Maxent and meter

A. Overview

(1) a. point of work on meter
   b. previous results
   c. what Hayes et al. do
   d. current results
   e. Welsh

B. Point of work on meter

(2) What do poetic restrictions tell us about the grammar?

(3) “The second coming” — William Butler Yeats
    Turning and turning in the widening gyre
    The falcon cannot hear the falconer;
    Things fall apart; the centre cannot hold;
    Mere anarchy is loosed upon the world,
    The blood-dimmed tide is loosed, and everywhere
    The ceremony of innocence is drowned;
    The best lack all conviction, while the worst
    Are full of passionate intensity.

(4) “The Cremation of Sam McGee” — Robert W. Service
    There are strange things done in the midnight sun
    By the men who moil for gold;
    The Arctic trails have their secret tales
    That would make your blood run cold;
    The Northern Lights have seen queer sights,
    But the queerest they ever did see
    Was that night on the marge of Lake Lebarge
    I cremated Sam McGee.
“Jolene” — Dolly Parton
Jolene, jolene, jolene, jolene
I'm begging of you please don't take my man
Jolene, jolene, jolene, jolene
Please don't take him just because you can
Your beauty is beyond compare
With flaming locks of auburn hair
With ivory skin and eyes of emerald green
Your smile is like a breath of spring
Your voice is soft like summer rain
And I cannot compete with you, jolene

Does art break the rules or does it reveal the rules?

C. Previous results

A stress maximum cannot occur in a weak position (Halle and Keyser, 1971).

Stress maximum: a stressed syllable surrounded by stressless syllables.

<table>
<thead>
<tr>
<th>non-maxima</th>
<th>maxima</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bóoks in...</td>
<td></td>
</tr>
<tr>
<td>Háppy are...</td>
<td></td>
</tr>
<tr>
<td>Fine háts...</td>
<td></td>
</tr>
<tr>
<td>...bíg mén are...</td>
<td>...the mén are...</td>
</tr>
<tr>
<td>...the mén sín...</td>
<td>...the ápples...</td>
</tr>
<tr>
<td>...bíg mén sín...</td>
<td>...complète the...</td>
</tr>
</tbody>
</table>

Positions
- iambic pentameter: $W W S W S W S W S W S S$
- trochaic tetrameter: $S W S W S W S W S$

Prosodic hierarchy (Hayes, 1989): word, clitic group, phonological phrase, intonational phrase, utterance.

- Clitic group formation
  a. Every content word (lexical category) belongs to a separate Clitic Group.
  b. Definition: The host of a Clitic Group is the content word it contains.
  c. Definition: $X$ and $Y$ share category membership in $C$ if $C$ dominates both $X$ and $Y$.
  d. Clitic words are incorporated leftward or rightward into an adjacent Clitic Group. The group selected is the one in which the clitic group shares more category memberships with the host.
(13) Phonological phrase construction. In the configuration \([X\ldots X^0 Y\ldots]\):
   a. The sequence \([\ldots X^0]\) obligatorily occupies the same P-phrase,
   b. \(Y\) may optinally adjoin to the P-phrase of \(X^0\) if it contains only one Clitic Group, and
   c. All Clitic Groups unaffected by rules (a) and (b) form P-phrases.

(14) Phonological phrases in Shakespeare
   [When lofty trées] [I sée] [bárren of léaves]
   *[With lofty birches] [quite bárren of léaves]
   *[When lofty birches] [are bárren of léaves]

D. The theory

(15) Construct lines whose phonological structure evokes the meter.

(16) Use Align constraints to put the metrical structure into conformity with the prosodic hierarchy.

(17) Align(\(PH,MH\))
   \(PH = \{W, CG, PP, IP, U\}\)
   \(MH = \{Foot, Line\}\)

(18) Stress matching
   *Stress in W Avoid stressed syllables in W position
   *Stressless in S Avoid stressless syllables in S position

(19) Sequences
   *Rise from S Do not rise in stress out of an S position
   *Fall from W Do not fall in stress out of a W position
   *No fall from S Stress must fall out of an S position
   *No rise from W Stress must rise out of a W position

(21) . . . and bunches of others.
E. Current results

(22) Scan by hand
Shakespeare’s Sonnets
Books VIII and IX of Milton’s *Paradise Lost*

(23) Calculate violations and weights for each of the 87 constraints in their set; choose
the best constraints.

(24) Which is better: \( \{C_1, \ldots, C_n\} \) or \( \{C_1, \ldots, C_n, C_{n+1}\} \)?

(25) Likelihood Ratio Test from paper (incorrect?)
\[
2 \times \log \left( \frac{\text{Probability of corpus under simpler grammar}}{\text{Probability of corpus under full grammar}} \right)
\]

(26) Calculating this in R: \texttt{pchisq(-2*log(P(simple)/P(full)),1)}

(27) Results
Everything matters except constraints built on stress maxima.

F. Welsh: background

(28) Dafydd ap Gwilym, *Y Gwynt*

(29) Yr wybrwynt, helynt hýlaw,
Agwrdd drwst a gerdda dráw,

Gwr eres wyd garw ei sáin,
Drud byd heb droed heb ádáin.

Uthr yw mor eres y’th réod
O bantí wybr heb úntroed,

A buaned y rhédy
Yr awr hon dros y fron frý.
Sky-wind, unhindered course,
mighty commotion passing yonder,
you are a harsh-sounding minstrel,
world’s fool without foot or wing.
It’s amazing how wondrously you were sent
from the pantry of the sky without any feet,
and how swiftly you run
now across the hilltop on high.

Yr wybrwyt | hêlynt | hyllaw,      sain
Agwrd drêst | a gerdwa drâw,   croes
Gwr eres wyd | garw ei sain,     croes
Drud býd | heb droed heb ádain.  traws
Úth rýw | mor eres y̲ th rôed   traws
O bântri | wybr heb úntroed,   traws
A buaned | y rhédy             llusg
Yr awr hón | dros y̲ frón | frý.         sain

Caesura/gorffwysfa in cynganedd sain
There are two and they can go anywhere in the line.

Caesura/gorffwysfa in cynganedd lusg
There is one and it can go anywhere in the line.

Caesura/gorffwysfa in cynganeddion croes and traws:
There is one. I can go after the first, second, third, or fourth syllables unless the
final word in the first half line is monosyllabic and the final word in the second half
line is polysyllabic. In that case, it may not fall after the fourth syllable.

G. Welsh: maxent

Can maxent discover these or other restrictions on the cywydd form?

Segment coding

<table>
<thead>
<tr>
<th>[syl]</th>
<th>[gorff]</th>
<th>[stress]</th>
</tr>
</thead>
<tbody>
<tr>
<td>u</td>
<td>+</td>
<td>0</td>
</tr>
<tr>
<td>s</td>
<td>+</td>
<td>0</td>
</tr>
<tr>
<td>w</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>g</td>
<td>-</td>
<td>+</td>
</tr>
</tbody>
</table>
Line coding
Yr wybrwynt, helynt hylaw, u w s u g s u g s u
Agwrrdd drwst a gerdda draw, s u w s g u w s u w s
Gŵr eres wyd garw ei sain, s w s u w u g s w u w s
Drud byd heb droed heb adain. s w s g u w s w u w s u
Uthr yw mor eres y’th roed s w u g u w s u w w s u w s
O bantri wybr heb untroed, u w s u g s w u w s u
A buaned y rhedy u w u s u g u w s u
Yr awr hon dros y fron fry. u w w u g u w w s g s

500 lines of cywydd coded by hand from the www.dafyddapgwilym.net website. For some reason, the software requires at least 3000 items, so all counts were multiplied by 10.

I’d actually wanted to code each line with a word boundary on each edge, but for some reason that crashed the software. Those were removed.

Here are the results for 20 constraints, including 4-grams, no projections.

<table>
<thead>
<tr>
<th>Constraint</th>
<th>Language</th>
<th>Count</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>[+word_boundary][-syl]</td>
<td>6.837 coding</td>
</tr>
<tr>
<td>2</td>
<td>[-syl][+word_boundary]</td>
<td>6.795 coding</td>
</tr>
<tr>
<td>3</td>
<td>[-syl][-syl]</td>
<td>7.38 coding</td>
</tr>
<tr>
<td>4</td>
<td>[+str][+str]</td>
<td>4.472 language</td>
</tr>
<tr>
<td>5</td>
<td>[+word_boundary][-str][+gor]</td>
<td>2.996</td>
</tr>
<tr>
<td>6</td>
<td>[+syl][+str][-word_boundary]</td>
<td>5.491</td>
</tr>
<tr>
<td>7</td>
<td>[+word_boundary][-word_boundary][-str][-str]</td>
<td>1.061</td>
</tr>
<tr>
<td>8</td>
<td>[+syl][-str][-str]</td>
<td>2.17 language</td>
</tr>
<tr>
<td>9</td>
<td>[+word_boundary][-str][+syl][+str]</td>
<td>1.023 language</td>
</tr>
<tr>
<td>10</td>
<td>[-word_boundary][+str][-syl][-str]</td>
<td>3.27 language</td>
</tr>
<tr>
<td>11</td>
<td>[-gor][-word_boundary][+str][-gor]</td>
<td>3.27 language</td>
</tr>
<tr>
<td>12</td>
<td>[+word_boundary][-word_boundary][+str][+gor]</td>
<td>2.409</td>
</tr>
<tr>
<td>13</td>
<td>[+word_boundary][-word_boundary][+str][+gor]</td>
<td>2.946</td>
</tr>
<tr>
<td>14</td>
<td>[+syl][-str][+gor]</td>
<td>2.443</td>
</tr>
<tr>
<td>15</td>
<td>[-word_boundary][-str][-str]</td>
<td>2.832</td>
</tr>
<tr>
<td>16</td>
<td>[+word_boundary][-str][-str]</td>
<td>3.163</td>
</tr>
<tr>
<td>17</td>
<td>[+gor][-str][+gor]</td>
<td>3.147</td>
</tr>
<tr>
<td>18</td>
<td>[+syl][-str][-syl]</td>
<td>3.432 language</td>
</tr>
<tr>
<td>19</td>
<td>[+gor][-str][+word_boundary]</td>
<td>2.443</td>
</tr>
<tr>
<td>20</td>
<td>[+syl][+syl][+str]</td>
<td>2.034</td>
</tr>
</tbody>
</table>

Files:
scanned.txt
features.txt
projections.txt
learningdata.txt
testingdata.txt
H. References


Hayes, Bruce, Colin Wilson, and Anne Shisko. In submission. Maxent grammars for the metrics of Shakespeare and Milton.
