. His diacritics sometimes indicate different scansions for very similar sequences in different lines, the choice of which is unpredictable.

A clear case is his treatment of sequences of two stressless syllables in a row, separated by a phrase break. Hopkins sometimes indicates by a diacritic that the first stressless syllable should be scanned as an outride, as in the examples of (33a). In other cases, however, he provides no outride diacritic, suggesting a multiply filled weak position as in (33b).

(33) *Treatment of disyllabic stressless sequences with phrase break*

a. *As outride*

$/_{w}$Not, $/_{s}$I'll $/_{w}$not, $/_{s}$car- $/_{w}$rion $/_{s}$com- $/_{o}$fort $/_{w}$De- $/_{s}$spair,
    $/_{w}$not $/_{s}$feast $/_{w}$on $/_{s}$thee; *CC* 1

$/_{w}$Ø $/_{s}$: Fe- $/_{w}$lix $/_{s}$Rán- $/_{o}$dal $/_{w}$the $/_{s}$fár- $/_{o}$rier, $/_{w}$O is he $/_{s}$déad $/_{o}$then?
    My $/_{s}$dú- $/_{w}$ty all $/_{w}$én- $/_{s}$ded, *FR* 1

$/_{s}$Mónths $/_{w}$Ø $/_{s}$eár- $/_{o}$lier, $/_{w}$since $/_{s}$Í had $/_{w}$our $/_{s}$swéet $/_{w}$re- $/_{s}$príeve
    $/_{w}$and $/_{s}$rán- $/_{w}$som *FR* 7

b. *As doubly filled W position*

$/_w$ Mells $/_s$ blue $/_w$ and $/_s$ snow- $/_w$ white $/_s$ through $/_w$ them, a $/_s$ fringe
$/_w$ and $/_s$ fray *AB* 9

$/_w$ There) – $/_s$ boy : $/_w$ Ø $/_s$ bu- $/_w$ gler, $/_s$ born, $/_w$ he $/_s$ tells $/_w$ me, of
$/_s$ I- $/_w$ rish *BC* 2

$/_w$ Who $/_s$ fíred $/_w$ Ø $/_s$ Fránce $/_w$ for $/_s$ Má- $/_w$ ry with- $/_s$ óut : $/_w$ Ø $/_s$ spót.
DO 14

We can discover no basis for the choice between these two possibilities, and conclude that this is a case of free variation – that both possibilities are legal sprung rhythm scansions. For further possible instances of free variation see tableau (46) below, as well as the full set of tableaux posted on the article website.

5.2 Variety of scansions

We turn next to the issue of the *number* of possible scansions – are there too many? To begin with the facts, our software indicates that the average number of scansions permitted by the amplified Kiparskyan grammar is 14·8 (minimum 1, maximum 932, median 6). The number of possible scansions increases with line length, since there are more loci of possible variation.

Crucially, it appears that none of these possible scansions can be ruled out by adding new inviolable constraints. Furthermore, the possible scansions vary widely in terms of markedness, and the Kiparskyan grammar offers no way of choosing a less marked (or totally unmarked) scansion over a highly marked but still possible one.

Here is an example. The counterintuitive scansion of *BC* 21 in (34a), which is legal under the amplified Kiparskyan analysis, is on equal footing with the much more natural scansion in (34b).

(34) a. $/_w$ Forth $/_s$ Christ $/_w$ from $/_s$ cup- $/_o$ board $/_w$ fetched, $/_s$ how $/_w$ Ø $/_s$ fain
$/_w$ I of $/_s$ feet

b. $/_w$ Forth $/_s$ Christ $/_w$ from $/_s$ cup- $/_w$ board $/_s$ fetched, $/_w$ how $/_s$ fain
$/_w$ I of $/_s$ feet

However, the scansion in (34a) cannot be eliminated by adding additional inviolable constraints. Its most obvious metrical defects are (a) the syllable *fetched*, bearing phrasal stress, in W position, and (b) the empty W, which, unusually for Hopkins, is not flanked by two stressed syllables (see §6.3 below). Any constraints that forbade these configurations, however, must be violated in other lines, which would otherwise not be scannable at all. For example, phrasal stresses appear in weak position in (35a), and an empty W not flanked by stresses appears in (35b).

(35) a. $/_{w}$ In $/_{s}$ wide $/_{w}$ the $/_{s}$ world's $/_{w}$ <u>weal</u>; $/_{s}$ ra̋re $/_{w}$ <u>gòld</u>, $/_{s}$ bőld $/_{w}$ <u>stèel</u>, $/_{s}$ ba̋re *TG* 17

b. $/_{s}$ Léaves, $/_{w}$ Ø $/_{s}$ <u>líke</u> $/_{w}$ the $/_{s}$ thíngs $/_{w}$ of $/_{s}$ mán, $/_{w}$ you *SF* 3

These scansions are not only mandated by Hopkins' diacritics, but are the only scansions allowed under the amplified Kiparskyan grammar.

In sum, it appears that any effort to use inviolable constraints to trim back the large number of scansions, or to remove highly marked scansions allowed under the amplified Kiparskyan grammar, would run afoul of exceptions.

6 The Parsability Principle and its consequences

We have just seen that the large number of scansions allowed under the amplified Kiparskyan grammar is problematic, in that some of the scansions seem intuitively unnatural. But the simple fact that the scansions are numerous may also be problematic. Here we explore the possibility that a genuine 'metrical pathology' is present in the Kiparskyan analysis. Specifically, we propose that metrical systems are subject to a naturalness requirement related to parsability; i.e. the ability of the reader or listener to recover from the phonological string a reasonably unambiguous sense of the scansion. We will call this the PARSABILITY PRINCIPLE.

(36) *Parsability Principle*

In a metrical system, the rules must be such that the reader or listener can in the great majority of cases recover the intended alignment of the verse with the metre.

The Parsability Principle is defined as a strong tendency, rather than an absolute requirement, because metrical systems may involve *modest* deviations from full parsability. We give one example here. Milton's mature blank verse involves a system of paraphonology (Bridges 1921, Sprott 1953, Kiparsky 1977) whereby vowels in hiatus undergo optional glide formation or elision. When the paraphonology is applied in appropriate ways, the resulting representations obey a very strict requirement for syllable count (ten positions plus an optional extrametrical). Occasionally – and only occasionally – Milton writes a line in which there are two locations where a vowel could be removed by paraphonology, and both scansions are reasonable. This is true, for instance, of *Paradise Lost* 10.974 in (37).

(37) *Two scansions of* Paradise Lost *10.974*

a. *Paraphonology applied to* dying
/ˈdaɪɪŋ/ → [ˈdaɪŋ] *by postvocalic stressless vowel drop*
(Kiparsky 1977: 240; 'PR 1')
$/_{w}$Liv- $/_{s}$ing $/_{w}$or $/_{s}$dying $/_{w}$from $/_{s}$thee $/_{w}$I $/_{s}$will $/_{w}$not $/_{s}$hide

b. *Paraphonology applied to* thee I
/ði aɪ/ → [ðjaɪ] *by prevocalic glide formation*
(Kiparsky 1977: 240; 'PR 3')
$/_{w}$Liv- $/_{s}$ing $/_{w}$or $/_{s}$dy- $/_{w}$ing $/_{s}$from $/_{w}$thee I $/_{s}$will $/_{w}$not $/_{s}$hide

Such cases do exist, but they are rare, and even among them, usually one scansion is more far more probable (in terms of Milton's characteristic metrical practice) than the other. In short, for readers with sufficient experience with Milton's metre, scansion is highly but not perfectly recoverable from the printed page.

The same holds, we believe, of metre in general: scansion is generally, if sometimes imperfectly, recoverable from text. Across traditions, this pattern arises from several factors: strict syllable-to-position correspondence (as in most European literary verse or in Chinese metres), strict quantity limitations, or the requirement that paraphonological rules apply as the norm rather than the exception, as in Italian (Elwert 1984).

Intuitively, there is good reason for the Parsability Principle to hold true: the reader takes pleasure in 'hearing the rhythm' of a verse line, and a line that is compatible with many different rhythmic interpretations will necessarily fail to produce a vivid impression for any particular one of them.[25]

If the Parsability Principle is true, then one of two things must hold of sprung rhythm. Either it was a misconceived project – essentially, unnatural poetry – or the Kiparskyan analysis is only a partial account of it. In fact, we think it unlikely that sprung rhythm is an unnatural metre, since so many readers experience strong intuitive appreciation of Hopkins' metrical practice. We posit instead that the Kiparskyan analysis

[25] An anonymous reviewer, addressing the question of parsability, asks whether this implies 'that metrical theory provides a discovery procedure … such that the theory in combination with the line should generate one of a small number of scansions'. If the rest of linguistics is taken as a guide, the answer would be no. It is a near-consensus position for syntax that a grammar specifies what can be a legal sentence, and that people in addition have the ability to parse sentences fluently, making use of their grammars as well as much other information. Our Parsability Principle is a pragmatic requirement on metrical grammars, which assumes the existence of an effective parser; but it does not incorporate the parser itself into metrical theory. For more on parsing, see §6.6.4 below.

is incomplete, and seek additional constraints that make the system more reliably parsable.

6.1 Amplifying the analysis: the role of stochastic grammar

Our goal is to discover a system that can produce multiple possible scansions for a line of sprung rhythm, but that (a) distinguishes marked from unmarked scansions in some way and (b) provides the reader/listener with a means of recovering the intended scansion with some reasonable probability. We will argue here that what is needed is a stochastic grammar that includes violable constraints.

The need for such grammars is adumbrated in much of the literature in metrics. For instance, the core constraints proposed for English iambic pentameter (notably in Kiparsky 1977) are exceptionful, as pointed out by, for example, Barnes & Esau (1978), Koelb (1979) and Tarlinskaja (2006). Youmans (1989) argues that exceptionality and gradience are pervasive in metrics, supporting his claim by showing how subtle, violable constraints lead Milton to use non-standard word order in his verse.

While the existence of violable constraints in metrics is widely acknowledged, we are also in need of an explicit formal theory that can accommodate such constraints. Recent theoretical developments in linguistics offer a number of candidate frameworks. What they all share is that output candidates (for us, scansions) are not classified on an up-or-down basis, but rather are assigned a probability, expressing gradient metricality. A probability-based approach has the potential to solve all of the problems we have encountered so far, as follows.

First, where there is evidence (as in (33)) that more than one scansion should be considered well-formed, then each such scansion receives an intermediate probability (that is, well above zero and well below one). For example, the stochastic grammar we will propose assigns to the three lines of (33a) the probabilities (in order) 0·589, 0·495 and 0·548, and to (33b) 0·383, 0·102 and 0·420. As we will see later on, such intermediate probabilities result from constraint conflict.

Second, a stochastic grammar can include non-stochastic behaviour as a special case. In the stochastic grammar we will propose, the constraints of the amplified Kiparskyan grammar are assumed to be inviolable. Lines violating them receive zero probability, and are thus treated as fully unmetrical, just as before.

Third, while a stochastic grammar cannot *entirely* rule out problematic bad scansions such as (34a), it can come close, by assigning them very low probabilities. As will emerge, such scansions violate constraints that, though not inviolable, are still rather strong. In the grammar we propose, (34a) is assigned the very low probability of 0·0065; its much better partner (34b) 0·973.

Fourth, a stochastic grammar can nevertheless assign high probabilities to scansions that have noticeable constraint violations, such as those in (35). These scansions win for want of a better alternative, since they have no competitor scansions that obey the relevant constraints. In the grammar we propose, both (35a) and (35b) are the unique candidates that obey the inviolable constraints, and thus they receive a probability of one.

Lastly, a stochastic grammar can solve the parsability problem: for the grammar we propose, it is generally the case that just one scansion (or at worst, just a few) dominates the probability distribution. If the reader simply guesses the highest-probability scansion, she will in the great majority of cases arrive at a single, clear metrical interpretation.

In what follows, we flesh out this idea. §6.2 describes the maxent framework for metrical grammars. §6.3 gives our proposed additional, violable constraints. §6.4 gives the data we use to set the weights of the maxent grammar, and §6.5 describes the computations by which the final grammar was arrived at. The remaining sections assess our analysis and offer general conclusions.

6.2 Constructing the stochastic grammar I: framework

As noted above, several formal models exist that make it possible to mould raw constraint sets into grammars that assign probabilities to candidates. In the main text we describe the results we have obtained using MAXENT GRAMMARS, an approach whose intellectual ancestry dates to Hopkins' time but whose application to linguistics is recent (see e.g. Goldwater & Johnson 2003, Smolensky & Legendre 2006, Wilson 2006, Hayes & Wilson 2008). Other frameworks are discussed in Appendix B.

Maxent grammars use constraints to assign probabilities to a set of candidates. The candidates may be assumed to be created for each input by a GEN function, as in Optimality Theory (Prince & Smolensky 1993). Each candidate is assessed for the number of times it violates each constraint. The candidates compete with one another for a share in the probability (total: 1) for any given input. Here, the inputs are pairs consisting of the phonological representation of a line and a sprung rhythm metre (trimeter, tetrameter, etc., according to the particular poem). The candidate outputs created by GEN are scansions; i.e. alignments of the syllables with the metre as in (2).

Each constraint bears a weight, a non-negative real number. Intuitively, the higher the weight, the stronger the effect of the constraint; all else being equal, candidates that violate highly-weighted constraints will have lower probabilities. The procedure that derives probabilities from weights and violations (see Goldwater & Johnson 2003) is given in (38).

(38) *Maxent probability computation*

Given a metre M, a phonological representation for a line L and a set of n possible scansions:

For each scansion:

Assess its constraint violations.

Multiply constraint violations by respective weights, and sum.

Reverse the sign and take e to the result.

Sum, over all n scansions, the result obtained in previous line.

Compute, for each scansion, its share of this sum. This is the PROBABILITY of this scansion as assigned by the grammar.[26]

Where do our constraint weights come from? For the inviolable constraints of the amplified Kiparskyan grammar, we simply assume infinite weights, and thus that scansions violating them receive zero probability. For the violable constraints, we need smaller weights, such that violations do not completely rule out a scansion, but have the effect of strongly disfavouring inferior scansions – for example, in preferring (34b) to (34a). The weights also need to be balanced against one another, so that rival scansions that violate different constraints (as in the cases in (33)) will each be able to receive substantial probability.

In principle, we could set the weights by hand, finding values that prefer the scansions that we intuitively prefer. In fact, a better method is available, one that accesses Hopkins' own intuitions. We cover this method in §6.4 below, then in §6.5 discuss how we computed the weights.

6.3 Constructing the stochastic grammar II: the violable constraints

In developing our augmented grammar, we tried to adhere to existing research literature in metrics. This is because we value a solution to sprung rhythm that situates it typologically within English metrics and within metrics generally. Thus, we sought constraints that would be applicable in other metrical genres and traditions as well as in sprung rhythm.

We explored a large number of constraints (30), but in the end we added only eight new violable constraints to those of the amplified Kiparskyan grammar. Our criterion for selecting them was that they all passed a

[26] In mathematical notation this is:

(i) $p(\omega) = \frac{1}{Z} e^{-\Sigma_i \lambda_i \chi_i(\omega)}$, where $Z = \sum_j e^{-\Sigma_i \lambda_i \chi_i(\omega_j)}$

(Della Pietra *et al.* 1997: 1). $p(\omega)$ denotes the predicted probability of scansion ω; e is the base of natural logarithms; Σ_i denotes summation across all constraints, λ_i denotes the weight of the ith constraint, $\chi_i(\omega)$ denotes the number of times ω violates the ith constraint; and Σ_j denotes summation across all possible scansions.

standard significance test (the likelihood ratio test).[27] Our proposed grammar is the largest we could find in which every constraint tested as significant.

The constraints we added to the grammar fall into four families. First, there seems to be a dispreference in Hopkins' verse for multiply filled W positions (Kiparsky 1989: 314). We suggest that any polysyllabic W position involves some cost, and that a triply filled W position is costlier still. (In principle, a quadruply filled W should be even more costly, but since the model fit does not improve with a constraint banning this we will not include it here.) Moreover, resolution seems to be disfavoured in S position, in comparison to W (Kiparsky 1989: 332). The constraints we suggest are given in (39).

(39) *Violable constraints on multiply filled metrical positions*

a. *2σ IN W
   Assign a violation for W positions filled by two or more syllables.

b. *3σ IN W
   Assign a violation for W positions filled by three or more syllables.

c. *RESOLUTION IN S
   Assign a violation for each resolved sequence in S position.

For examples of violations, see (3), (5) and (13). *2σ IN W is meant as 'at least two'; hence if (39b) is violated, so is (39a).

The typological status of the constraints in (39) should be uncontroversial: a great many metrical traditions require W positions to be filled just once, and a great many English poets never use resolution in their standard-metre verse.

A second source of violable constraints is the principle (common to all stress-based metres) that scansions should match the stress contour of the line to the SW pattern of the metre (Jespersen 1933, Hayes 1983). Our grammar uses the constraint in (40).

[27] The test is described in Pinheiro & Bates (2000: 83). Using the target grammar, one calculates the predicted probability of the entire set of Hopkins-preferred scansions (see §6.4), by multiplying out the probabilities assigned to each individual line. Then, one removes from the target grammar the constraint one wants to test and again computes the combined probability. The formula in (i) can be approximated by a chi-square distribution with one degree of freedom, and one can then look up from this distribution the probability of the null hypothesis that the simpler grammar is adequate.

(i) $$2 \times \ln\left(\frac{\textit{Probability of corpus under simpler grammar}}{\textit{Probability of corpus under full grammar}}\right)$$

The least significant p-value obtained was $p = 0{\cdot}009$ for constraint (39c). We rejected constraints for which p was greater than $0{\cdot}05$.

(40) MATCH SW

Assign a violation if the (first) syllable occupying W position has more stress than the (first) syllable occupying the preceding S.

A sample violation is the sequence $/_s$ *and* $/_w$ *God's* / in (14). It has long been noticed that the sequence SW tends to be matched more strictly (that is, with a 'fall' in stress contour) than WS (Magnuson & Ryder 1970, Kiparsky 1977), and this appears to be so for sprung rhythm; the analogous constraint MATCH WS fails to pass our significance test. MATCH SW is one of the constraints that results in low probability for scansion (34a) above.

Third, we find that empty W positions strongly tend to be flanked by S positions that are filled with stressed syllables, a distribution pointed out by Hopkins himself (Abbott 1935a: 23). This restriction perhaps relates to the Parsability Principle, as it may help the listener identify these positions as being empty. In English folk verse forms that also allow empty W positions, such a requirement appears to be normally obeyed (cf. Attridge 1982: 97), as in the lines of children's verse in (41).

(41) *Empty W in folk verse flanked by stress*

a. $/_s$ $\underline{\text{Tom}}$, $/_w$ Ø $/_s$ $\underline{\text{Tom}}$, $/_w$ the $/_s$ pi- $/_w$ per's $/_s$ son (Opie & Opie 1951: 411)
b. $/_s$ $\underline{\text{Hinx}}$, $/_w$ Ø $/_s$ $\underline{\text{minx}}$, $/_w$ the $/_s$ old $/_w$ witch $/_s$ winks (1951: 208)
c. $/_s$ Mat- $/_w$ thew $/_s$ $\underline{\text{Mark}}$, $/_w$ Ø $/_s$ $\underline{\text{Luke}}$, $/_w$ and $/_s$ John (1951: 303)

Thus we posit the constraint in (42).

(42) *NO CLASH EMPTY W

Assign a violation for an empty W position if the S positions that flank it are not both filled by stressed syllables.

(43) gives a sample violation.

(43) *Sample violation*

$/_w$ Ø $/_s$ Már- $/_w$ ga- $/_s$ $\underline{\text{rét}}$ $/_w$ Ø $/_s$ $\underline{\text{áre}}$ $/_w$ you $/_s$ gríe- $/_w$ ving *SF* 1

The term 'clash' in the name of the constraint refers to the stress clash created by the two adjacent stressed syllables in S. *NO CLASH EMPTY W is the other constraint that contributes to the low probability for scansion (34a) above.

The remaining constraints we posit govern outrides. Outrides are themselves assessed as metrically costly by the constraint *OUTRIDE (44a). As Kiparsky notes (1989: 324), outrides that precede only a weak phonological break are disfavoured; we express this with (44b), which penalises outrides that are only at the end of a Clitic Group, not at the end of a Phonological Phrase or higher category. Lastly, departing slightly from Kiparsky, we have marginally permitted outrides in short lines (tetrameter and shorter), but we penalise this with what turned out to be a very strong constraint, (44c).

(44) *Constraints on outrides*

a. *OUTRIDE
Assign a violation for every outride.

b. *OUTRIDE-WEAK BREAK
Assign a violation for every outride that is only at the end of a Clitic Group.

c. *OUTRIDE-SHORT LINE
Assign a violation for every outride in a line with four or fewer S positions.

Outrides in general are abundant (though rare enough that a constraint against them improves the model fit). Examples of less usual lines violating (44b, c) are given below.

(45) *Examples of unusual outrides*

a. *Violating* (44b)
$/_{\mathrm{w}}$And the $/_{\mathrm{s}}$az- $/_{\mathrm{o}}$<u>urous</u> $/_{\mathrm{w}}$hung $/_{\mathrm{s}}$hills $/_{\mathrm{w}}$are his $/_{\mathrm{s}}$world= $/_{\mathrm{w}}$Ø $/_{\mathrm{s}}$: wiel-
$/_{\mathrm{w}}$ding $/_{\mathrm{s}}$shoul- $/_{\mathrm{w}}$der *HH* 9

b. *Violating* (44c)
$/_{\mathrm{w}}$A- $/_{\mathrm{s}}$cross $/_{\mathrm{w}}$my $/_{\mathrm{s}}$foun- $/_{\mathrm{w}}$der-ing $/_{\mathrm{s}}$deck $/_{\mathrm{o}}$<u>shone</u>[28] *HF* 18

Constraints (44a) and (44b) arguably have a solid typological status. Outrides (also known as extrametrical syllables) have a marked status in standard English metres; for example, they are often used sparingly, or limited to line endings, or limited to fully stressless syllables, and so on. The restriction of outrides to just before strong prosodic breaks (44b) is also standard (Steele 1999: 87). The limitation of outrides to the longer metres (44c) is harder to defend typologically; Kiparsky (1989: 336–337) suggests an artistic basis for this choice.

In sum, we propose to add to the amplified Kiparskyan grammar a set of eight violable constraints, given in (39), (40), (42) and (44). These constraints all receive finite weights in the maxent grammar, reflecting their violability.

[28] *HF* 18 must be scanned with *shone* as an outride in order to avoid an overreaving violation in the immediately following line. The other line violating (44c) is *LE* 20.1.

6.4 Setting the weights

The violable constraints of a maxent grammar are assigned weights, expressing their potency in reducing the probability of scansions that violate them. Informal experimentation indicates that, provided we give these constraints weights that are neither very large (which would render them essentially inviolable) nor very small (which would render them ineffective), we can make substantial headway on the problems laid out in §6.1 (permitting free variation, allocating low probabilities to implausible scansions and so on). Yet setting the weights by hand relies on the metrical intuitions of the analyst – perilous in the case of sprung rhythm – and strikes us as somewhat unprincipled in any event.

We can do better than this, because Hopkins has left us an invaluable source of his own metrical intuitions, in the form of the diacritic markings he attached to much of his sprung rhythm verse (§3.5). Over half the time, the diacritic markings are compatible with just one of the many scansions permitted by the amplified Kiparskyan grammar. These diacritically marked lines thus constitute unambiguous testimony for Hopkins' own metrical preferences.

Our goal, then, is to set the weights in a way that maximises the probability assigned to scansions which we have reason to believe from the diacritics that Hopkins himself preferred. By assigning the weights in this way, we can maximise the faithfulness of the analysis to Hopkins' intuitions, rather than relying on our own.

In addition to the diacritics, there are two other sources of information that can be used to identify Hopkins' preferred scansions. First, our inspection of the corpus leads us to believe that, without exception, Hopkins places the rhyming syllable (or initiates a sequence of rhyming syllables) in the rightmost S position of the line. We have relied on this regularity and reject as Hopkins-incompatible any scansions that would violate it.[29] In addition, we have assumed that any outride intended by Hopkins is marked as such.

In sum, our criteria for diagnosing Hopkins' own preferences are his diacritics, the assumption that rhymes are in S position and the assumption that all outrides are marked. From these, we established with our software that there are 311 lines in which the diacritics unambiguously indicate a 'Hopkins-preferred' scansion by our criteria – somewhat more than half the corpus.

The 311 Hopkins-preferred scansions demonstrate the violability of the constraints proposed in §6.4. Table I gives the violation counts for each constraint in this set of lines.

[29] A few rhymes, like *FR* 3: *Pining, pining, till time when reason rambled in it and some*, require stressings that go counter to the English norm (here [ˈændsəm], with an oddly stressed proclitic in order to rhyme with *handsome*) if they are to sound like rhymes at all. We assume this is what Hopkins meant (cf. his letters; Abbott 1935a: 180); it also rationalises the scansions of a number of lines.

constraint	violations
*2σ IN W	211
*3σ IN W	23
*RESOLUTION IN S	21
MATCH SW	117
*NO CLASH EMPTY W	21
*OUTRIDE	88
*OUTRIDE-WEAK BREAK	3
*OUTRIDE-SHORT LINE	2

Table I

Violations of the violable constraints in Hopkins-preferred scansions.

6.5 Grammar computation

Assembling all of the above ingredients, we carried out the following computation. First, we assumed that the inviolable constraints of the amplified Kiparskyan grammar have infinite weights, thus excluding any candidates that violate them. We implemented this with our scansion software, which simply refrained from outputting any scansions that violate these constraints.[30] The software did output all scansions that obey these constraints (a total of 8633, distributed among 583 lines), along with the number of times each scansion violated our eight violable constraints (hence 8×8633 values). The software also singled out the 311 lines that have a unique 'Hopkins winner' and designated which scansion for each of the lines was the winner.

The next step was to set the weights, with the goal of assigning as much probability as possible to the Hopkins-preferred scansions. This is an instance of the widely-used 'maximum likelihood criterion' for model fitting.[31] There are several algorithms that can find the maximum-likelihood weights for a maxent grammar (Malouf 2002). We used the conjugate gradient algorithm, implemented in a computer program called the Maxent Grammar Tool (Wilson & George 2009). We fed this program an input file containing the information above, and used it to calculate the weights. Once the weights were calculated, we used them to assign a probability to every scansion, following the procedure in (38). These

30 The full set of candidates (every logically possible assignment of syllables to positions) is often very large, making our procedure or some equivalent a practical necessity. For instance, for line (46) below there are about 350,000 candidates. For the longest lines in the corpus the count exceeds 10^{11}.

31 The danger inherent in the maximum-likelihood criterion is overfitting; that is, excessive customisation of the grammar to the data set (Duda *et al.* 2001). For evidence that our grammar is not overfitted, see Appendix A: §2.

probabilities were calculated not just for the 311 lines having an unambiguous Hopkins scansion, but for the entire set of 583 lines.[32]

In the following section, we assess the degree of success this procedure achieved, both in matching Hopkins' preferred scansions and in rendering the system parsable.

6.6 Results

6.6.1 *Weights.* The weights for the violable constraints as calculated are given in Table II.

constraint	violations	weight
*2σ IN W	211	1·75
*3σ IN W	23	1·75
*RESOLUTION IN S	21	1·44
MATCH SW	117	1·05
*NO CLASH EMPTY W	21	1·74
*OUTRIDE	88	2·22
*OUTRIDE-WEAK BREAK	3	1·69
*OUTRIDE-SHORT LINE	2	11·47

Table II
Weights of the violable constraints.

The nearly inviolable constraint *OUTRIDE-SHORT LINE received an extremely high weight, the others more modest ones. Note that the constraints were cumulative in their effects. For example, every line that had an outride before a weak break violated not just *OUTRIDE-WEAK BREAK, but also *OUTRIDE. Hence by the procedure in (38) any such line received a penalty of at least 1·69 + 2·22, or 3·91. The constraint pairs {*OUTRIDE, *OUTRIDE-SHORT LINE} and {*2σ IN W, *3σ IN W} work similarly.

6.6.2 *Sample tableau.* We give in (46) a tableau for *BC* 9. The second row gives the constraint weights and the second column the probabilities predicted by the grammar, reflecting the calculations specified in (38). In the last four rows, the legal candidates for this line (obeying inviolable constraints) are listed. The Hopkins-preferred scansion (Hopkins marked the word *in* as being scanned in S) is marked with a pointing finger. For all four of the legal candidates, the first W position is not shown, because it must be filled by the final syllable of *communion* in the preceding line, reflecting overreaving (§3.2).

32 Constraint weighting employed a Gaussian prior (Goldwater & Johnson 2003: 3), $\sigma = 100{,}000$, $\mu = 0$. The σ value is high but appears not to have resulted in overfitting; see Appendix A: §2.

(46) *Tableau for* BC *9*

/*Here he knelt then in regimental red;* WSWSWSWSWSW/		*2σ IN W	*3σ IN W	*S RESO-LUTION	MATCH SW	*NO CLASH EMPTYW	OUTRIDE	OUTRIDE-WB	OUTRIDE-SL
	weights	1·75	1·75	1·44	1·05	1·74	2·22	1·69	11·47
	prob								
☞ a. $/_s$Here $/_w$he $/_s$knelt $/_w$then $/_s$ín $/_w$re-gi- $/_s$ment- $/_w$al $/_s$red.	0·665				*				
b. $/_s$Here $/_w$he $/_s$knelt $/_w$then ín $/_s$re- $/_w$gi- $/_s$ment- $/_w$al $/_s$red.	0·332	*							
c. $/_s$Here $/_w$he $/_s$knelt $/_w$then $/_s$ín $/_w$re-gi- $/_s$ment- $/_o$al $/_w$Ø $/_s$red.	0·002				*	*	*	*	
d. $/_s$Here $/_w$he $/_s$knelt $/_w$then ín $/_s$re- $/_w$gi- $/_s$ment- $/_o$al $/_w$Ø $/_s$red.	0·001	*				*	*	*	

For this line, the grammar worked fairly well, matching the Hopkins-preferred scansion with the highest-probability candidate. This is because MATCH SW, violated by the winner, has a somewhat lower weight than *2σ IN W, violated by the second candidate. We believe that this candidate would strike many readers (perhaps including Hopkins) as also fairly acceptable. The remaining candidates have a gratuitous outride and empty W and receive very low predicted probabilities.

The tableaux for the remaining 582 lines are too voluminous to print but may be consulted in the supplementary online materials.

6.6.3 *Evaluation of the maxent grammar*. In evaluating our grammar, we will first consider just the 311 lines where we can diagnose Hopkins' own preferences from his diacritic usage. For this purpose, it will be useful to have a baseline grammar with which the maxent grammar can be compared. We use our amplified Kiparskyan grammar; i.e. the eleven Kiparskyan constraints summarised in §3.1 along with the two we added in §4.4. We assume all of these constraints to be inviolable; hence a legal scansion is one which obeys them all.

To compare this baseline grammar with our maxent grammar, we assign it the most sensible stochastic interpretation: every output it designates as legal for a given line is assumed to have the same probability, which is simply its share of the total probability of one, divided by the number of candidates. This 'equiprobable' grammar sometimes does assign a probability of 1, but only for the 47 lines in which it outputs one single well-formed candidate (as it happens, every one of these matches the Hopkins-preferred scansion).

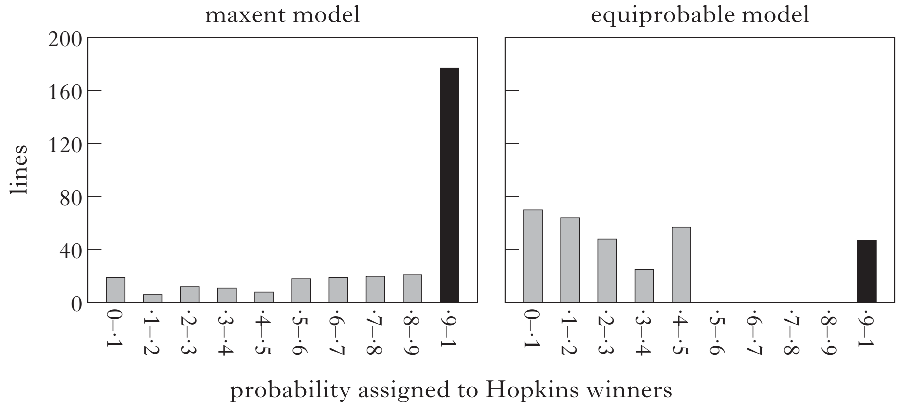


Figure 1
Error 'tails' for the maxent and equiprobable models.

As a first comparison, we consider the median probability assigned to the 311 Hopkins-preferred scansions.[33] For the equiprobable grammar, this is p = 0·25. Our stochastic grammar improves on this considerably, at p = 0·979. Statistical testing, reported in Appendix A, indicates that the difference in performance between the two models is highly significant.

While the median value for our model, 0·979, is encouraging, it is also true that our model has a long, thin 'tail' of cases where it outputs a probability for a Hopkins-preferred scansion that is well below one. It is unknown to us how our model (or data coding) might be improved to reduce this degree of error. However, the error for our model can be usefully contrasted with the comparable error for the equiprobable model, where inappropriate low probabilities are output for Hopkins-preferred scansions in great abundance. The two models are contrasted in Fig. 1 with histograms indicating the frequency of their predicted probabilities for the Hopkins-preferred scansions; values less than 0·9 are shown in grey.

We can also evaluate our grammar against the 272 lines in which we cannot tell unambiguously what scansion Hopkins preferred. The goal here is more modest; we seek a model that seldom predicts scansions that are *incompatible* with Hopkins' diacritics and rhyme patterns. The median total probability assigned by our model to Hopkins-incompatible scansions is 2·7 %, again with a long tail of higher values comparable to that seen in Fig. 1. In contrast, for the equiprobable model, the corresponding value is 69·7 %.

[33] Since the distributions of probabilities assigned to Hopkins-preferred scansions is heavily skewed in both the equiprobable and the maxent model (see the histograms in Fig. 1), we use median probabilities which, unlike means, are resistant to the effects of extreme outliers (Myers *et al.* 2010).

These results lead us to conclude, with Kiparsky (1989), that Hopkins' diacritics primarily serve to direct the reader to a scansion that would be predictable on independent grounds. The exceptional cases may simply be error or random noise, or perhaps reflect idiosyncratic choices made by Hopkins (but only occasionally) for artistic reasons.

6.6.4 *Sprung rhythm and the Parsability Principle.* We hypothesised that the Kiparskyan analysis is in violation of a metrical universal, the Parsability Principle (36), which requires that readers should be able to recover the intended scansion with reasonable probability. We consider now to what extent our stochastic grammar has solved this problem.

As mentioned above (note 25), parsability results only in part from the grammar: it also depends on the parsing strategy employed. The simplest parsing strategy is to examine all the candidate scansions that the grammar assigns to a text line and adopt as one's favoured parse the grammar's highest-probability candidate, picking at random when there are ties. For our 311 lines where the Hopkins-preferred scansion can be inferred, this strategy coupled with our maxent grammar succeeds 84·5% of the time. This strikes us as reasonably high, though probably on the low side for metrical systems in general. It is much higher than is achieved under the equiprobable model, for which the comparable value is 35·4%.

We doubt that there is any grammar that could achieve near-100% success under this strategy, for a reason already given (§5): Hopkins himself was sometimes ambivalent in establishing the correct scansion for a line. Such cases impose a probably irreducible but modest level of unparsability for the Hopkins corpus.

A second way to assess parsability is to assume that an adequate grammar should have *few* possible scansions per line. Although the maxent grammar technically assigns a non-zero probability to every candidate scansion, it assigns probabilities that are very near zero to most of them. Here, we count the number of 'contender' scansions per line – the number of candidate scansions which get some reasonable amount of probability. We define contenders somewhat arbitrarily as candidate scansions which get a probability of more than 5%. With this criterion, most lines have only one or two contender scansions, and no line has more than ten (average 2·4, median 2). 82% of lines have three or fewer scansions that a listener really has to consider. The distribution of numbers of contenders is illustrated in Fig. 2 with a histogram. Again, we conclude that our stochastic grammar produces a system that essentially obeys the Parsability Principle.

6.6.5 *What do the results mean?* We return now to the original concerns that led us to revise and amplify the Kiparskyan system as a stochastic grammar. Recall our evaluation of the Kiparskyan system of inviolable constraints (plus our own two): it is valid to the extent that it predicts what can be a line of sprung rhythm, and that exceptions are far rarer than could ever occur by chance. However, we also characterised the Kiparskyan

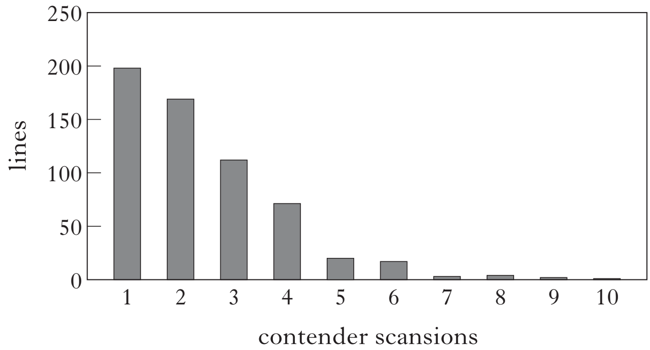


Figure 2
Numbers of 'contender' scansions.

system as one that an adequate general theory of metrics should exclude; this is because it fails to satisfy our Parsability Principle (36), which requires that readers/listeners should in most cases be able to recover the intended scansion from the text.

Our strategy was to augment the grammar with violable constraints, maintaining its rigour by casting it in a formal framework for gradient grammar. When tested, our augmented grammar largely succeeded in matching Hopkins' own preferred scansions. It also provided mostly unambiguous parsability, largely by assigning low probability to implausible scansions. Our results enable us to reaffirm the claim, made intuitively by Hopkins and defended by Kiparsky, that sprung rhythm is a strict metre – but with a stochastic twist: the strictness of our grammar is found not just in the inviolable constraints, but in the strongly biased output probability distributions assigned by the violable constraints.

It remains true that sprung rhythm is not an *easy* metre, either for readers or for metrists, and we address here the question of why this should be so. First, while our gradient grammar succeeds in making most lines unambiguously parsable, we nevertheless judge that they are *less* unambiguously parsable than in most metrical traditions; the very fact that Hopkins invented his system of diacritics attests to this. The degree of parsability seems to be sufficient for experienced modern readers to perceive the verse as metrical and to get a good intuitive sense (for most lines) of the intended scansion. But the sense of correct scansion is often uncertain enough to make the verse difficult for the reader – and indeed, to form a barrier, as Kiparsky (1989) demonstrated, to metrical analysis.[34]

[34] The case of sprung rhythm is an unusual one, in which the actual posthumous reputation of a poet depends, at least in principle, on the correctness of a linguistic analysis. Throughout the period sprung rhythm has been studied, researchers have proposed highly unrestrictive characterisations of this metre, blaming not themselves but Hopkins for writing ill-regulated verse. Kiparsky (1989: 305–309) critiques a number of accounts along these lines; and the recent analysis of Fabb & Halle (2008: 89–90) might also reasonably be included among their number.

In addition, it strikes us as unusual that the full grammar could not be learned until we consulted Hopkins' diacritic markings. From our viewpoint, the diacritics evidently have a double purpose. For individual lines, they serve as a direct crib for indicating the scansion. More speculatively, they could be regarded as *training data*: the reader who has read enough sprung rhythm with the diacritics included could – in principle – develop an intuitive sense of correct scansion, and ultimately would be able to use this sense to find the intended scansion even without the assistance of the diacritics.[35] A metrical system that requires diacritic training data to learn is quite unusual, and it may well be that people more easily learn metrical systems in which the poet's intent is recoverable solely from the text.[36]

7 Implications

Although sprung rhythm is a somewhat recherché topic, we think our work has implications for metrics and for linguistics in general.

Concerning metrics, we think the comparison we have made is informative concerning the prospects of a theory of metrics based entirely on all-or-nothing constraints, as proposed recently by Fabb (2001, 2002, 2006) and Fabb & Halle (2008). The ability of our analysis to predict, with fair accuracy, what Hopkins' preferred scansion would be is entirely dependent on the use of violable constraints.

Concerning linguistic theory, we present our work as an example of a tradition established by Bresnan *et al.* (2007), who demonstrated that by using a set of violable constraints, deployed as a stochastic grammar, it is possible to predict with surprising accuracy which version of the dative construction (NP PP or NP NP) English speakers will employ. We have likewise used stochastic grammar to predict which of the legal scansions of his system Hopkins will employ. While the accuracy of our system falls short of Bresnan *et al.*'s (84·5% *vs.* 94%), our system is also making its choices from a set that can include dozens or even hundreds of possibilities, rather than just two. We suggest the term 'quasi-prediction' for what Bresnan *et al.* and we have done, and suggest that it would be profitable in general for linguists to seek out and analyse phenomena that are quasi-predictable.

35 Failing this, we note that most of our added stochastic constraints are standard-issue items of English metrics, and readers of Hopkins could in principle extrapolate from their intuitive knowledge of other English poetry to arrive at an approximation of Hopkins' intended scansions.

36 Another form of verse in which the hearer may benefit from 'external' evidence for scansion is rap, where a musical accompaniment helps guide the hearer to the location of the intended strong metrical positions. Our experience reading rap lines set as text suggests that this metrical genre might also be quite difficult to learn and scan in the absence of this external cue.

Lastly, we note some methodological implications of our work. By expressing the lines of Hopkins' poetry in prosodically annotated digital form and scanning them with software, we found it was feasible to test a much larger number of hypotheses than we otherwise might, at the same time reducing the possibility of error. The digital approach, combined with the maxent framework for gradient grammars, also made it possible to test our constraints statistically, thus excluding those that could be true by accident.

Appendix A: Statistical tests

1 Comparison of the maxent and equiprobable models

In this appendix, we address the question of whether our stochastic model significantly outperforms the 'equiprobable model', in which all candidate scansions output by the amplified Kiparskyan grammar are given an equal probability of occurrence. Using the equiprobable model as our null hypothesis, we examine the proportion of 'correct guesses' that the stochastic model makes. A correct guess is defined as a case in which the stochastic grammar's most probable candidate is the scansion which agrees with Hopkins' diacritic markings. We test only the 311 lines of the corpus for which only one candidate agrees with the diacritics. We use a z-test for proportions (Healey 1999: 196), comparing the proportion of correct guesses of the stochastic grammar ($\hat{p}$) to the proportion of correct guesses expected by chance according to the null hypothesis (p_0). The statistic z is calculated according to the formula in (47).

(47)
$$z = \frac{(\hat{p} - p_0)}{\sqrt{\left(\frac{p_0(1-p_0)}{n}\right)}}$$

Because different lines have different numbers of candidates, the proportion correct expected by chance differs from line to line. For example, in a line with four candidate scansions, guessing at chance would result in a correct guess 25 % of the time. In a line with five candidate scansions, chance would be 20 %. To solve this problem, we divide the corpus into groups based on number of candidate scansions. Within each group, we test the proportion of correct guesses made by the stochastic model against the expected number of correct guesses made by chance. There are 33 groups, so this means running 33 separate z-tests. To correct for family-wise error, we use the Bonferroni correction, and set our significance level at 0·05/33 = 0·0015.

Table III lists the groups, proportion 'correct', expected proportion 'correct' and p-values. Ignoring groups with only one or two lines in them, the proportion of correct guesses made by the stochastic grammar is always significantly greater than chance. From this we conclude that our model performs significantly better than the equiprobable model at selecting Hopkins-preferred scansions as winners.

C	n	g	$\hat{p}$	p_0	z	p-value
2	57	54	0·95	0·50	6·76	0·0000 *
3	25	23	0·92	0·33	6·22	0·0000 *
4	48	39	0·81	0·25	9·00	0·0000 *
5	14	13	0·93	0·20	6·82	0·0000 *
6	25	21	0·84	0·17	9·03	0·0000 *
7	2	1	0·50	0·14	1·44	0·0745
8	20	15	0·75	0·13	8·45	0·0000 *
9	3	2	0·67	0·11	3·06	0·0011 *
10	5	5	1·00	0·10	6·71	0·0000 *
11	4	3	0·75	0·09	4·59	0·0000 *
12	16	12	0·75	0·08	9·65	0·0000 *
13	4	4	1·00	0·08	6·93	0·0000 *
14	3	3	1·00	0·07	6·24	0·0000 *
15	2	2	1·00	0·07	5·29	0·0000 *
16	12	10	0·83	0·06	11·03	0·0000 *
17	1	0	0·00	0·06	−0·25	0·4013
18	1	1	1·00	0·06	4·12	0·0000 *
20	6	5	0·83	0·05	8·80	0·0000 *
25	1	1	1·00	0·04	4·90	0·0000 *
30	2	0	0·00	0·03	−0·26	0·3964
31	1	1	1·00	0·03	5·48	0·0000 *
32	1	0	0·00	0·03	−0·18	0·4287
40	1	1	1·00	0·03	6·24	0·0000 *
48	1	0	0·00	0·02	−0·15	0·4420
51	1	0	0·00	0·02	−0·14	0·4438
52	1	0	0·00	0·02	−0·14	0·4443
72	1	0	0·00	0·01	−0·12	0·4528
74	1	1	1·00	0·01	8·54	0·0000 *
77	1	0	0·00	0·01	−0·11	0·4543
100	1	0	0·00	0·01	−0·10	0·4600
102	1	0	0·00	0·01	−0·10	0·4604
128	1	1	1·00	0·01	11·27	0·0000 *
932	1	0	0·00	0·00	−0·03	0·4869

Table III

Z-test results for 33 line groups. C is the number of candidate scansions per line, n is the number of instances of lines with C candidate scansions, and g is the number of 'correct guesses' in that group. $\hat{p}$ is equal to g/C. An * indicates significance ($p < 0{\cdot}0015$).

2 Checking the model against overfitting

Overfitting arises when grammars are tested against the same data on which they have been trained. We need to guard against the possibility that the weights of the model have been adjusted to match idiosyncrasies of the training data, meaning that that they would not yield as accurate predictions were it possible to test them against additional data.

One way to check against overfitting in a limited data sample is through tenfold cross-validation (Witten & Frank 2005). In this method, the lines are divided randomly into ten equal groups. For each group, the constraint weights are computed on the basis of the remaining nine-tenths of the data, then, using these weights, the probabilities are calculated for the lines in the group. This is done ten times, so that in the end every line L in the training data has been assigned predicted probabilities for its possible scansions using data that did not include L. This makes it possible to check how well the grammar works when it lacks the advantage of training on its own testing data.

When we carried out this procedure, we found that the weights obtained were generally similar to what we had found earlier: the median difference between the weights of (47) and those obtained in the ten cross-validation trials was only 4·2%. Moreover, when we examined the predictions of the cross-validation procedure, we found that performance declined only slightly: indeed, the median probability assigned to a Hopkins-preferred scansion remained identical, at 0·979. With regard to parsability, the effectiveness of the 'first-guess' strategy outlined in §6.6.4 declines from 84·5% to 84·2%. It appears, then, that our analysis works largely because it matches general patterns in the data rather than being accidentally fitted to the particular lines of the corpus.

Appendix B: Other frameworks

As mentioned above, maxent is just one framework that can be used to form constraint-based stochastic grammars. We tried three others, employing the same constraints and training data.

Noisy Harmonic Grammar (Boersma & Pater 2008) is similar to maxent, but instead of directly generating output probabilities, it always outputs the most harmonic candidate; stochastic behaviour is induced by letting the constraint weights vary at random about their central values in a Gaussian distribution. We calculated the constraint weights using Praat (Boersma & Weenink 2010).

Stochastic Optimality Theory (Boersma 1997, Boersma & Hayes 2001) uses Gaussian probability distributions over a 'ranking scale' to establish the probabilities of constraint rankings; otherwise it works like classical Optimality Theory (Prince & Smolensky 1993). We calculated the ranking values for the constraints with the Gradual Learning Algorithm (Boersma 1997) using the implementation in OTSoft (Hayes *et al*. 2010).

In the theory of partially ordered grammars (Anttila 1997a, b, Kiparsky 2005, 2006), only some rankings are specified; in calculating outputs, equal probability is assumed for all total rankings that are compatible with the specified rankings. This theory obtains quantitative predictions without having to include numbers in the grammar itself. There is as yet no ranking algorithm for this theory and the ranking we adopted ({inviolable constraints} ≫ *OUTRIDE-SHORT LINE ≫ {all others}) was found by experimentation.[37]

All three frameworks gave results similar to maxent. Some heuristic comparisons appear in Table IV; full sets of predictions are on the article website.

constraint	median probability assigned to Hopkins-preferred candidates	'first-guess' parsability (%)	bad high guesses	bad low guesses
Equiprobability (§6.6.3)	0·25	35·4	0	3
Maxent	0·979	84·5	0	2
Maxent cross-validated (§6.2)	0·979	84·2	0	3
Noisy HG	1	83·0	20	17
Stochastic OT	1	84·2	9	17
Partially ordered OT	1	86·1	8	17

Table IV

Predicting the Hopkins scansions with different constraint-based models.

In the table, 'median probability' is defined as in §6.6.3 and 'first guess parsability' as in §6.6.4. A 'bad high guess' is defined as predicting a probability $>0{\cdot}9$ for a scansion not preferred by Hopkins, and a 'bad low guess' as predicting a probability $<0{\cdot}01$ for a Hopkins-preferred scansion.

In the comparisons, the three additional models examined here emerged as somewhat 'edgier' than maxent: they succeeded in allocating more probability to the Hopkins-preferred scansions, but when they erred, they erred more drastically. This is perhaps a consequence of the fact that these models are 'winner take all' at the level of candidate evaluation, whereas maxent allocates probability more broadly, permitting even harmonically bounded winners to receive a minority share.

Since the performance is so similar, we do not see these results as useful for the purpose of evaluating the relative performance of these frameworks in general.

[37]Learning parameters for GLA: evaluation noise 2, plasticity on a gradually descending scale from 2 to 0·001, 1,000,000 training trials. For Noisy HG: evaluation noise 2, plasticity 1, 100,000 training trials. OT with partial rankings was approximated with an equivalent Stochastic OT grammar with highly dispersed ranking values. All models were tested with 100,000 trials. Harmonically bounded winners (see immediately below) did not damage the performance of these models, so we included them.

Of the 311 Hopkins-preferred scansions, 15 are harmonically bounded in the sense of Prince & Smolensky (1993: §9.1.1). The three models under discussion always assign zero probability to harmonically bounded candidates; hence the harmonically bounded winners are a substantial source of error for these models. They are somewhat less troublesome for the maxent model, in which harmonically bounded candidates receive a positive probability (never higher than the candidates that bound them): our maxent grammar of Table II assigns these 15 scansions probabilities ranging from 0·001 to 0·5 (average 0·142).

Since it is a question of some theoretical interest whether grammatical theory should ever assign positive frequencies to harmonically bounded candidates, we inspected these fifteen scansions. The most interesting case is *HH* 8 ((48a)), where Hopkins has assigned an outride before a rather weak phonological break when he could have placed it before a stronger (punctuated) one, as in (48b). In our grammar, (48a) incurs violations for constraints penalising outrides before weak breaks; (48b) does not, and all else is equal. Line *HH* 1 works similarly.

(48) *A possible case of a harmonically bounded winner* (*HH* 8)

a. *Hopkins-preferred scansion* (*harmonically bounded*)

$/_{s}$Rap- $/_{o}$turous $/_{w}$love's $/_{s}$greet- $/_{o}$<u>ing</u> $/_{w}$of $/_{s}$real- $/_{w}$er, of $/_{s}$round-
$/_{w}$er re $/_{s}$plies?

b. *Bounding candidate*

$/_{s}$Rap- $/_{o}$turous $/_{w}$love's $/_{s}$greet- $/_{w}$ing of $/_{s}$real- $/_{o}$<u>er</u>, $/_{w}$of $/_{s}$round-
$/_{w}$er re $/_{s}$plies?

Thus, it is possible that for (presumably) purely artistic reasons Hopkins chose to use his diacritics to annotate a less probable candidate as his preferred one. (In our maxent grammar, the candidate probabilities are 0·057 for (48a) *vs.* 0·309 for (48b).) This seems a more plausible claim than that Hopkins chose to mark a completely unmetrical scansion, as must be said in other frameworks. Given the thinness of the data, this argument must count as almost conceptual in character, but it does point out an important point of difference among these frameworks that deserves further study.

Appendix C: poems studied, with title abbreviations

AB	Ashboughs
BC	The Bugler's First Communion
BP	Binsey Poplars
BR	Brothers
CC	Carrion Comfort
CS	The Caged Skylark
DS	Duns Scotus's Oxford
FR	Felix Randal
HF	That Nature is a Heraclitean Fire and of the Comfort of the Resurrection
HH	Hurrahing in Harvest

HP	Harry Ploughman
HR	Henry Purcell
ID	Inversnaid
KF	As Kingfishers Catch Fire
LE	The Loss of the Eurydice
MM	The May Magnificat
NW	No Worst
PB	Pied Beauty
RB	Ribblesdale
SD	The Soldier
SF	Spring and Fall
SS	Spelt from Sibyl's Leaves
TG	Tom's Garland
WH	The Windhover
WM	At the Wedding March

REFERENCES

Abbott, Claude Colleer (1935a). *The correspondence of Gerard Manley Hopkins and Richard Watson Dixon*. London: Oxford University Press.

Abbott, Claude Colleer (1935b). *The letters of Gerard Manley Hopkins to Robert Bridges*. London: Oxford University Press.

Abbott, Claude Colleer (1956). *Further letters of Gerard Manley Hopkins*. 2nd edn. London: Oxford University Press.

Anttila, Arto (1997a). *Variation in Finnish phonology and morphology*. PhD dissertation, Stanford University.

Anttila, Arto (1997b). Deriving variation from grammar. In Frans Hinskens, Roeland van Hout & W. Leo Wetzels (eds.) *Variation, change and phonological theory*. Amsterdam & Philadelphia: Benjamins. 35–68.

Attridge, Derek (1982). *The rhythms of English poetry*. London & New York: Longman.

Bailey, James (1975). *Toward a statistical analysis of English verse : the iambic tetrameter of ten poets*. Lisse: de Ridder.

Barnes, Mervin & Helmut Esau (1978). English prosody reconsidered. *Language and Style* **11**. 212–222.

Biggs, Henry (1996). *A statistical analysis of the metrics of the classic French decasyllable and the classic French alexandrine*. PhD dissertation, University of California, Los Angeles.

Boersma, Paul (1997). How we learn variation, optionality, and probability. *Proceedings of the Institute of Phonetic Sciences of the University of Amsterdam* **21**. 43–58.

Boersma, Paul & Bruce Hayes (2001). Empirical tests of the Gradual Learning Algorithm. *LI* **32**. 45–86.

Boersma, Paul & Joe Pater (2008). Convergence properties of a gradual learning algorithm for Harmonic Grammar. Ms, University of Amsterdam & University of Massachusetts, Amherst. Available as ROA-970 from the Rutgers Optimality Archive.

Boersma, Paul & David Weenink (2010). *Praat : doing phonetics by computer* (version 5.1.42). http://www.praat.org/.

Bresnan, Joan, Anna Cueni, Tatiana Nikitina & R. Harald Baayen (2007). Predicting the dative alternation. In Gerlof Bouma, Irene Krämer & Joost Zwarts (eds.)

Cognitive foundations of interpretation. Amsterdam: Royal Netherlands Academy of Arts and Sciences. 69–94.
Bridges, Robert (1921). *Milton's prosody: with a chapter on accentual verse*. London: Oxford University Press.
Callow, John C. (1965). Kasem nominals: a study in analyses. *Journal of West African Languages* **2**. 29–36.
Chomsky, Noam & Morris Halle (1968). *The sound pattern of English*. New York: Harper & Row.
Della Pietra, Stephen, Vincent Della Pietra & John Lafferty (1997). Inducing features of random fields. *IEEE Transactions: Pattern Analysis and Machine Intelligence* **19**. 380–393.
de Lacy, Paul (2004). Markedness conflation in Optimality Theory. *Phonology* **21**. 145–199.
Dresher, B. Elan & Nila Friedberg (eds.) (2006). *Formal approaches to poetry: recent developments in metrics*. Berlin & New York: Mouton de Gruyter.
Dresher, B. Elan & Aditi Lahiri (1991). The Germanic foot: metrical coherence in Old English. *LI* **22**. 251–286.
Duda, Richard O., Peter E. Hart & David G. Stork (2001). *Pattern classification*. 2nd edn. New York: Wiley.
Elwert, W. Theodor (1984). *Italienische Metrik*. 2nd edn. Wiesbaden: Steiner.
Fabb, Nigel (2001). Weak monosyllables in iambic verse and the communication of metrical form. *Lingua* **111**. 771–790.
Fabb, Nigel (2002). *Language and literary structure: the linguistic analysis of form in verse and narrative*. Cambridge: Cambridge University Press.
Fabb, Nigel (2006). Generated metrical form and implied metrical form. In Dresher & Friedberg (2006). 77–91.
Fabb, Nigel & Morris Halle (2008). *Meter in poetry: a new theory*. Cambridge: Cambridge University Press.
Goldwater, Sharon & Mark Johnson (2003). Learning OT constraint rankings using a Maximum Entropy model. In Jennifer Spenador, Anders Eriksson & Östen Dahl (eds.) *Proceedings of the Stockholm Workshop on Variation within Optimality Theory*. Stockholm: Stockholm University. 111–120.
Gussenhoven, Carlos (1986). English plosive allophones and ambisyllabicity. *Gramma* **10**. 119–141.
Halle, Morris & Samuel Jay Keyser (1966). Chaucer and the study of prosody. *College English* **28**. 187–219.
Hanson, Kristin (1992). *Resolution in modern meters*. PhD dissertation, Stanford University.
Hanson, Kristin (2002). Quantitative meter in English: the lesson of Sir Philip Sidney. *English Language and Linguistics* **5**. 41–91.
Hanson, Kristin & Paul Kiparsky (1996). A parametric theory of poetic meter. *Lg* **72**. 287–335.
Hayes, Bruce (1983). A grid-based theory of English meter. *LI* **14**. 357–393.
Hayes, Bruce (1989). The prosodic hierarchy in meter. In Kiparsky & Youmans (1989). 201–260.
Hayes, Bruce (1995). *Metrical stress theory: principles and case studies*. Chicago: University of Chicago Press.
Hayes, Bruce (2009). *Introductory phonology*. Malden, Mass. & Oxford: Wiley-Blackwell.
Hayes, Bruce, Bruce Tesar & Kie Zuraw (2010). *OTSoft 2.3.1*. Software package. http://www.linguistics.ucla.edu/people/hayes/otsoft/.
Hayes, Bruce & Colin Wilson (2008). A maximum entropy model of phonotactics and phonotactic learning. *LI* **39**. 379–440.

Healey, Joseph F. (1999). *Statistics: a tool for social research*. Belmont, Ca.: Wadsworth.

Holder, Alan (1995). *Rethinking meter: a new approach to the verse line*. Lewisburg, Pa.: Bucknell University Press.

House, Humphry (1959). *The journals and papers of Gerard Manley Hopkins*. London: Oxford University Press.

Hurley, Michael D. (2005). Darkening the subject of Hopkins' prosody. *Victorian Poetry* **43**. 485–496.

Jespersen, Otto (1933). Notes on metre. In Otto Jespersen. *Linguistica: selected papers in English, French and German*. Copenhagen: Levin & Munksgaard. 249–274. Paper originally presented in Danish in 1900.

Kahn, Daniel (1976). *Syllable-based generalizations in English phonology*. PhD dissertation, MIT. Published 1980, New York: Garland.

Kenstowicz, Michael (1994). *Phonology in generative grammar*. Cambridge, Mass. & Oxford: Blackwell.

Kessler, Brett (1992). *External sandhi in Classical Sanskrit*. MA thesis, Stanford University.

Kiparsky, Paul (1975). Stress, syntax, and meter. *Lg* **51**. 576–616.

Kiparsky, Paul (1977). The rhythmic structure of English verse. *LI* **8**. 189–247.

Kiparsky, Paul (1982). Lexical phonology and morphology. In The Linguistic Society of Korea (ed.) *Linguistics in the morning calm*. Seoul: Hanshin. 3–91.

Kiparsky, Paul (1989). Sprung rhythm. In Kiparsky & Youmans (1989). 305–340.

Kiparsky, Paul (2005). Where Stochastic OT fails: a discrete model of metrical variation. *BLS* **31**. 409–430.

Kiparsky, Paul (2006). A modular metrics for folk verse. In Dresher & Friedberg (2006). 7–49.

Kiparsky, Paul & Gilbert Youmans (eds.) (1989). *Rhythm and meter*. San Diego: Academic Press.

Koelb, Clayton (1979). The iambic parameter revisited. *Neophilologus* **63**. 321–329.

Ladd, D. Robert (1986). Intonational phrasing: the case for recursive prosodic structure. *Phonology Yearbook* **3**. 311–340.

McCarthy, John J. (2003). OT constraints are categorical. *Phonology* **20**. 75–138.

MacKenzie, Norman H. (ed.) (1990). *The poetical works of Gerard Manley Hopkins*. Oxford: Clarendon Press.

MacKenzie, Norman H. (1991). *The later poetic manuscripts of Gerard Manley Hopkins in facsimile*. New York & London: Garland.

Magnuson, Karl & Frank G. Ryder (1970). The study of English prosody: an alternative proposal. *College English* **31**. 789–820.

Malouf, Robert (2002). A comparison of algorithms for maximum entropy parameter estimation. In Dan Roth & Antal van den Bosch (eds.) *Proceedings of the 6th Workshop on Computational Language Learning* (*CoNLL-2002*). New Brunswick: Association for Computational Linguistics. 49–55.

Mester, Armin (1994). The quantitative trochee in Latin. *NLLT* **12**. 1–61.

Moore-Cantwell, Claire (2009). Gerard Manley Hopkins' diacritics: a corpus-based study. Ms, University of California, Los Angeles.

Myers, Jerome L., Arnold D. Well & Robert F. Lorch, Jr. (2010). *Research design and statistical analysis*. 3rd edn. New York: Routledge.

Nespor, Marina & Irene Vogel (2007). *Prosodic phonology*. 2nd edn. Berlin: Mouton de Gruyter.

Opie, Iona & Peter Opie (1951). *The Oxford dictionary of nursery rhymes*. London: Oxford University Press.

Oxford English Dictionary. Online edition (accessed May 2011). Oxford: Oxford University Press. http://oed.com.

Pater, Joe & Anne-Michelle Tessier (2003). Phonotactic knowledge and the acquisition of alternations. In M. J. Solé, D. Recasens & J. Romero (eds.) *Proceedings of the 15th International Congress of Phonetic Sciences*. Barcelona: Causal Productions. 1177–1180.
Pinheiro, José C. & Douglas M. Bates (2000). *Mixed-effects models in S and S-PLUS*. New York: Springer.
Prince, Alan (1989). Metrical forms. In Kiparsky & Youmans (1989). 45–80.
Prince, Alan & Paul Smolensky (1993). *Optimality Theory: constraint interaction in generative grammar*. Ms, Rutgers University & University of Colorado, Boulder. Published 2004, Malden, Mass. & Oxford: Blackwell.
Russom, Geoffrey (1987). *Old English meter and linguistic theory*. Cambridge: Cambridge University Press.
Selkirk, Elisabeth (1980). Prosodic domains in phonology: Sanskrit revisited. In Mark Aronoff & Mary-Louise Kean (eds.) *Juncture*. Saratoga: Anma Libri. 107–129.
Selkirk, Elisabeth (1984). *Phonology and syntax: the relation between sound and structure*. Cambridge, Mass: MIT Press.
Sihler, Andrew L. (1995). *New comparative grammar of Greek and Latin*. Oxford: Oxford University Press.
Siptár, Péter & Miklós Törkenczy (2000). *The phonology of Hungarian*. Oxford: Oxford University Press.
Smolensky, Paul & Géraldine Legendre (eds.) (2006). *The harmonic mind: from neural computation to optimality-theoretic grammar*. 2 vols. Cambridge, Mass.: MIT Press.
Sprott, S. Ernest (1953). *Milton's art of prosody*. Oxford: Blackwell.
Steele, Timothy (1999). *All the fun's in how you say a thing: an explanation of meter and* versification. Athens: Ohio University Press.
Tarlinskaja, Marina (1976). *English verse: theory and history*. The Hague & Paris: Mouton.
Tarlinskaja, Marina (2006). What is 'metricality'? English iambic pentameter. In Dresher & Friedberg (2006). 53–76.
Tarlinskaja, Marina & L. M. Teterina (1974). Verse – prose – metre. *Linguistics* **129**. 63–86.
Wells, John C. (1982). *The accents of English*. 3 vols. Cambridge: Cambridge University Press.
Wilson, Colin (2006). Learning phonology with substantive bias: an experimental and computational study of velar palatalization. *Cognitive Science* **30**. 945–982.
Wilson, Colin & Benjamin George (2009). Maxent grammar tool. Computer program. Available at http://www.linguistics.ucla.edu/people/hayes/MaxentGrammarTool.
Wimsatt, James I. (1998). Alliteration and Hopkins's sprung rhythm. *Poetics Today* **19**. 531–564.
Wimsatt, James I. (2006). *Hopkins's poetics of speech sound: sprung rhythm, lettering, inscape*. Toronto: University of Toronto Press.
Witten, Ian H. & Eibe Frank (2005). Data mining: practical machine learning tools and techniques. 2nd edn. San Francisco: Morgan Kaufmann.
Young, George (1928). *An English prosody on inductive lines*. Cambridge: Cambridge University Press.
Youmans, Gilbert (1989). Milton's meter. In Kiparsky & Youmans (1989). 341–379.