Automatic Parsing for Shakespearean Meters

Term Paper for Metrics with Paul Kiparsky

Sam Bowman

December 16, 2011

1 Introduction

Hanson and Kiparsky (1996, H&K) lay out a formal grammar for Shakespearean iambic pentameter defined by a small set of parameter settings, framed as constraints. In this paper, I investigate the possibility that different sub-genres of Shakespeare's writing vary, categorically or on a gradient, in their settings for these parameters.

To this end, I adapt the still under development Anttila et al. (2011) PROSODIC metrical parser for Shakespearean meter, and describe here both the structure of the parser, and the adjustments to the H&K model that are needed to make it implementable.

2 Parsing Shakespeare Automatically

2.1 The PROSODIC Parser

PROSODIC generates metrical parses in three steps: In the first step, it reads in lines of text, associates each word of each line with a pronunciation (or multiple candidate pronunciations) from the CMU Pronouncing Dictionary (The Carnegie Mellon University Speech Group, 2007). In the next step, it generates a set of parses—assignments of syllables to alternating strong and weak metrical positions—consistent with a set of hard constraints built into the model. In the final step, it evaluates the parses for each line according to a set of user-configurable weighted constraints and selects the parse with the lowest violation score. A successful parse looks like this:

(1) `and from the powdering tub of infamy`
→ `and|FROM|the|PO|wd.ering|TUB|of|IN|fa|MY`

Two things are immediately striking in this representation. The parser uses a very simple heuristic to syllabify orthographic text in the output, leading to degenerate syllables like `wd`, but its parses reflect a high quality, hand-corrected syllabification

of the word in CMU's phonemic notation. The syllabification and footing above is simply an unfaithful representation of the underlying `|'P AW | d er . ih ng |`. More substantively, while the parser does recognize genuine pronunciation variation, it does not allow for the kind of adaptive re-syllabification that is often used in Shakespearean poetry, of the sort that might render *fire* as two syllables or *being* as one. In this case, the parser does not consider the metrical disyllabic pronunciation `|'P AW D | r ih ng |`.

2.2 Constraints and Parameters for Shakespeare

In this section I lay out how best to configure PROSODIC in order to generate plausible parses of Shakespearean iambic pentameter. Hanson and Kiparsky (1996) lay out five formal properties of Shakespearean meter, which I use as the basis for my model:

(2) A weak metrical position may not contain a strong syllable.

(3) A metrical position may contain one or two syllables.

(4) A disyllabic metrical position may not contain a lexical monosyllable.

(5) A disyllabic metrical position may contain a strong syllable of a lexical word only if the syllable is (i) light and (ii) followed within the same position by an unstressed syllable normally belonging to the same word.

(6) Line initial inversion is permitted.

Prosodic encodes the stress, weight, and strength status of each syllable, and the size of each metrical position, so constraints 2, 4, and 5 can be written as violable constraints in the model precisely as described.

Constraint 3 can be formulated in two ways: as a violable constraint against polysyllabic feet, or as a hard constraint preventing the parser from generating candidates with polysyllabic feet. Since constraints 4 and 5 are not well-defined for feet of more than two syllables, it is possible to define them such that it would be optimal for a parser to create trisyllabic feet in order to avoid violating them in a line that is slightly unmetrical. To avoid this potential problem, which I take to be more fatal to metricality than an overly weighty disyllabic position, I implement constraint 3 as a hard constraint, and never generate trisyllabic positions.

Two additional properties are implicated in the H&K model, but not included in the above list:

(7) The first metrical position must be weak.

(8) Metrical positions alternate from strong to weak.

Both of these are ostensibly built in to PROSODIC as inviolable constraints, but 7—essentially the stipulation of an iambic meter—does not appear to have been fully implemented, and I replace it with a heavily weighted violable constraint.

Because the model holds constraint 8 to be inviolable, and because constraint 7 is the only means we have to enforce an iambic meter, it is impossible to formulate constraint 6, allowing inversion, directly: The parser will never consider the sequence of metrical positions *sws*, and will avoid selecting any sequence beginning with *sw*. I address this by instead modifying constraint 2, the only constraint which refers directly to strong metrical positions, and replace it with the following:

(9) A weak metrical position may not contain a strong syllable, *unless it is in the first metrical position of a line.*

This has the effect of preserving a nominally alternating meter, but allowing for any arrangement of lexical primary stresses in the first foot, including a disyllabic word with an initial strong heavy syllable. This approach is consistent with the H&K model, and shares with that model the possibility of allowing one configuration which would be illegal under an analysis with true trochaic substitution (an *sw* first foot): an initial weak position containing a disyllabic word by resolution, followed by a strong position containing a strong syllable. I could not find an example of this configuration in my corpora, but this modified Shakespearean line (originally from Sonnet 50) would be fully metrical under the present approach, despite containing strong syllables in both of the first two metrical positions:

(10) `very heavy do i journey on my way`
`ve.ry|HE|avy|DO|i|JOU|rney|ON|my|WAY`

Two more low-weighted violable *preference* constraints help to select among otherwise equally good parses:

(11) Function words may not appear in strong positions.

(12) Metrical positions may not contain multiple syllables.

For Prosodic to use a set of violable constraints, it must be ranked. For many lines, there exist parses that violate none of these constraints, and for these lines, the ranking is immaterial. However, when no such parse is available for a line, due either to a failure of the parser's syllabifications or to genuine unmetricality, the weighting of the constraints determines which unmetrical parse is the closest fit to the line. Given the noise that the dictionary introduces in this model, the choice of weighting may actually be essential to recovering correct parses that do not appear fully metrical for the syllabification being used.

Without a corpus of expert parses of difficult lines, it is difficult to determine precisely what assigment of weights yields the *best* parse of a difficult line, or what *best* would mean in this situation. Purely subjective experimentation yielded the following assignment of weights to the violable constraints:

- The first metrical position must be weak. [**10,000**]

- A weak metrical position may not contain a strong syllable, *unless it is in the first metrical position of a line.* [**100**]
- A disyllabic metrical position may not contain a lexical monosyllable. [**100**]
- A disyllabic metrical position may contain a strong syllable of a lexical word only if the syllable is (i) light and (ii) followed within the same position by an unstressed syllable normally belonging to the same word. [**100**]
- Function words may not appear in strong positions. [**1**]
- Metrical positions may not contain multiple syllables. [**1**]

Since the multiple between differing weights (100) is larger than the maximum possible number of violations of any one constraint in a line, the gang effect that are characteristic of weighted-constraint harmonic grammar are absent here, and the constraint grammar mimics an OT grammar with a stratified constraint ranking.

2.3 Parsing Performance

The parser generally did quite well: For the vast majority of metrical lines, it produced a reasonable output with no violation of the crucial (100-weighted and 10,000-weighted) constraints. Here is a selection of the parser's output, selected more-or-less at random:

(13) `the spirit of love with a perpetual dullness:`
`the|SPI.RIT|of|LOVE|with|A|pe|RP|et.ual|DUL|lness`

(14) `let this sad interim like the ocean be:`
`let|THIS|sad|INT|erim|LIKE|the|OC|ean|BE`

(15) `which parts the shore where two contracted new:`
`which|PARTS|the|SHORE|where|TWO|con|TRA|cted|NEW`

(16) `come daily to the banks that when they see:`
`come|DAI|ly|TO|the|BANKS|that|WHEN|they|SEE`

(17) `and he is bred out of that bloody strain:`
`and|HE|is|BRED|out|OF|that|BLO|ody|STRAIN`

(18) `witness our too much memorable shame:`
`wit|NESS|ou.r|TOO|much|ME|mo|RA|ble|SHAME`

(19) `when cressy battle fatally was struck:`
`when|CRE|ssy|BA|ttle|FA|ta|LLY|was|STRUCK`

The bulk of the problematic parses can be attributed to failures in syllabification. The line below, for example, is metrical if *even* is contracted to a single syllable, but such a contraction is unavailable to the parser. Instead, it is forced to place the

primary stress of *even* in a strong position, illegally wedging *eyes* into a disyllabic weak position.

(20) `thy hungry eyes even till they wink with fullness:`
`thy|HUN|gry.eyes|EV|en|TILL|they|WINK|with|FUL|lness`

Errors like this can even lead to pentameter lines being parsed with four or six strong positions, as here. Note the problematic *desire* and *blessed*:

(21) `then can no horse with my desire keep pace`
`then|CAN|no|HORSE|with|MY|de|SI|re|KEEP|pace`

(22) `blessed are you whose worthiness gives scope`
`blessed|ARE|you|WHOSE|wor|THI|ness|GIVES|scope`

Fortunately, these errors are both infrequent enough and uniform enough, that it may be possible to discover some consistent differences in genres in spite of them.

3 Experiment 1: Case Study

3.1 Corpus

I draw my texts from The Moby Project and MIT's (1993) free complete works of Shakespeare. This choice was a pragmatic one: PROSODIC cannot parse a line in which any word is not in the pronouncing dictionary. Without building a sophisticated system for handling unknown words and typographical flaws, I thus require an edited electronic text in modern spelling.

For a preliminary experiment, I compare the parser's behavior on three styles of Shakespearean text: sonnets, casual dramatic scenes with short pieces of dialog among lower-class characters, and more pivotal dramatic scenes with longer pieces of dialog and intermittent soliloquies. The sources I used, which I selected for the purpose, were the following:

- Sonnets: Sonnets 50–60
- Pivotal dialog: *Henry V*, 2.4 (France. The KING's palace.), 3.4 (The Heath. Before a hovel.)
- Casual dialog: *Henry V*, 2.1 (London. A Street.), 2.3 (London. Before a tavern.); Comedy of Errors, 2.2 (A public place.), 4.1 (A public place.)

I stripped the dialogs of all but the actual spoken lines (removing stage directions), and removed the first and last typographical lines of each dramatic line to avoid short lines. I then used a script to scan the resulting text and remove lines which contain words with unknown pronunciations. I finally selected the first 100 lines of each genre for testing.

3.2 Results

In light of the frequent occurrence of minor errors introduced by the dictionary, considering the pure percentage of metrical lines would likely be as misleading as it would be uninformative. Instead, I consider the ways in which unmetrical lines diverge from the meter: the different constraints which are violated. Proportional differences in these are more likely to reflect real variation in the meter: A constraint that is frequently violated in optimal parses for some text may be nonexistent, or at least minimally salient, in the meter of that text.

The table below tests this. It should be noted that in weighting the first constraint, against trochee-initial lines, so highly, I essentially bar violating lines from ever being optimal. Even a clearly trochaic line would be optimally parsed iambically, incurring numerous violations of other constraints.

Text	7: *InitS	9: *SinW	4: *Mono-in-Di	5: Disyl-Strong	12: *Disyl	11: FinS
Sonnets	0	.14	.03	.01	.1	.57
Pivotal	0	.11	.02	0	.1	.68
Casual	0	.14	.01	0	.09	1.07

Average non-weighted constraint violations per line.

These results serve to at least tentatively refute the notion that some constraints are heavily relaxed in some genres: For most lines in all three genres, there is a parse available that satisfies all of the crucial constraints.

The only result in the above table that seems potentially significant is the higher proportion of function words in strong positions in the casual text, but this is only a quantitative phenomenon. If there is any variation in the basic parameters of Shakespearean meter, they are not represented by these constraints, or are distributed evenly among these texts.

4 Experiment 2: Larger Samples

The results from the small case study were not especially conclusive, but there are several factors which may be responsible for this. The choice of constraint ranking in the parser is arbitrary, the pronouncing dictionary is incapable of capturing some variation that is important to parsing, and my selection of data is motivated by my subjective perception of register. I attempt to bypass at least some of this by investigating the properties of larger texts in their entireties.

4.1 Data

In this experiment I test the same constraints on three larger datasets, to see if the one generalization noted in the previous section holds, and to seach for more robust variation in the violations of the crucial constraints. Here, I use two very different

full plays, the tragedy *King Lear* and *The Comedy of Errors*, as well as Shakespeare's full corpus of sonnets. I also do not filter as heavily here, leaving in first and last lines of dialog passages, and cutting lines only if they are shorter than 28 characters (and thus very unlikely to be pentameter) or contain out-of-dictionary words.

4.2 Results

Text	7: *InitS	9: *SinW	4: *Mono-in-Di	5: Disyl-Strong	12: *Disyl	11: FinS
Sonnets	0	0.0813	0.0106	0.0258	0.0709	0.614
Lear	0	0.0928	0.0584	0.0606	0.220	0.776
C. of Errors	0	0.115	0.04	0.0344	0.126	0.893

Average non-weighted constraint violations per line.

There are still no clear and systematic differences in metricality between texts, however, two results seem to stand out:

- The trend observed in the previous experiment persists, with function words appearing in strong positions about twice as often in the lower register text (here the comedy) than in the sonnets.
- One new trend appears here: The rate of parses disyllabic metrical positions in *King Lear* is nearly double that of either other text. In all texts, between a quarter and a half of those violate one of the two constraints on such positions.

5 Conclusions

If there are any systematically distinct metrical subgenres within the text I have selected, they either do not follow the divisions I have selected, or they are not easily represented in the simple constraint-based meter I have laid out here.

In any case, there remains some room for further experimentation. An improved dictionary-syllabifier pair could enable more accurate, and hence more precise, comparisons of sub-genres on gradient scales. In addition, a larger range of non-crucial low-weighted violable constraints could allow for the identification of more systematic trends in optimal parses, and help to identify what aspect of a line and of a meter are most salient in composition and parsing.

References

Arto Anttila, Joshua Falk, and Ryan Heuser. Prosodic. software package, 2011.

K. Hanson and P. Kiparsky. A parametric theory of poetic meter. *Language*, 72, 1996.

The Carnegie Mellon University Speech Group. The carnegie mellon pronouncing dictionary, 2007.

The Moby Project and MIT. The complete works of william shakespeare, 1993.