

# Supervised Machine Learning for Hybrid Meter

Alex Estes & Christopher Hench

University of California, Berkeley

Department of German

Berkeley, CA 94720, USA

{estes, chench}@berkeley.edu

## Abstract

Following classical antiquity, European poetic meter was complicated by traditions negotiating between the prosodic stress of vernacular dialects and a classical system based on syllable length. Middle High German (MHG) epic poetry found a solution in a hybrid qualitative and quantitative meter. We develop a CRF model to predict the metrical values of syllables in MHG epic verse, achieving an F-score of .894 on 10-fold cross-validated development data (outperforming several baselines) and .904 on held-out testing data. The method used in this paper presents itself as a viable option for other literary traditions, and as a tool for subsequent genre or author analysis.

## 1 Introduction

The divergence of Latin into distinct regional dialects had profound linguistic and literary implications for all of Europe. Even before the Middle Ages, the syllable length of classical Latin had been nearly forgotten in the vernacular.<sup>1</sup> Latin poetry had used quantitative meter, whereby syllable length was the organizing principle. However, the emerging dialects differed from Latin in that stress became a phonologically important feature, and so-called qualitative meter predominated. In order to reconcile these linguistic differences, poetic forms emerged in which meter relied on both stress and syllable length. These hybrid metrical forms pose unique challenges to automated scansion (the pro-

cess of determining the metrical value of each syllable for a line of poetry). In applying machine learning techniques to scan syllables of a hybrid meter, we believe we can contribute to the study of both metrics and poetics, medieval and otherwise. Our system serves not only pedagogical purposes by introducing students to the meter of medieval German epic poetry, but also presents itself as a tool for further research in author identification or topic modeling metrical form.

To illustrate quantitative meter, we consider the epic poetry of Latin and Greek. Each line consists of six feet, each foot typically a dactyl (a long syllable followed by two short syllables) or spondee (two long syllables). A syllable is considered long if it has a long vowel or diphthong, or ends in a consonant (Hayes, 1989). All other syllables are short. The first line of Virgil’s *Aeneid* serves as example:<sup>2</sup>

*ārma vi|rumque ca|nō, Tro|jae quī|prīmus ab|ōrīs*  
— ◡ | — ◡ ◡ | — — | — — | — ◡ ◡ | — —

Shakespeare’s verse, on the other hand, exhibits qualitative meter, structured in iambic pentameter, where each line has five iambs (a bisyllabic foot consisting of an unstressed first syllable followed by a stressed second syllable). The first line of *Romeo and Juliet* is scanned below:<sup>3</sup>

|Two house|holds, both |alike |in dig|nity.|  
| ×    × | ×    × | × × | × × | × × |

## 2 Middle High German Meter

This paper considers the meter of twelfth and thirteenth century Middle High German (MHG) epic

<sup>1</sup>Augustine writes toward the end of the fourth century that while he recognizes time intervals, he can no longer distinguish between long and short syllables: *syllabarum longarum et brevium cognicionem me non habere...* “I cannot recognize long and short syllables...” cf. Augustinus, *De musica*, III, 3, 5.

<sup>2</sup>— represents a long syllable and ◡ a short syllable.

<sup>3</sup>Every syllable is marked with ×. The acute accent ´ marks stress.

verse. Although written in a Germanic language, MHG poetry was greatly influenced by the Romance tradition. This heritage is evident in its hybrid metrical structure: MHG verse patterns according to both syllable stress and length (Bostock, 1947).

The predominating pattern is an alternation between stressed and unstressed syllables (Tervooren, 1997).<sup>4</sup> MHG epic verse employs trochaic tetrameter: each line has four feet (Bostock, 1947), and each foot is a trochee. Phonologically, a trochee consists of two syllables; the first syllable is stressed, and the second is unstressed. For example, the English word “better” is a trochee, but the word “alive” is not. The famous Longfellow epic poem *The Song of Hiawatha* is written in trochaic tetrameter, and the first line serves to illustrate this rhythm:

Should you ask me, whence these stories?  
 | ×      × | ×   × |      ×      × | × × |

Similarly, the typical MHG epic verse foot is two syllables in length, a stressed syllable followed by an unstressed syllable. However, feet can also be filled by one or three syllables (Domanowski et al., 2009). If a foot is filled by one syllable, the syllable must be phonologically long. If the foot is filled by three syllables, either the first two or the last two syllables must both be phonologically short.

It is in these atypical feet that the influence of quantitative meter, where syllable length is the key factor, becomes evident. We must slightly redefine the foot to account for this. Syllable length is measured in morae. Phonologically, a mora is a unit of time such that a short syllable has one mora and a long syllable has two morae (Fox, 2000).<sup>5</sup> A foot in this meter is more precisely defined as having two morae, not necessarily two syllables.<sup>6</sup> Indeed, the mora, not the syllable, has been called the fundamental unit of MHG verse (Tervooren, 1997, p. 1), although the mora functions differently in this po-

etic tradition than in its phonological definition. If a foot has only one syllable, the syllable must be long because a long syllable is two morae and the MHG foot requires two morae. A short syllable cannot be the only syllable in a foot, since it cannot be two morae. If a foot has three syllables, two must be short because only short syllables can be scanned as half morae, together forming one mora.<sup>7</sup> The other syllable is analyzed as one mora, yielding the required two morae in the foot. To summarize, a syllable can have one of three length values: mora, half mora, or double mora. A half mora must be phonologically short, and a double mora must be phonologically long. Phonological length is otherwise irrelevant and any syllable can be one mora.

In addition to length, as a function of morae, syllables are also assigned stress. There are three stress values: primary, secondary, or unstressed. Primary stress is assigned to the first or only stressed syllable in a word. Secondary stress is assigned to any following stressed syllable(s) in that word. All other syllables are unstressed.<sup>8</sup>

The final mora of the final foot of a line is omitted by convention. This is construed as a pause, and receives its own symbol in the scansion, even though there is no corresponding word or syllable. A short, word final syllable may also be elided before a word beginning with a vowel. MHG epic verse permits up to three syllables in anacrusis (a series of syllables at the beginning of a line that do not count in the meter). These syllables may or may not carry lexical or syntactic stress, but they are always scanned as unstressed morae.

The above features yield eight possible metrical values for any syllable:

<sup>4</sup>There is no consensus view on MHG meter. For this work we have most closely followed the viewpoints presented by Domanowski et al. (2009) and Heusler (1956), as well as more explicitly addressed the function of morae.

<sup>5</sup>For example, the English word “red” has two morae since it ends in a consonant, whereas the first syllable in the English word “reduce” has one mora, since it ends in a short vowel.

<sup>6</sup>It can be helpful to think of MHG meter in the musical sense. Each foot is a measure of 2/4 meter, where one mora is equivalent to one quarter note (Bögl, 2006).

<sup>7</sup>Occasionally very weakly stressed long syllables can also count as a half mora.

<sup>8</sup>The metrical distinction between different degrees of stress is rooted in phonological reality (Giegerich, 1985): in a word with many syllables, one syllable usually has a primary stress, and the others have either secondary or weak stress. For example, many pronounce the English word “anecdotal” with secondary stress on the first syllable, primary stress on the third syllable, and weakest stress on the second and fourth syllables.

- **mora - primary stress** (×): a syllable with primary stress
- **mora - secondary stress** (×̇): a syllable with secondary stress
- **mora - unstressed** (×): an unstressed syllable
- **half mora - primary stress** (◌́): a short syllable with primary stress; according to metrical convention the preceding syllable must be long
- **half mora - secondary stress** (◌̇́): a short syllable with secondary stress
- **half mora - unstressed** (◌): a short unstressed syllable
- **double mora** (—): a stressed long syllable; double morae always carry primary stress
- **elision** (e): an elided syllable

Line 1 of Hartmann von Aue's *Der arme Heinrich* is prototypical. Each foot consists of a stressed syllable followed by an unstressed syllable. There is a one-syllable anacrusis:<sup>9</sup>

Ein | rîter | sô ge|lêret | was<sup>10</sup>  
 × | ×̇ × | ×̇ ×|×̇ × | ×̇ ^

Line 6 also begins with one syllable in anacrusis. The second foot has a stressed mora consisting of two syllables, each one a half mora. The third foot has one syllable; a diphthong allows it to be scanned as long. The final foot has a mora with secondary stress, since the preceding syllable is stressed and in the same word:

der |nam im |manege |schou|we<sup>11</sup>  
 × | ×̇ × | ×̇ ◌ ◌ | — | ×̇ ^

Line 34 has no anacrusis, and in the second foot two half mora syllables form the unstressed mora:

|die ein |ritter in |sîner |jugent<sup>12</sup>  
 | ×̇ × | ×̇ ◌ ◌ | ×̇ × | ◌ ◌ ^

Line 8 shows an elided syllable in the second foot:

dar |an be|gundē er |suo|chen<sup>13</sup>  
 × | ×̇ × | ×̇ × | — | ×̇ ^

<sup>9</sup> ^ represents the rest for the empty mora at the end of a line. Note that this notation differs slightly from that which is used for classical and Shakespearian verse.

<sup>10</sup>“There was a knight so learned”

<sup>11</sup>“he looked extensively,”

<sup>12</sup>“which a knight [should have] in his youth.”

<sup>13</sup>“in [these books] he began to search,”

### 3 Previous Computational Approaches to Meter

There are two prevailing treatments of meter in the literature concerned with computational poetic text analysis. One approach takes a known meter and assigns syllables to stress patterns based on such parameters (Hartman, 1996). The second approach assumes nothing of the meter, and seeks to determine it by marking syllables and identifying patterns (Plamondon, 2006; McAleese, 2007; Greene et al., 2010; Agirrezabal et al., 2013; Navarro, 2015). Our approach draws more on the latter.

Previous scholarship has focused on relatively simple systems of meter and adopted rule-based, statistical, or unsupervised approaches. The hybrid nature of MHG meter, and other complex systems developing out of classical antiquity, makes it difficult to scan poetry using these methodologies.<sup>14</sup>

### 4 Data

As supervised machine learning is a novel approach to scansion, annotated metrical data do not exist for MHG or most other languages. Following the scansion categorization system outlined above, the authors annotated syllables of MHG epic poetry into the eight categories of metrical value.

The annotated data consist of 450 lines from Hartmann von Aue's *Der arme Heinrich*, 200 lines from Wolfram von Eschenbach's *Parzival*, and 100 lines from Wirnt von Grafenberg's *Wigalois*.<sup>15</sup> An additional 10% (75 lines of Hartmann von Aue's *Iwein*) was annotated to be held-out for testing, yielding a total of 825 annotated lines. Summary statistics for all annotated data are given in Table 1.

Syllabification was performed prior to annotation. The principles of onset maximization (early formulation in Vennemann (1972)) and sonority sequencing (early formulation in Jespersen (1904)) govern syllabification in many languages, including, to a great degree, MHG. The division of words

<sup>14</sup>A strictly rule-based approach was undertaken by Friedrich Dimpel (2004). While Dimpel's approach is accurate, it is an arduous task, restrictive, and extremely language specific. Moreover, by identifying only the stressed syllables, it does not encompass the full complexity of MHG meter.

<sup>15</sup>Incorporating different poems from different poets accommodates varying styles of writing, but it also introduces more variability, an issue to be addressed in subsequent work.

	mean	std.	min.	max.
char. per line	21.34	3.39	9	32
syll. per line.	7.62	1.04	5	11
words per line	5.01	1.13	1	8
char. per word	4.26	1.96	1	17
syll. per word	1.52	.71	1	7
char. per syll.	2.80	.81	1	7

**Table 1:** Summary statistics for annotated dataset

with only one intervocalic consonant such as *ta-ge* “days” poses no difficulties. Only certain consonant clusters necessitate further information from MHG phonology. For example, the orthographic sequence of a nasal followed by a velar obstruent, although representing simply a velar nasal in modern German, were in fact still two separate phonemes in MHG (Paul et al., 1982). Thus the word *lange* “long” is syllabified as *lan-ge*. Intervocalic affricates can be viewed as either ambisyllabic or biphonemic, for example in *sitzen* “sit”; under both interpretations the first syllable has a coda and the second has an onset. There are also instances where morpheme boundaries interfere with the otherwise normal processes of syllabification. For example, the common MHG suffix *-lich* in *wîplich* “female” results in the syllabification *wîp-lich*, not *wîplich*, despite onset maximization preferring the latter. Accounting for these idiosyncrasies, beyond onset maximization and sonority sequencing, with additional rules resulted in syllabification with an accuracy of 99.4% on the first 1,000 words in Hartmann von Aue’s *Iwein*, yielding a 95% confidence interval of 98.9% to 99.9%.

Annotation was carried out by both authors, who are trained in MHG scansion.<sup>16</sup> In the case that a line exhibits multiple permissible scansions, priority is given to the scansion which best preserves the alternation of stressed and unstressed syllables. If a decision still cannot be made, then stress is given to semantic importance. An additional consideration is the syntactic stress of a particular line. Clearly, such evaluations allow some room for interpretation. Nevertheless, on a sample of 100 lines from the annotated data (739 syllables), the Cohen’s kappa co-

<sup>16</sup>Although neither author is a native speaker of New High German, the two phases of the language and the metrical traditions are sufficiently different that both native and non-native speakers require training in MHG scansion.

		Annotator 2							
Annotator 1		×	×	×	—	˘	˘	e	˘
	×	285	4	0	1	3	0	0	0
	×	0	225	1	0	1	0	0	0
	×	1	2	74	0	2	0	0	0
	—	1	2	0	72	0	0	0	0
	˘	1	0	0	0	36	0	0	0
	˘	0	1	0	0	0	17	0	0
	e	0	0	0	0	0	0	9	0
	˘	0	0	0	0	0	0	0	1

**Table 2:** Inter-annotator agreement confusion matrix

efficient for the inter-annotator agreement is .962. The greatest disagreement for the human annotators was among unstressed and stressed morae, and between unstressed morae and unstressed half morae (Table 2), implying both some stress and some value disagreement.

## 5 Conditional Random Fields Model

Our efforts were focused on constructing a Conditional Random Fields (CRF) model to predict scansion for MHG meter (Lafferty et al., 2001).<sup>17</sup> Features are applied to the focal syllable, as well as to the ten preceding syllables and the ten following syllables of the line, if present.<sup>18</sup> Our resulting MHG features and their motivations are:

- **Position within line:** the last mora of a line is always stressed, and double morae occur most often in the third foot.
- **Length of syllable in characters:** longer syllables are more likely to be stressed. Unstressed prefixes and suffixes tend to be maximally three characters.
- **Syllable characters:** the characters in a syllable can help identify grammatical morphemes that are often unstressed. Slices were taken of the first character, first and second characters, last three characters, last two characters, and last character.
- **Elision:** the last two characters of the previous syllable and the first two characters of the cur-

<sup>17</sup>Future work might consider an LSTM neural network model. The decision to implement a CRF model was predicated on the interpretability of CRF modeling and understanding the primary features for MHG scansion.

<sup>18</sup>The implementation of the CRF model was expedited with the help of Python’s *crfsuite* (Okazaki, 2007).

rent syllable are identified to detect conditions for elision.

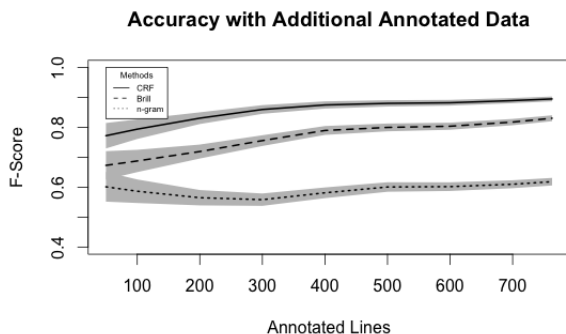
- **Syllable weight and length:** syllables ending in a vowel or consonant are open and closed, respectively. Syllables ending in a short vowel are short; otherwise they are long. Such values are useful in identifying double or half mora syllables, which must be long or short respectively. For example, the syllable “schou” in line 6 of *Der Arme Heinrich* above is a double mora, and is accordingly long.
- **Word boundaries:** stress usually occurs on the first syllable of a word.

The model was tuned only on the development data and the best performing model was chosen. The resulting best model uses an L1 coefficient of 1.3 and L2 coefficient of .001. No further changes were made after the model features and parameters were selected.

## 6 Results and Comparison to Other Models<sup>19</sup>

To evaluate the performance of our CRF model, we compare against two baselines: an n-gram model cascading into regular expressions, and a Brill transformation-based model on top of the n-gram model, both using syllables as units, just as the CRF model does. The n-gram model consists of cascading trigram, bigram, unigram, and regular expressions models, i.e. first a label is predicted based on the previous two labels, if possible; otherwise it is predicted based on the previous one label, and if the first two models fail it is predicted solely based on the label probability for the syllable itself. If the syllable did not appear in the training data, and it cannot be predicted by the first three models, it resorts to regular expressions. Based on MHG scansion theory and observations while annotating, syllables with long vowels were assigned to double mora, short syllables to unstressed mora, and the remaining syllables to mora with primary stress (which proved important to recognize stress alternation). The n-gram model was implemented with default settings.

<sup>19</sup>The results for all models are 10-fold cross-validated.



**Figure 1:** F-score of different models with 95% confidence intervals

The Brill model (Brill, 1995), implemented with the help of NLTK (Bird et al., 2009), first assigns the most likely label from the n-gram model described above, and then generates rules to improve the initial estimate of the n-gram model according to the training data. It then iterates over these rules, correcting labels until accuracy no longer increases. The Brill model was implemented with a maximum of 200 rules.

The n-gram model found little success even with added training data, ending with an accuracy of only 61.8%. The transformation-based Brill model improved quickly upon the n-gram model, but plateaued at 82.8% accuracy. Figure 1 shows the increase in accuracy with an increase in the number of annotated lines for all models, suggesting that marginal returns to annotation begin to diminish significantly after around 400 lines, or, in the case of MHG, about 3,000 syllables. Supervised machine learning thus proves to be an economical option for languages with complex meter.

The final results of the cross-validated CRF model are given in Table 3 in descending order of frequency in the data, along with a final held-out test set of 75 lines from Hartmann von Aue’s *Iwein*. The model achieves an F-score of .894 on the cross-validated development data and .904 on the held-out testing data.

Apart from the infrequent half morae with secondary stress and elisions, the confusion matrix highlights four other problematic situations: (1) stressed morae marked as unstressed morae, (2) unstressed morae marked as stressed morae, (3) dou-

			held-out	
metrical value	F	obs.	F	obs.
mora - unstr.	.922	2403	.938	253
mora - prim.	.918	2025	.947	216
mora - sec.	.856	437	.880	37
double mora	.842	425	.865	34
half mora - unstr.	.574	231	.585	41
half mora - prim.	.771	103	.737	11
elision	.763	65	.500	2
half mora - sec.	0	4	0	0
(macro) average	.894		.904	

**Table 3:** CRF model F-score for individual metrical values and (macro) average in development and on held-out data

ble morae marked as unstressed morae, and (4) unstressed half morae marked as unstressed morae. Three of these issues mirror those of the human annotators. Issues (1) and (2) are attributable not only to the frequency with which these features appear, but also to the pattern of alternating stress. If the stress assignment on one syllable is wrong, then it will likely be wrong for all following syllables in the line. (1) and (2) also affect (3), as shown below in line 14 of *Der arme Heinrich*, where the CRF model labels the penultimate syllable as an unstressed mora, when a double mora is correct:

CRF model: |  $\acute{x}$   $\times$  |  $\acute{x}$   $\times$  |  $\acute{x}$   $\times$  |  $\grave{x}$   $\wedge$   
correct:  $\times$  |  $\acute{x}$   $\times$  |  $\acute{x}$   $\times$  | — |  $\grave{x}$   $\wedge$

This error can be attributed to the overgeneralization of stress alternation. Formally, the model's prediction constitutes a metrically valid line of verse, but it cannot be correct for this particular line because the first syllable in *möhte* must be stressed. The model notably identifies the elision correctly. The worst performance, on unstressed half morae (4), paints a different picture, as illustrated in 21,7 of *Parzival*:

unsern goten, die in uns brähten<sup>21</sup>

CRF model:     $\times \times \mid \acute{\times} \times \mid \times \mid \acute{\times} \times \mid - \mid \acute{\times} \wedge$

correct:        $\times \times \mid \acute{\times} \smile \smile \mid \acute{\times} \times \mid - \mid \acute{\times} \wedge$

The model correctly predicts stress, but does not hold to the constraints of a foot as defined above. While these errors may be common, they are not as

<sup>20</sup>“and with which he might”

<sup>21</sup>“to our gods, who brought him to us”

CRF	Brill
(1) not $\acute{\times}$ if next syll. is end of line	(1) $\acute{\times} \rightarrow \times$ if at word boundary and following syll. is $\acute{\times}$
(2) — if end of line is next syll.	(2) $\acute{\times} \rightarrow \text{—}$ if followed by $\grave{\times}$ and word boundary
(3) $\text{e}$ if last char is “e” and first char. of next syll. is “e”	(3) $\times \rightarrow \grave{\times}$ if end of line
(4) not — if syll. is open and light	(4) $\grave{\times} \rightarrow \acute{\times}$ if monosyllabic
(5) $\times$ if syll in pos. +7 is not end of line	(5) $\times \rightarrow \acute{\times}$ if following syll. is “ge”

**Table 4:** Top five CRF features and Brill rules

severe as the situation depicted in the first example.

The top five highest scoring features of the CRF model and rules of the Brill model are given in Table 4. The CRF feature scores show that syllable quality is helpful (4), something the Brill model cannot generalize. Both models recognize alternation, but while the Brill model adopts a more general rule (1), CRF features (1) and (5) recognize alternation using the number of syllables and the end of the line. CRF feature (5) also accounts for anacrusis. Notably, the Brill model took more advantage of word boundaries in (1) and (2), while these features rank lower in the CRF model. Features for phonemes do not rank high individually, unsurprising considering their impact among the other features. Nevertheless, “ver” and “ge” (MHG prefixes) both score highly in the CRF for marking a syllable unstressed, but only “ge” ranks as a top five rule for the Brill model.

The scores from both models confirm extant MHG metrical theory and suggest new methods of approach for students of MHG meter. Instead of first marking stress, as suggested by *Minimalmetrik* (1997) and the pedagogically oriented website Mittelhochdeutsche Metrik Online (2009), it may be useful to first determine the cadence and anacrusis, as noted in CRF features (1), (2), (5), and Brill rule (3). The advantage of this approach is not mistakenly marking stress in anacrusis. Stress can then be marked in the remaining syllables, and metrical values can be assigned based on phonological features. These results and insights support our feature decisions and our implementation of a CRF model.

## 7 Conclusion

This paper has presented a new application of machine learning models to poetry, specifically to traditions with hybrid meter. It promises to contribute to other literary interests in computational linguistics such as author and genre analysis. Research has shown that a proper account of meter must consider its variation throughout an entire text (Golston, 2009); automated scansion makes this a realistic enterprise. In applying this method, researchers can survey large scale variation within and across texts to discover the patterns that characterize authors and genres. Indeed, subtle differences in meter may prove to be distinct authorial voice or reveal significant stylistic choices. This paper also paves the way for further work such as: cluster analysis of meter over a large corpus of texts, topic modeling cadence across genre, and charting the literary affect of meter.<sup>22</sup>

## Acknowledgments

We are grateful to David Bamman for commenting on drafts, and to the anonymous reviewers for useful feedback. This work was generously supported by Digital Humanities at Berkeley.

## References

- [Agirrezabal et al.2013] Manex Agirrezabal, Bertol Arrieta, Aitzol Astigarraga, and Mans Hulden. 2013. ZeuScansion: a tool for scansion of English poetry. *Finite State Methods and Natural Language Processing*, pages 18–24.
- [Bird et al.2009] Steven Bird, Edward Loper, and Ewan Klein. 2009. *Natural Language Processing with Python*. O'Reilly Media Inc.
- [Bögl2006] Herbert Bögl. 2006. *Abriss der mittelhochdeutschen Metrik: mit einem Übungsteil*. G. Olms, Hildesheim ; New York.
- [Bostock1947] J. Knight Bostock. 1947. *Der Arme Heinrich* by Hartman von Ouwe. Basil Blackwell and Mott, Oxford.
- [Brill1995] Eric Brill. 1995. Transformation-based error-driven learning and natural language processing: A case study in part-of-speech tagging. *Computational linguistics*, 21(4):543–565.
- [Dimpel2004] Friedrich Dimpel. 2004. *Computergestützte textstatistische Untersuchungen an mittelhochdeutschen Texten - Bd. II. Daten und Programme*. A. Francke, Tübingen.
- [Domanowski et al.2009] Anna Domanowski, Yochanan Rauert, Hanno Rüther, and Tomas Tomasek. 2009. Mittelhochdeutsche Metrik Online. <https://www.unimuenster.de/MhdMetrikOnline/>. Accessed on 04-02-2016.
- [Fox2000] Anthony Fox. 2000. *Prosodic Features and Prosodic Structure: The Phonology of Suprasegmentals*. Oxford University Press, New York.
- [Giegerich1985] Heinz J. Giegerich. 1985. *Metrical Phonology and Phonological Structure: German and English*. Cambridge University Press, Cambridge.
- [Golston2009] Chris Golston. 2009. Old English Feet. In *Versatility in Versification: Multidisciplinary Approaches to Metrics*, pages 105–122. Peter Lang.
- [Greene et al.2010] Erica Greene, Tugba Bodrumlu, and Kevin Knight. 2010. Automatic analysis of rhythmic poetry with applications to generation and translation. In *Proceedings of the 2010 conference on empirical methods in natural language processing*, pages 524–533. Association for Computational Linguistics.
- [Hartman1996] Charles O. Hartman. 1996. *Virtual Muse: Experiments in Computer Poetry*. Wesleyan University Press, Hanover, N.H.
- [Hayes1989] Bruce Hayes. 1989. Compensatory Lengthening in Moraic Phonology. *Linguistic Inquiry*, 20:253–306.
- [Jespersen1904] Otto Jespersen. 1904. *Lehrbuch der Phonetik*. Teubner, Leipzig.
- [Lafferty et al.2001] John D. Lafferty, Andrew McCallum, and Fernando C. N. Pereira. 2001. Conditional

<sup>22</sup>The data and source code for this paper are available online at <https://github.com/henhc/CLFL.2016>.

- Random Fields: Probabilistic Models for Segmenting and Labeling Sequence Data. In *Proceedings of the Eighteenth International Conference on Machine Learning*, ICML '01, pages 282–289. Morgan Kaufmann Publishers Inc.
- [Longfellow1932] Henry Wadsworth Longfellow. 1932. *The poems of Longfellow, including Evangeline, The song of Hiawatha, The courtship of Miles Standish, Tales of a wayside inn*. B. A. Cerf, D. S. Klopfer, New York.
- [McAleese2007] Gareth McAleese. 2007. Improving Scansion with Syntax: an Investigation into the Effectiveness of a Syntactic Analysis of Poetry by Computer using Phonological Scansion Theory.
- [Navarro2015] Borja Navarro. 2015. A computational linguistic approach to Spanish Golden Age Sonnets: metrical and semantic aspects. In *Proceedings of the Fourth Workshop on Computational Linguistics for Literature*, pages 105–113, Denver, Colorado, USA, June. Association for Computational Linguistics.
- [Okazaki2007] Naoaki Okazaki. 2007. CRFsuite: a fast implementation of Conditional Random Fields (CRFs). <http://www.chokkan.org/software/crfsuite/>.
- [Paul et al.1982] Hermann Paul, Hugo Moser, Ingeborg Schröbler, and Siegfried Grosse. 1982. *Mittelhochdeutsche Grammatik*. Max Niemeyer, Tübingen.
- [Plamondon2006] M. R. Plamondon. 2006. Virtual Verse Analysis: Analysing Patterns in Poetry. *Literary and Linguistic Computing*, 21:127–141.
- [Shakespeare and Gibbons1980] William Shakespeare and Brian Gibbons. 1980. *Romeo and Juliet*. The Arden edition of the works of William Shakespeare. Methuen, London ; New York.
- [Tervooren1997] Helmut Tervooren. 1997. *Minimalmetrik zur Arbeit mit mittelhochdeutschen Texten*. Kümmerle Verlag, Göppingen.
- [Vennemann1972] Theo Vennemann. 1972. On the Theory of Syllabic Phonology. *Linguistische Berichte*, 18:1–18.
- [von Aue2004] Hartmann von Aue. 2004. *Der arme Heinrich*. Number 6 in Bibliothek des Mittelalters. Deutscher Klassiker Verlag.
- [von Eschenbach et al.1994] Wolfram von Eschenbach, Karl Lachmann, Eberhard Nellmann, and Dieter Kuhn. 1994. *Parzival*. Bibliothek deutscher Klassiker. Deutscher Klassiker Verlag, Frankfurt am Main, 1 aufl edition.
- [von Grafenberg2014] Wirnt von Grafenberg. 2014. *Wigalois: Text der Ausgabe von J.M.N. Kapteyn*. De Gruyter Texte. De Gruyter, Berlin, 2., überarb. aufl edition.