Title
Resonances in Middle High German: New Methodologies in Prosody

Author
Hench, Christopher Leo

Publication Date
2017

Peer reviewed|Thesis/dissertation

Resonances in Middle High German: New Methodologies in Prosody

by

Christopher Leo Hench

A dissertation submitted in partial satisfaction of the

requirements for the degree of

Doctor of Philosophy

in

German

and

in

Medieval Studies

in the

Graduate Division

of the

University of California, Berkeley

Committee in charge:

Professor Niklaus Largier, Chair
Professor David Bamman
Professor Frank Bezner
Professor Anton Kaes

Fall 2017

Resonances in Middle High German: New Methodologies in Prosody

Abstract

Resonances in Middle High German: New Methodologies in Prosody

by

Christopher Leo Hench

Doctor of Philosophy in German
and in
Medieval Studies

University of California, Berkeley

Professor Niklaus Largier, Chair

For years, scholars of medieval German have grappled with how to analyze formal characteristics of the lyric and epic poetry while taking into consideration performance, musicality, *mouvance*, linguistic variation, and aggressive editing practices. The scholarship has justifiably resorted to restricted explorations of specific texts or poets, or heavily criticized region-specific descriptions with several caveats. The many challenges this multifaceted poetry presents has also obscured one of its most central features——the medieval voice. With the little evidence we have often being ambiguous or contradictory, how are we to understand the role of the medieval voice in the German corpus as a whole? This project seeks to shed light on this forgotten aspect by taking advantage of computational methods to demonstrate relative formal and thematic relationships based on sound and voice. In doing so, it presents several new prosodic analytical methods.

The first chapter of this project underlines the importance of sound and voice to medieval performance and composition. It additionally justifies the syllable as the foundation of the novel formal methods presented in the following chapters. Chapter two presents a new syllabification algorithm that combines two linguistic principles: the Sonority Sequencing Principle and the Legality Principle. This algorithm is then customized and computationally implemented to perform accurately on medieval German across dialects and editing practices. Chapter three employs this syllabification algorithm to characterize different phonological soundscapes in reference to theme and voice using a modest corpus of only medieval German lyric poetry. While chapter three intends to quantify the vocal affect of a soundscape, chapter four aims to account for the sequencing of these soundscapes within the larger text and corpus. Chapter three therefore lays the foundation for a structural model of medieval German form. Finally, chapter five reduces the corpus once again to a very small subset of medieval German *Vierheber* (epic poetry with four stresses per line). It presents a supervised machine learning model to predict scansion for these texts according to the theory proposed

by Andreas Heusler, and draws conclusions from how different poets took advantage of the freedom this structure allowed.

In each chapter, I present the aggregate statistics to confirm and supplement our knowledge of the medieval German corpus as a whole. Yet more importantly, I return to individual texts in order to demonstrate how these newly discovered formal soundscapes manifest and function within a smaller narrative. This combination of "distant reading" and "close reading" supports my overarching argument that medieval German form and content were demonstrably and quantifiably highly intertwined. In many cases markedly different formal, phonologically-influenced structures were implemented to trigger connections to related texts and formats. This project thus creates a new understanding of intertextual relations in medieval German literature.

For my family

Contents

List of Figures

List of Tables

Acknowledgments

Above all, I would like to thank my advisor Niklaus Largier and the German Department at UC Berkeley. I have never felt so intellectually free to explore curiosities and interests. I am so thankful for the time I had to pursue any and everything.

I would also like to thank Digital Humanities at Berkeley and the D-Lab for supporting me throughout my time at Berkeley, and for giving me the opportunities that have truly shaped my life. I feel like I finally found that small part of world that excites me the most and the people who share that interest. From the beginnings of my DH explorations, David Bamman has always been happy to talk and guide me.

I am also thankful for my conversations with Frank Bezner, who sparked my passion for the Latin language and poetry. Although my forays were always too brief, it is something to which I am certain I will return.

Since I first arrived at Berkeley, Tony Kaes has constantly reminded me of how to be engaged with your work in all aspects of life. What interests us in the past can always find relevance in the present.

I will always be indebted to the DAAD and the University of Stuttgart, and Professor Manuel Braun and his colleagues for their wonderful hospitality. My conversations with Nils Reiter were especially helpful for the later parts of my work.

Of course, none of this would have been possible without the *Mittelhochdeutsche Begriffsdatenbank* (MHDBDB). Professor Klaus Schmidt was incredibly kind from the beginning, and I was so fortunate to have worked with them.

I also thank my family and friends for their support throughout the program. They managed to keep me sane.

Chapter 1

Voice and Rhythm in the Middle Ages

A heightened interest in the intersection of literature and sensory perception in literary and cultural studies has evolved from an intense rethinking of the body and identity at the turn of the 21st century.[1] With the parallel advances in technology, new studies have attempted to quantify, aggregate, and visualize these previously unmeasurable aspects.[2] Undertaking such a study on the literature of the Middle Ages poses unique challenges to scholars due to the lack of primary source descriptions. The project at hand seeks to revive the voice of the medieval German corpus for a modern reader through computational techniques using only the surviving evidence. This chapter first explains the central role of voice in medieval literature and culture and then justifies using the syllable as the base unit of analysis for investigating this voice.

[1]For the Middle Ages, Mary Carruthers convincingly describes the difference in palate between the Middle Ages and today, while Rachel Fulton Brown reveals how this sense was used to strengthen one's relationship to God.Mary Carruthers, "Sweetness," *Speculum* 81, no. 4 (2006): 999–1013; Rachel Fulton, ""Taste and see that the Lord is sweet" (Ps. 33:9): The Flavor of God in the Monastic West," *The Journal of Religion* 86, no. 2 (April 2006): 169–204 See also Emma Dillon, *The sense of sound: musical meaning in France, 1260-1330* (New York: Oxford University Press, 2012); M. Nolan, "Medieval Sensation and Modern Aesthetics: Aquinas, Adorno, Chaucer," *the minnesota review* 2013, no. 80 (January 1, 2013): 145–158

[2]This is the main objective of the Neuro Humanities Studies project at the University of Cantania. One investigation considers using neuroscience techniques to quantitatively identify aesthetic appreciation for lyric from brain imaging data.Isabel C. Bohrn et al., "When we like what we know – A parametric fMRI analysis of beauty and familiarity," *Brain and Language* 124, no. 1 (January 2013): 1–8 A further noteworthy study is Tanya Clement's measurement and correlation of applause.Tanya Clement and Stephen McLaughlin, "Measured Applause: Toward a Cultural Analysis of Audio Collections," *Journal of Cultural Analytics*, May 23, 2016, `http://culturalanalytics.org/2016/05/measured-applause-toward-a-cultural-analysis-of-audio-collections/` Clement's work proves to be especially relevant here due to the importance she places on the audience's involement in the performance, and the similar aspects of oral performances.

1.1 The Medieval Voice

The voice assumed a central role in Medieval Studies with Paul Zumthor's writings of the 1970s and 1980s and has continued to demand consideration through perspectives in performance studies and music. For Zumthor, studying the voice is based on the understanding that sound is received by an interlocutor before meaning. Zumthor and Hans Ulrich Gumbrecht formulate this similarly: the phone only "prepares" for meaning to be realized and assumes meaning only in a specific environment.[3] Zumthor's *La Poésie et la Voix dans la civilisation médiévale* is one of the most influential studies of voice and orality in the Middle Ages, which elevates the human voice to its own "dimension" within a poetic text.[4] Zumthor argues that despite the intended medium of transmission, every medieval literary text must have been crafted with a vocal component in mind for an audience (be it also an audience of one). Moreover, these poems, due to their performative aspect, inherently keep the text "hidden" from that audience.[5] This is, of course, independent of whether it was composed as a written text or not. Naturally, there existed various types of performance and audiences, as noted by Zumthor (and Thomas Cramer for the German tradition).[6]

Voice and performance are complicated in Middle High German (MHG) literature by subgenre. MHG lyric poetry (broadly categorized into *Minnesang* (love poetry), *Sangspruchdichtung* (didactic strophic songs), *Leich* (spiritual or worldy lais), and epic poetry (both heroic epics and courtly epics) each have varying traditions of performance and relationships to the audience, and thus invoked voice differently. Apart from the textual tradition, audience expectations, social class, size, and gender composition differed.[7] Originally researched in quite different manners, these categories have recently begun to be considered in unison. This issue was addressed for the MHG lyric poetry by a special issue of the *Zeitschrift für deutsche Philologie*.[8]

While Zumthor argues that even those who read these texts undoubtedly articulated the sounds aloud to themselves, the medium of transmission still plays a crucial role in the recognition of other formal elements.[9] Cramer maintains that the debate over oral and

[3]Paul Zumthor and Marilyn C. Engelhardt, "The text and the voice," *New literary history* 16, no. 1 (1984): 69; Hans Ulrich Gumbrecht, Karl Ludwig Pfeiffer, and Monika Elsner, eds., *Materialität der Kommunikation*, 1st ed., Suhrkamp-Taschenbuch Wissenschaft 750 (Frankfurt am Main: Suhrkamp, 1988), 718

[4]Zumthor and Engelhardt, "The text and the voice," 67.

[5]ibid.; Paul Zumthor, *Die Stimme und die Poesie in der mittelalterlichen Gesellschaft* (München: Fink, 1994), 32; Recalling Zumthor's language, Günther Schweikle characterizes oral and sung performances in the medieval German tradition as "tied to the moment".Günther Schweikle, *Minnesang*, vol. 244 (Metzler, 1995), 24

[6]Thomas Cramer, *Waz hilfet âne sinne kunst?: Lyrik im 13. Jahrhundert Studien zu ihrer Ästhetik*, Philologische Studien und Quellen, Heft 148 (Berlin: E. Schmidt, 1998), 22.

[7]Horst Brunner and Helmut Tervooren, "Einleitung: Zur Situation der Sangspruch- und Meistergesangsforschung," *Zeitschrift für deutsche Philologie* 119, no. 2000 (2000): 5; It is important to note that many of the largest manuscript sources for MHG texts do not distinguish, or only slightly distinguish, differing subgenres. ibid.

[8]Ibid.

[9]Zumthor and Engelhardt, "The text and the voice," 68.

written delivery of these texts really only concerns epic texts, because it is clear that the lyric texts were sung. Yet it remains unclear how and to whom they were sung.[10] Cramer contrasts these performances with formal readings. He argues that many of the *Minnesänger* express the opinion that their craft is only for the *wîse*, viz. 'wise', to understand, and that it is very rare that this understanding ever happens.[11] Thus Cramer understands MHG lyric as a multifaceted craft, one that has social value as a performance, but one that also demands an aesthetic appreciation.[12] Cramer also argues against the widely-held notion that *Minnesang* was exclusively read or sung to a large audience, or even to a small group of interested listeners in an exclusive social circle.[13] Cramer rather claims that according to the evidence we have from the German, Latin, and French traditions, there must have been multiple forms of transmission, including: sung in a private audience, sung by a messenger, or read by the intended.[14] Cramer notes that the only Manesse miniature that could be interpreted as performing in front a public audience is Der von Kürenberg, which he suggests may be indicating an archaic inability to write, while many other *Minnesänger* are depicted sending letters containing the songs to their ladies.[15] Acrostic, palindromes, and other wordplay, Cramer argues, simply could not have been discerned by even the greatest listener and must have been intended for a reading audience.[16] Holger Runow presents two cosmological poems by Der Kanzler and Rumelant to demonstrate this complexity. He argues that it would be difficult, if not impossible, to communicate the formal complexity through an oral medium. Even if it were written down and read, Runow asks: "Who understood it? Nobody?"[17] Runow follows Cramer's example by identifying puzzles and wordplay within texts that could only be discovered visually by reading. He concludes that the reception of not only *Minnesang*, but also *Sangspruchdichtung* (especially beginning with Walther von der Vogelweide) must be considered as having had a varied medium.[18] Cramer adds that, unlike the current belief that text *mouvance* was precisely due to the orality of the genre, a written form of composition would have also fostered such *mouvance*.[19] Poets, however, can

[10]Cramer, *Waz hilfet âne sinne kunst?: Lyrik im 13. Jahrhundert Studien zu ihrer Ästhetik*, 9.

[11]Ibid., 10, 16-17.

[12]Ibid., 17.

[13]He justifies this argument primarily with verses from Ulrich von Liechtenstein's *Frauendienst*.ibid., 14

[14]Ibid., 35.

[15]Ibid., 41.

[16]Ibid., 45.

[17]"Wer hat's verstanden? Niemand?" Holger Runow, "Hât ieman sîn sô snellen... Rezeptionsbedingungen des Sangspruchs um 1300 zwischen Mündlichkeit und Schriftlichkeit," in *Sangspruchdichtung um 1300: Akten der Tagung in Basel vom 7. bis 9. November 2013*, ed. Gert Hübner, Dorothea Klein, and Universität Basel, Spolia Berolinensia 33 (Hildesheim: Weidmann, 2015), 96

[18]Ibid.

[19]"The individual versions with their relative linear, even bold messages were suited for oral presentation, while the differentiations and complexity of the problems can only be exploited after consideration of the written version of the presentation." ("Die einzelnen Fassungen mit ihren relativ geradlinigen, manchmal sogar plakativen Botschaften boten sich für den mündlichen Vortrag an, während sich die Differenzierungen und die Komplexität der Probleme erst dem Nachdenken bei der Lektüre der schriftlich fixierten Fassungen erschließt.")Cramer, *Waz hilfet âne sinne kunst?: Lyrik im 13. Jahrhundert Studien zu ihrer Ästhetik*, 98

also take measures to hinder mouvance, e.g. through a strict form pattern, so that a song cannot be adapted without it needing to be rewritten entirely.[20]

As if subgeneres and transmission complexities were not enough, there is also debate over the performance's influence on the lyrical 'I'. In MHG lyric there are two opposing interpretations of the lyrical 'I' with a wide spectrum in between.[21] In the past decades, the scholarship has moved toward the interpretation of the 'I' as a performer, a traveling poet, who creates a 'song-internal' world through 'song-internal' references, while occasionally drawing in 'song-external' references to reality.[22] The extreme interpretation of this would be that the MHG lyrical 'I' is entirely fictional, created for entertainment or more likely for pedagogical purposes. It follows that the women in *Minnesang*, and even the political views expressed in other poetry, cannot be understood as 'real'; we know nearly nothing about the real authors and performers.[23]

Günther Schweikle, while not advocating this extreme stance, does argue that *Minnesang* is a constructed realm of stage play. Schweikle claims that *Minnesänger* assume specific roles.[24] Horst Brunner adopts a similar viewpoint: "The voices, which are spoken from a love song, are rarely or never the voices of the author, rather those of the role crafted by him in a fictional game."[25] Even when speaking self-reflexively, these poets refashion themselves many times over. Sabine Obermaier explains this lucidly in her exposition 'Der Dichter als Handwerker - Der Handwerker als Dichter'.[26]

Schweikle's position has been softened significantly by Jan-Dirk Müller, who, while acknowledging and emphasizing the fictional aspect of *Minnesang*, seeks to reconcile this with the many attempts by the *Minnesänger* to integrate a very real aspect of their life, society, and work into their songs. Müller claims that the lyrical 'I' is complicated severely because

---

All translations in this project are the author's own unless otherwise noted.

[20]Cramer, *Waz hilfet âne sinne kunst?: Lyrik im 13. Jahrhundert Studien zu ihrer Ästhetik*, 110-114.

[21]This debate began with Rainer Warning, "Lyrisches Ich und Öffentlichkeit bei den Trobadors," in *FS Hugo Kuhn (1979)* (1979), 120–159. See also Peter Strohschneider, ""nu sehent, wie der singet!": vom Hervortreten des Sängers im Minnesang," in *"Aufführung" und "Schrift" in Mittelalter und früher Neuzeit* (1996), 7–30, Jan-Dirk Müller, "Die Fiktion höfischer Liebe und die Fiktionalität des Minnesangs," in *Text und Handeln. Zum kommunikativen Ort von Minnesang und antiker Lyrik* (2004), 47–64 and Jan-Dirk Müller, ""Ir sult sprechen willekomen": Sänger, Sprecherrolle und die Anfänge volkssprachlicher Lyrik," in *Minnesang und Literaturtheorie*, ed. Ute von Bloh and Armin Schulz (Tübingen: Niemeyer, 2001), 111-112

[22]Harald Haferland, "Minnesang als Posenrhetorik," in *Text und Handeln: zum kommunikativen Ort von Minnesang und antiker Lyrik*, ed. Albrecht Hausmann, Cornelia Logemann, and Christian Rode, Beihefte zum Euphorion 46 (Heidelberg: Winter, 2004), 65; Schweikle, *Minnesang*, 92; Maria Dobozy, *Re-membering the present: the medieval German poet-minstrel in cultural context*, Disputatio 6 (Turnhout, Belgium : Abingdon: Brepols; Marston [distributor], 2005)

[23]Naturally, the configuration of the *Minnesang* lyrical 'I' is very different from that of the *Sangspruchdichtung*.

[24]Schweikle, *Minnesang*, 92.

[25]"Die Stimmen, die aus dem Minnelied sprechen, sind selten oder nie die Stimme des Autors, sondern von ihm gestaltete Rollen in einem fiktionalen Spiel." Brunner and Tervooren, "Einleitung: Zur Situation der Sangspruch- und Meistergesangsforschung," 7

[26]Sabine Obermaier, "Der Dichter als Handwerker - Der Handwerker als Dichter. Autorkonzepte zwischen Sangspruchdichtung und Meistersang.," *Zeitschrift für deutsche Philologie* 119, no. 2000 (2000): 59–72.

of the performative aspect of the poetry, especially when the poets name themselves as the artist and composer of the song. This aspect introduces, and assumes, some element of extratextual reality that is absent in written form—the audience must first assume that the 'I' is the 'I' standing before them, and must be convinced otherwise.[27] This is true not only for self-reflexive songs, but also to add credibility in the cases of pedagogical verse (*Lehrdichtung*). Müller argues that such elements of reality, including self-referential elements, are necessary to convince the audience. Müller explains the several levels of the lyrical 'I' that exist in any given song, and that to interpret this 'I' the relation of each to reality and fiction must be considered. One great difference in *Minnesang* to other forms of poetry is that the poets commonly represent not a single fictional or non-fictional person, but an entire community (ladies, knights, singers, etc.). These roles often collapse into one through an abundance of references.[28] Müller argues for an inclusive interpration of this 'I' as an individual, an author, a singer, and community, also emphasizing the agenda behind the 'ich'. Much of *Minnesang* reveals apparent agendas behind a single, or several, *Minnesänger*, which for purposes of convincing the audience, needs to alter reality.[29] Müller explains that *Minnesang* has always been a referential phenomenon, not hinting at reality, but at what reality could or should be.[30]

Intensifying some of Müller's claims, and revising some of his previous work, Harald Haferland agrees with Dirk-Müller that fictionality in *Minnesang* is not what we consider fiction today.[31] Yet Haferland maintains that the fiction of *Minnesang* is actually much closer to reality. Haferland argues that the effect of the songs would be much stronger if the poet could relate situations directly to the audience without a "song-internal" 'I': "If Reinmar sang 'I love ...', then the audience was certainly not prepared to allege that he was only performing a temporary textual state beyond his physical self by performing the textual 'I'."[32] Only in actual *Wechsellieder*[33], when Reinmar impersonates a lady, is it clear that Reinmar is switching roles. Haferland takes this claim one step further and asks whether Reinmar is not simply placing himself in a hypothetical situation. In such a case, Reinmar remains Reinmar, albeit now Reinmar as a lady.

In accordance with the authenticity of *Minnesang*, Haferland emphasizes the true existence of the 'ladies'. It is possible that the *Minnesänger* were referencing real, or semi-real, situtations and ladies, embellishing only as much as to not arouse suspicion from the audience. The lyrical 'I' thus only became a communal 'I' in Müller's sense through a process

[27]Müller, ""Ir sult sprechen willekomen": Sänger, Sprecherrolle und die Anfänge volkssprachlicher Lyrik," 109.

[28]Ibid., 110, 127.

[29]Ibid., 127.

[30]Ibid., 113.

[31]See Harald Haferland, *Hohe Minne: zur Beschreibung der Minnekanzone*, Beihefte zur Zeitschrift für deutsche Philologie 10 (Berlin: Erich Schmidt Verlag, 2000)

[32]"Sang Reinmar 'ich minne [...]', so war das Publikum offensichtlich nicht darauf vorbereitet, zu unterstellen, er fülle nur vorübergehend eine Textinstanz über seine korperliche Präsens aus, indem er das textuelle 'ich' performierte." Haferland, "Minnesang als Posenrhetorik," 93

[33]Songs clearly switching between male and female voices.

of folklorization, according to Haferland, the summation of which leaves the song unrecognizable to the poet.[34] Haferland does not believe that poets attempted to craft an 'I' for themselves throughout their entire œuvre, but rather, following Romance models, imagined and staged themselves as role models in matters of love, which they fostered over a longer period of time.[35]

No matter how we wish to interpret the 'ich minne', there is some subject 'ich' that is loving. To get at the nuance we must examine all levels of what this 'ich' meant in *Minnesang*.[36] Whether the 'ich' is a direct reference to the poet himself, or only to the assumed song-internal role, in both cases he offers his words with an agenda in mind. *Minnesang* only assumes its artful practice when this 'I' steps away from these references.[37] While Haferland's argument does represent the more extreme viewpoint and is not without fault, his emphasis on the *reception* of the poetry and not the author is important to this project. What did this performance mean for the audience (even an audience of one) and what were they intended to believe? Is this aspect more important than what the author was actually intending or *doing*? How can the modern reader identify with this audience through the voice of the reader or performer?

When Reinmar informs us that other *Minnesänger* and nobles accuse him of just "posing" and exaggerating, they can only do this if they already accepted the assumption that most *Minnesänger* were serious about what they sung.[38] As I discuss sound and voice in the context of reception in Chapter 3, and recalling Zumthor and Gumbrecht's emphasis that sound arrives before meaning, I will establish a connection between reception, the audience's auditory experience, and meaning.

Different lyrical 'I's, performance, and authenticity add to this climate of deceit surrounding verse in the Middle Ages. The idea that the voice can lie, but the eyes cannot, is well supported in the scholarship and contemporary literature. Horst Wenzel understands speaking and writing as crucial to a medieval bilateral mode of perception, though he argues that there is a preference for sight.[39] Hearing is often contrasted with seeing and denoted as the more unreliable sense.[40] One must *see* for oneself if what was *heard* is true. In the Middle Ages, if not also today, deceit was more easily achieved by speech than by sight. Yet this precisely underlines the importance of speech and sound, and provides further motivation for this project. The voice has the power to deceive and convince through painting pictures, which may not conform to reality.

[34]Haferland, "Minnesang als Posenrhetorik," 71.

[35]Ibid., 72-73.

[36]Ibid., 93.

[37]Müller, ""Ir sult sprechen willekomen": Sänger, Sprecherrolle und die Anfänge volkssprachlicher Lyrik," 113.

[38]Haferland, "Minnesang als Posenrhetorik," 96.

[39]Horst Wenzel, *Hören und Sehen, Schrift und Bild: Kultur und Gedächtnis im Mittelalter*, 1st ed. (München: C.H.Beck, February 27, 1995), 54, 57.

[40]Cramer also discusses the idea of obfuscating lies in language, though not necessarily with the oral and auditory components Wenzel does.Cramer, *Waz hilfet âne sinne kunst?: Lyrik im 13. Jahrhundert Studien zu ihrer Ästhetik*, 181

It was through speech, after all, that most people understood their religion. The *Welsche Gast* tells us how the scripture was mediated through the priest, by seeing, to the people, by hearing:

> Der pfaffe sehe die schrift an,
> sô sol der ungelêrte man
> diu bilde sehen, sît im niht
> diu schrift zerkennen geschiht.[41]

These central qualities of truth and validation depended on voice, as Zumthor argues, and rhythm, as Mertens argues for *Minnesang*.[42] Rhyme and rhythm together can be restrictive or prescriptive forces acting on a medieval poet. They can force poets to write a verse slightly different than the original thought in order to fill or fit the meter, or complement the rhyme. Thus the selected words were not only selected for sematic suitability, but for formal effect. They may not communicate the full truth. These restraints, combined with the need for validation, make it particularly difficult for the performance-driven lyric poets more than the epic poets.[43] In a similar vein, Zumthor argues that if sound play through formal elements is too dense, as highly formal later *Minnesang* often exemplifies, it can result in a purely auditory experience of an entirely new meaning. This experience strives to free itself from the language entirely, thus obscuring the meaning of the language used.[44] In fact, the loss of faith in the authenticity of lyric is Rüdiger Schnell's argument for the rise of prose writing in the later Middle Ages.[45] As Heinrich von Morungen illustrates:

> Wolte sî mîn denken vür daz sprechen
> und mîn trûren vür die klage verstân,
> sô müese in der niuwen rede gebrechen.
> owê, daz iemen sol vür vuoge hân,
> Daz er sêre klage,
> daz er doch von herzen niht meinet,
> alse einer trûret unde weinet

[41] Thomasin and F. W. von Kries, *Der Welsche Gast*, Göppinger Arbeiten zur Germanistik 425 (Göppingen: Kümmerle-Verlag, 1984), ll. 1103-1106. "The priest reads the scripture, so should the unlearned man see the images, since he does not manage to recognize the words."

[42] Zumthor, *Die Stimme und die Poesie in der mittelalterlichen Gesellschaft*, 32; Volker Mertens, "Was ist Rhythmus in Minnesang?," in *Aus dem Takt: Rhythmus in Kunst, Kultur und Natur* (Bielefeld: Transcript, 2005), 194

[43] Cramer, *Waz hilfet âne sinne kunst?: Lyrik im 13. Jahrhundert Studien zu ihrer Ästhetik*, 180.

[44] Zumthor, *Die Stimme und die Poesie in der mittelalterlichen Gesellschaft*, 74.

[45] Rüdiger Schnell, "Prosaauflösung und Geschichtsschreibung im deutschen Spätmittelalter," in *Literatur und Laienbildung im Spätmittelalter und in der Reformationszeit: Symposion Wolfenbüttel 1981*, ed. Ludger Grenzmann, Karl Stackmann, and Deutsche Forschungsgemeinschaft, Germanistische Symposien, Berichtsbände 5 (Stuttgart: J.B. Metzler, 1984), 217; To justify this claim, he points to the Phillip-August *Prosachronik*, which reasons its prose form in the style of Lancelot so that the author must not be "forced to lie", as all rhyming literature must.ibid., 218

unde er sîn niemen kan gesagen.[46]

Despite voice, sound, and rhythm presenting themselves as promising areas of research, Zumthor questions what has hindered scholarship on the medieval voice. He argues that it is not due to the little knowledge of how it actually sounded, but rather simply the wrong approach.[47] Zumthor sees the greatest problem being that the character of the human voice is dependent on the individual and that the "force" of a poem is derived from the unique contact between the performer and the audience.[48] Zumthor is most concerned with the physical aspect of voice, and how its presence becomes an "object of perception."[49] This voice, therefore, is a part of medieval poetic form (foremost it is sound play), and in this respect the text fits itself to the voice, not vice-versa.[50]

Unfortunately, Zumthor does not believe that we can garner anything about the actual voice from the text, rather only, as he does, emphasize the extent to which medieval literature was oral. Anyone who believes themselves able to pinpoint the "living language" from a manuscript enters what Zumthor labels "pseudo-scientific speculation."[51] Recent scholarship on medieval German has disagreed. Markus Stock argues that with MHG texts we are not so concerned with orality (*oralité*), because according to Zumthor orality is the vocal communication of culture because *there is no feasible alternative*. Yet with MHG we are more concerned with *vocalité*, showing that a vocal presentation was chosen and preferred over a written contribution. These poets *could have* written, and the elite *could have* read them, but that was not the preferred medium.[52] Furthermore, while Zumthor believes these manuscripts are empty because they are no longer full with spoken words, Stock points out that this would leave medievalists with nothing to study. While we may not be able to get at the true or full picture, "traces" of this picture can be found.[53]

[46] *MF* XIb 3. "If she wishes to understand my thinking as speaking and my pain as lamentation, then she would have to forgo this speech. Woe! That one considers it artful when someone laments about something that doesn't come from the heart, just as the one who grieves and cries and cannot tell anyone about it."

[47] Zumthor and Engelhardt, "The text and the voice," 68.

[48] Ibid., 73, 75.

[49] Zumthor, *Die Stimme und die Poesie in der mittelalterlichen Gesellschaft*, 13.

[50] Ibid., 35-36.

[51] Ibid., 86.

[52] Markus Stock, "Das volle Wort - Sprachklang im späteren Minnesang," in *Text und Handeln: zum kommunikativen Ort von Minnesang und antiker Lyrik*, ed. Albrecht Hausmann, Cornelia Logemann, and Christian Rode, Beihefte zum Euphorion 46 (Heidelberg: Winter, 2004), 193.

[53] ibid., 193-194; Andreas Heusler similarly writes that the musical evidence we have of neumes are not capable of depicting rhythm, but it can lead us closer. ("Die Buchstabenschrift hat keine Mittel, die rhythmischen Werte abzubilden; sie führt uns bis auf einigen Abstand an diese Werte hinan.") Andreas Heusler, *Deutsche Versgeschichte: mit Einschluss des altenglischen und altnordischen Stabreimverses*, 2nd ed., vol. 2, Grundriss der germanischen Philologie (Berlin: W. De Gruyter, 1956), 13.

1.2 Rhythm and Memory

While voice is at the origin of these medieval performances, rhythm contributes an additional corporeal aspect of sound. In his essay "Rhythmus und Sinn", Hans Ulrich Gumbrecht discusses the connection between rhythm and the body, and suggests adding memory to this constellation.[54] Gumbrecht seeks to attach a similar importance to rhythm as Zumthor did for the voice,[55] proposing three poetological functions of rhythm: 1) the mnemonic function (*die gedächtnisstützende Funktion*), 2) the affective function (*die affektive Funktion*), and 3) the coordinating function (*die koordinierende Funktion*).[56] Aside from the first function, these describe primarily physical relations to rhythm, and how these relations can be "coordinated" among a group of individuals. Gumbrecht defines rhythm as "the achievement of form under the (complicating) conditions of temporality."[57] According to this definition, rhythm can only begin when a sequence of sounds is produced (the "complicating" condition), which addresses the Husserlian issue of ephemeral, temporal objects (*Zeitobjekte*).[58] In a spoken language, this is the "return of (arbitrary) sequences of accompanying sound qualities."[59] Gumbrecht does not go into further detail concerning these "sound qualities", but undoubtedly means phonetic sequences, similar to Zumthor's characterization of the voice. Gumbrecht continues to describe two planes of "consensual order".[60] The higher order describes a further abstraction from the material. Gumbrecht uses language as an example, which requires an interaction (two subject interlocutors), and creates observers of the event. In the case of MHG poetry, this would consist of the audience, however large or of whatever consistency, and the poet, whether this is identified as the lyrical 'I' or someone else. The higher order always creates new elements by assigning semantic meaning. The lower order is more material, and does not generate new elements; no semantic meaning is assigned, and observers are not created in the lower order. Gumbrecht's example for this lower order are the human organs and rhythm. Rhythm, here sequences of sound, do not generate any semantic meaning inherently, a claim in agreement with Saussure's famous declaration that the sounds of words are arbitrary, and consistent with Zumthor's assertion that sound paves the way for meaning, but is not meaning itself.[61] The power of verse is thus the *combination of the higher and lower orders*, or the material and the abstract. Hence the German phrase *gebundene Sprache* (bound language). Gumbrecht claims that these three functions are produced by the different way each order negotiates them.[62] Thus prosody can be one of the

[54] Hans Ulrich Gumbrecht, "Rhythmus und Sinn," in *Materialität der Kommunikation*, 1st ed., ed. Hans Ulrich Gumbrecht, Karl Ludwig Pfeiffer, and Monika Elsner, Suhrkamp-Taschenbuch Wissenschaft 750 (Frankfurt am Main: Suhrkamp, 1988), 714–729.

[55] Zumthor and Engelhardt, "The text and the voice."

[56] Gumbrecht, "Rhythmus und Sinn," 716–717.

[57] "das Gelingen von Form unter der (erschwerenden) Bedingung von Zeitlichkeit." ibid., 717

[58] Ibid., 718.

[59] "[die] Wiederkehr von (beliebigen) Sequenzen begleitender Laut-Qualitäten." ibid., 719

[60] *konsensuelle Bereiche* ibid., 724–726

[61] Ferdinand de Saussure et al., *Course in general linguistics* (LaSalle, Ill: Open Court, 1986).

[62] Gumbrecht, "Rhythmus und Sinn," 725–726.

highest levels of semantic meaning in the higher order, but still be tied to the lower order via the inherent materiality in the rhythm. Gumbrecht describes the connection between sound, rhythm, form, and memory:

> On the more complex level of the verse (constituted of feet) and—more clearly—on the level of stanzas, it already requires a non-random act of memory and anticipation to experience a verse or stanza form. Rhyme (at the end of a verse) and switching rhymes (between the stanzas) likely function as singals, which trigger such acts of memory and anticipation and thus make possible the identification of forms.[63]

While rhyme and cadence aid memory as markers of form, in Gumbrecht's view, rhythm actually helps *reduce* the information required to remember a sequence:

> When one wishes to remember a sequence of language not shaped by rhythm, then it is only possible in a polythetic manner, namely only by bringing to mind the successive and individual sounds, words, and sentences of the utterance to be remembered. When remembering rhythmically formed language, however, one has the "opportunity to transform the many strands of the conscious mind into one." The rhythmical pattern, which gives specific form to a linguistic utterance, can then—quasi 'metonymically'—stand for the complexity that unfolds in the primary temporality. And the rhythm that was remembered offers a form for the mneumonic reproduction of any linguistic sequence, which drastically reduces the amount of syllables, words, and sentences, from which respectively the reproducing parts of the whole can be selected.[64]

It is much easier to remember sequences of sound *quality* rather than the actual sounds or words themselves. We work to *reduce* the amount of information we must process or remember, so instead of remembering entire words or stanzas, we remember the rhythm first. This

[63] "Auf der komplexeren Ebene des (aus mehreren Versfüßen konstituierten) Verses und—noch deutlicher—auf der Ebene von Strophen bedarf es offenbar schon nicht-willkürlicher Akte der Erinnerung und der Vorwegnahme, um eine Vers- oder Strophen-Form zu erleben. Vermutlich wirken der Reim (am Versende) und der Reimwechsel (zwischen den Strophen) als Signale, welche solche Akte der Erinnerung und der Antizipation auslösen und mithin die Identifikation von Formen ermöglichen." Gumbrecht, "Rhythmus und Sinn," 719

[64] "Wenn man sich an eine Sequenz nicht rhythmisch geformter Sprache erinnern will, so ist dies allein in polythetischer Weise möglich, nämlich nur dadurch, daß man sich sukzessiv die einzelnen Laute, Wörter, Sätze der zu erinnernden Aüßerung vergegenwärtigt. Beim Erinnern rhythmisch geformter Sprache jedoch hat man die "Möglichkeit, das vielstrangig Bewußte in ein schlicht in einem Strang Bewußtes zu verwandeln." Das rhythmische Muster, welches der sprachlichen Äußerung ihre spezifische Form gibt, kann dann—sozusagen 'metonymisch'—für deren primär in Zeitlichkeit entfaltete Komplexität stehen. Und der erinnerte Rhythmus gibt für die erinnernde Reproduktion jener Sprachsequenz eine Gestalt vor, welche die Menge der Silben, Wörter und Sätze, aus denen jeweils zu reproduzierende Teil-Einheiten selegiert werden können, drastisch reduziert." ibid., 720.

is why many people can begin to hum a song before they can recall the words.[65] Both Mertens and Gumbrecht suggest that this lower order of rhythm, which is inferred by the audience but not given verbatim by the performer, is eventually stored in memory and recalled in subsequent performances.[66] They further argue that this lower order is easier to remember than the actual words, and thus is acquired faster. Much of early literary form and poetry contributed to enhancing mnemonic capabilities. This mnemonic aspect functioned on two levels: (1) it allowed for continual transmission of the stories, i.e., facilitated performance by the speaker, and (2) it allowed for a story's solidification in the collective memory, i.e., on the end of the receivers.

Collective memory and communicative and oral practices are discussed by Jan Assmann, who emphasizes in a broader context one of Deutoronomy's eight methods of strengthening collective memory—the power of keeping stories in memory through the mouth.[67] This mnemonic function was especially important for the religiously devout to remember prayers or the praises of God. As Wenzel writes, the lay people "read" with their ears; encounters with text were primarily auditory.[68] Rhyme and rhythm propelled the energy of these encounters. Wenzel describes this as the "pulse of life," going so far as to argue that church music, through this rhythmical singing, changed the individual's feeling of time.[69]

Formulaic Theory

Zumthor and others also refer to a heightened element of repetition in medieval texts. This repetition stems from the oral component, which draws from other stories, but also calls upon the memory function. However, Zumthor argues that "formulas" and other markers of repetition are not exclusively linked to the oral tradition.[70] It has been shown that especially in the Homeric tradition, artificial forms were created in the language to retain the meter and aid memory.[71] The larger debate among medievalists, however, concerns formulaicity, also originally theorized for the Homeric tradition. Zumthor describes the cycle of a work as production, transmission, reception, conservation, and repetition.[72] Formulaicity highlights

[65]Christoph März, "Metrik, eine Wissenschaft zwischen Zählen und Schwärmen?," in *Mittelalter: neue Wege durch einen alten Kontinent*, ed. Jan-Dirk Müller and Horst Wenzel (Stuttgart: Hirzel, 1999), 325.

[66]Mertens, "Was ist Rhythmus in Minnesang?," 186.

[67]Jan Assmann, "Die Katastrophe des Vergessens. Das Deuteronomium als Paradigma kultureller Mnemotechnik," 1994,

[68]Wenzel, *Hören und Sehen, Schrift und Bild: Kultur und Gedächtnis im Mittelalter*, 380.

[69]Ibid., 93, 106.

[70]Zumthor, *Die Stimme und die Poesie in der mittelalterlichen Gesellschaft*, 66; See also Michael Curschmann, "The Concept of the Oral Formula as an Impediment to Our Understanding of Medieval Oral Poetry," *Medievalia et Humanistica* 8 (1977): 63–76

[71]See Joh. Ernst Ellendt, *Ueber den Einfluss des Metrums auf Wortbildung und Wortverbindung bei Homer.* (Königsberg: Druck der Universitäts-Buch- und Steindruckerei von E.J. Dalkowski, 1861); Heinrich Düntzer, *Homerische Abhandlungen von Heinrich Düntzer*, ed. National Central Library of Florence (Hahn, 1872); Milman Parry and Adam Milman, *The making of Homeric verse: the collected papers of Milman Parry* (Oxford: Clarendon Press, 1971)

[72]Zumthor, *Die Stimme und die Poesie in der mittelalterlichen Gesellschaft*, 36.

the importance of each of these stages.

The formulaic theory was first presented by Milman Parry and refined by his student Albert Lord.[73] Parry's original aim is to sift out originalities of Homeric diction through identifying formulas in Homeric epics. He argues that ancient poets had a large repertoire of syntactic phrases fitting an array of metrical patterns, which could be plugged in when necessary to preserve the meter and express the correct sense of the story.[74] Parry focuses on formulas of nouns and adjectives, as well as the indirect speech formula, e.g. τὸν δ' ἠμείβετ'ἔπειτα ("then he replied"). Parry is able to show that both the Greek and to a lesser extent Latin traditions are formulaic to a high degree, exhibiting their oral quality and declaring them as belonging to the oral tradition. While Parry does not venture into medieval territory, Lord refines many of Parry's claims and applies them in a more detailed manner to the Homeric and even medieval epics. At the conclusion of *The Singer of Tales*, Lord suggests that much of this theory can be applied to (early) medieval texts such as *Beowulf* and the *Chanson de Roland*.[75] Lord also suggests the early Germanic epic *Das Nibelungenlied* as a possible candidate for application of the formulaic theory.[76] But Lord then refocuses the issue on whether medievalists are working with "oral" or "written" products as mutually exclusive domains.[77] This idea is perpetuated for some time, until Michael Curshmann, in a response to the accusation that German medieval Germanists had not sufficiently engaged with the scholarship on orality, discusses "The Concept of the Oral Formula as an Impediment to Our Understanding of Medieval Oral Poetry."[78] Curschmann poses a central question to medieval studies: "whether it is legitimate at all to apply a theory developed pragmatically in the field of a living tradition to medieval literary production."[79]

Despite Curschmann's clearly formulated concern, some investigations were undertaken in the medieval realm both before and after his article. Edward Haymes was one of the first to engage with oral formulas in MHG. Haymes' dissertation "Mündliches Epos in mittelhochdeutscher Zeit" undertakes a quantitative analysis of formulas in the *Nibelungenlied* and other Dietrich epics. He compares their use of formulas to evidently more literary texts such as Gottfried's *Tristan*.[80] Curschmann importantly incorporates metrical aspects

[73]Parry and Milman, *The making of Homeric verse: the collected papers of Milman Parry*; Albert Bates Lord, Stephen A. Mitchell, and Gregory Nagy, *The singer of tales*, 2nd ed., Harvard studies in comparative literature 24 (Cambridge, Mass: Harvard University Press, 2000).

[74]This argument simultaneously supports the now widespeard belief that the Homeric epics could not have been authored by a single poet.Parry and Milman, *The making of Homeric verse: the collected papers of Milman Parry*, 12, 17-18

[75]Lord, Mitchell, and Nagy, *The singer of tales*, 198-220.

[76]Ibid., 206.

[77]Ibid., 220.

[78]Curschmann, "The Concept of the Oral Formula as an Impediment to Our Understanding of Medieval Oral Poetry."

[79]ibid., 64 Curschmann is also concerned by the use of computers and statistics in making claims about a tradition's orality. At what point is a text considered to be from an oral tradition, and can a computer metric truly decide this accurately? ibid., 67, 69

[80]Edward R. Haymes, *Mündliches Epos in mittelhochdeutscher Zeit*, vol. 164, Göppinger Arbeiten zur Germanistik (Göppingen, 1975).

into this analysis, noting the freer restrictions of the *Vierheber* (four stresses per verse) in contrast to the more restricted and formula-conducive structure of the *Nibelungenstrophe*.

Werner Schwarz studies specific noun-adjective phrases in the original tradition of Parry, noting that in MHG formulas are often used to describe "typical situations".[81] Due to the less strict meter of MHG (compared to classical hexameter), Schwarz places less emphasis on the meter's importance in formulas, and emphasizes that changing the pattern is more common and workable in MHG because it is not strictly "a rigid pattern of words".[82] Using the formula "liep als der lip" ("dear as the body/life"), Schwarz ventures to show not whether MHG was strictly oral or written, but rather how these two media *interacted*.[83] Whether or not the MHG poets employed oral formulaic techniques, they must have been at least aware of them from classical poetry. Schwarz traces the evolution of the meaning of "liep als der lip," beginning in the more oral form of the *Nibelungenlied* and the Dietrich epics into its use in the courtly epics.

Franz Bäuml and Agnes Bruno present a computational study, in which through supervised and unsupervised classification they classify 41 stanzas from the *Nibelungenlied* as either highly formulaic (62% and above) or less formulaic (50% and below).[84]. While their study was successful, it still does not satisfactorily address the question of formal density. Their statistics show more nouns, adjectives, and adverbs in more formulaic stanzas, and more verbs in less formulaic stanzas; most of their findings are consistent with the strict verse structure.[85] Their model confirms the investigations of Parry and Lord, in that most formulaic language is associated with descriptive, noun-adjective relationships. Bäuml later argues that while the formulaic theory cannot confidently establish the oral transmission of these texts, it can reveal characteristics they share with texts we know were orally transmitted.[86]

Whether or not one subscribes to the formulaic theory for medieval literature, or the extent to which one does, it further supports the link between memory and form as explicated by Gumbrecht. The musicologist Leo Treitler similarly emphasizes the underlying patterning of form and its connection to memory. Treitler argues that form can be repeated without being repetitive, e.g, the repetition of word formulas was conducted at a much higher rate in antiquity than in the Middle ages, but the repeated rhythmical *pattern* is still central to both.[87] For this reason, the emphasis should move from lexical, and even syntactical patterning, to rhythmical and metrical patterning: "Form, as well as salient detail, is persistent

[81]Werner Schwarz, "Notes on formulaic expressions in MHG poetry," in *Beiträge zur mittelalterlichen Literatur* (1984), 60.

[82]Ibid., 60, 68.

[83]Ibid., 61.

[84]For their measure at least half of a line must be deemed formulaic. Franz Heinrich Bäuml and Agnes M. Bruno, "Weiteres zur mündlichen Überlieferung des Nibelungenliedes," *Deutsche Vierteljahrsschrift für Literaturwissenschaft und Geistesgeschichte* 46 (1972): 489, 490

[85]Ibid., 492.

[86]Franz H. Bäuml, "The Oral Tradition and Middle High German Literature," *Oral Tradition* 1 (1986): 416.

[87]Leo Treitler, "Homer and Gregory: the transmission of epic poetry and plainchant," *The Musical Quarterly* 60, no. 3 (1974): 355-356.

and is therefore an important factor in what makes remembering possible."[88]

Mertens addresses *Minnesang*'s rhythm in a more musicological setting. He acknowledges its specific problems, notably the relatively little musical notation extant and the ambiguous neumes. Like Gumbrecht, Mertens understands rhythm as separate from meter, but places it in a performative and temporal, not material, model.[89] He claims that the audience "perceives the meter from the rhythmical (performed) impulses," which corresponds to Gumbrecht's distinction of the lower order.[90] Neumes, however, do not provide specific instructions as to performative aspects of a song, rather more the material Gumbrechtian relationship between notes: "neumes represent music as a physical-vocal realization of word and melody."[91] Mertens believes that the neumes and most indicators of structure simply served to remind the performer of a specific musical meter or melody, not to give specific directions of how it was to be performed. Thus any given performer interpreted the vague directions differently. Unfortunately, not only does *Minnesang* provide scant musical notation, but even less description of its musical tradition. It is not even clear whether or not instruments were employed or songs were accompanied by dance, which would strongly suggest a tighter rhythmical pattern. Nevertheless, many experts agree that natural word stress had to be employed as a governing feature in order for concepts to be conveyed and metrical patterns to be inferred.[92] But stress does not dictate tempo or emphasis within an individual performance. Mertens writes that *Minnesang*, because it was sung, had to at least adhere to limits of breathing (dictated by vowels and syllables).[93] Yet the principles of music theory and practice, and their integration with rhythmical poetry, were just being rewritten before the *Minnesänger*'s time.

1.3 Music and Rhythmic Poetry in the Middle Ages

In Hugo Kuhn's *Text und Theorie*, he begins by investigating a phenomenology of oral literature, what it can and cannot be, how we are to relate to it, and whether and how it can be categorized or classified.[94] Considering work from biologists and how to incorporate the necessary component of music into an analysis, Kuhn understands human language to be much closer to music than sounds from other animals. It can take on a rhythm, tone, and melody, just as music can, but it differs in its unique relationship to meaning.[95] Mov-

[88]Treitler, "Homer and Gregory: the transmission of epic poetry and plainchant," 345.

[89]Mertens, "Was ist Rhythmus in Minnesang?," 176.

[90]Ibid., 177.

[91]Ibid., 178.

[92]ibid., 184; for the relationship between neumes and melismatic song see Karlheinz Schlager, "Die Neumenschrift im Licht der Melismentextierung," *Archiv für Musikwissenschaft* 38, no. 4 (1981): 296; Kohrs also discusses the relationship between text and music, specifically for MHG Klaus Heinrich Kohrs, "Zum Verhältnis von Sprache und Musik in den Liedern Neidharts von Reuental," *Deutsche Vierteljahrsschrift für Literaturwissenschaft und Geistesgeschichte* 43, no. 4 (1969): 615

[93]Mertens, "Was ist Rhythmus in Minnesang?," 191.

[94]Hugo Kuhn, *Text und Theorie*, vol. 2, His Kleine Schriften (Stuttgart: Metzler, 1969), 19.

[95]Ibid., 48.

ing beyond this comparison, Kohrs, using Saussure's terminology of the difference between *signifié* and *signifiant*, investigates the difference in language through music: music adds an additional element that language "due to its acoustic range cannot."[96]

Several major changes in poetics and music in the Middle Ages were important to the vocal dimension of the performance. There were marked shifts away from the traditional classical authorities and toward new writers, who theorized the changing contemporary medieval music. The dominance of Donatus' *Ars minor* and *Ars maior* were replaced by late 12th century writers Alexander of Villediu and Eberhard of Béthune.[97] Not only was grammar retheorized, but the place of verse and poetics within grammar was no longer secure. Rhyme and rhythm began to be discussed outside of its traditional realm, and commonly in connection to music, e.g. in Matthew of Vendôme's *Ars versificatoria*, Geoffrey of Vinsauf's *Poetria nova*, Gervais of Melkley's *Ars versificaria*, Johannes de Garlandia's *Parisiana poetria*, and Eberhard the German's *Laborintus*.[98]

A major change in music occurred with the shift from chant style monophonic music to polyphonic music in the late 13th and early 14th centuries. Much of this change was pioneered in France, around Paris. Prior to this change, the emphasis on rhythm in monophonic music was much more interconnected with words and syllables. Unfortunately, there is little discussion in the primary sources of such formal evolution in the period around 1200.[99] Nevertheless there were multiple treatises composed in the 12th and early 13th centuries under the category "*De rithmis*" that begin to shed light on an evolving understanding of rhythm.

In late antiquity the rhythm of the Latin language also changed, and differences between long and short syllables were no longer discernable.[100] Whether the source of this change were the dialects themselves has been debated, but the rise of rhythmical poetry owes a great deal to the trouble required to regain the knowledge of the durations of classical Latin.[101] 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.[102] Isidore of Seville, defining the three parts of music in his 7th century *Etymologiae*, quite clearly distinguishes rhythm from meter in the context of words, an unnecessary distinction if one presumes the accent of classical meter:

> The three parts of music (*De tribus partibus musicae*) 1. Music has three parts,

[96]Kohrs, "Zum Verhältnis von Sprache und Musik in den Liedern Neidharts von Reuental," 605.

[97]Philipp Jeserich and Michael J. Curley, *Musica naturalis: speculative music theory and poetics, from Saint Augustine to the late Middle Ages in France*, Rethinking theory (Baltimore, Maryland: Johns Hopkins University Press, 2013), 254.

[98]Ibid., 253-254.

[99]Margot E. Fassler, "Accent, Meter, and Rhythm in Medieval Treatises "De rithmis"," *The Journal of Musicology* 5, no. 2 (April 1987): 164.

[100]Dag Ludvig Norberg and Jan M. Ziolkowski, *An introduction to the study of medieval Latin versification* (Washington, D. C: Catholic University of America Press, 2004), 81.

[101]Ibid.

[102]cf. Augustinus, *De musica*, III, 3, 5. "syllabarum longarum et brevium cognicionem me non habere." ("I cannot recognize the difference between long and short syllables.")

> that is, the harmonic (*harmonicus*), the rhythmic (*rhythmicus*), and the metric (*metricus*). The harmonic part is that which differentiates high and low sounds. The rhythmic part inquires about the impact of words, whether the sound agrees well or badly. 2. The metrical part is that which recognizes, by means of a demonstrable system, the measure of different meters, as for example the heroic, the iambic, the elegiac, and so on.[103]

Isidore explictly ties rhythm to the words themselves, and whether they *cohaereant* well or poorly with the sound (delineated by harmonics). Bede, however, is first to distinguish and coin "rhythmic poetry", though he does not recognize it yet as a respectable craft in its own right. He dedicates only a small section to *De rithmis* in his *De arte metrica et de schematibus et tropis*:

> Rhythmic verse resembles metrical verse. Rhythmic verse is a harmonious arrangement of words which is scanned, not by a quantitative system of meter, but by the number of syllables judged in accordance with the way they sound to the ear, as are the verses of common poets. Rhythm can certainly exist by itself without meter, but meter cannot exist without rhythm. This can be defined more clearly as follows: metrical verse is a quantitative system with a rhythmical beat, while rhythmic verse has a rhythmical beat without a quantitative system. However you will commonly find measured quantities by chance in rhytmic verse, not because the regular artistic arrangement has been preserved, but from the influence of the sound and rhythmical beat itself...[104]

Bede writes that these are the *the songs of the common poets*[105], likely referring to their inferior skills in comparision to their learned, dactylic-hexameter writing counter-parts.[106] Bede's distinction is similar to Isidore's, arguing that rhythmical verse is more focussed on

[103]Isidori Etymologiae Liber III Caput XVIII, "[1] Musicae partes sunt tres, id est, harmonica, rhythmica, metrica. Harmonica est, quae decernit in sonis acutum et gravem. Rhythmica est, quae requirit incursionem verborum, utrum bene sonus an male cohaereat. [2] Metrica est, quae mensuram diversorum metrorum probabili ratione cognoscit, ut verbi gratia heroicon, iambicon, elegiacon, et cetera." Isidore and Stephen A Barney, *The etymologies of Isidore of Seville* (Cambridge, UK; New York: Cambridge University Press, 2006), 96

[104]"Videtur autem rithmus metris esse consimilis, que est verborum modulata conpositio, non metrica ratione, sed numero syllabarum ad iudicium aurium examinata, ut sunt carmina vulgarium poetarum. Et quidem rithmus per se sine metro esse potest, metrum vero sine rithmo esse non potest. Quod liquidius ita definitur: metrum est ratio cum modulatione, rithmus modulatio sine ratione. Plerumque tamen casu quodam invenies etiam rationem in rithmo, non artifici moderatione servata, sed sono et ipsa modulatione ducente..." Bede, Bede, and Calvin B. Kendall, *Libri II De arte metrica ; et, De schematibus et tropis: The art of poetry ; and, Rhetoric*, ed. Stiftsbibliothek Sankt Gallen, Bibliotheca Germanica 2 (Saarbrücken [Germany]: AQ-Verlag, 1991), 160-161

[105]*carmina vulgarium poetarum*

[106]Bede, Bede, and Kendall, *Libri II De arte metrica ; et, De schematibus et tropis: The art of poetry ; and, Rhetoric*, 160-161.

a well-grouped sequence of words and not on syllable duration within those words. This contrasts strongly with definitions of *rhythmus* before late antiquity, produced primarily by grammarians. While the phenomenon of rhythmical poetry worked its way down from the north[107], many musical treatises on the continent still focussed on Latin chant, such as Aurelian of Réôme's *Musica disciplina*. Nevertheless Aurelian importantly expands upon Isidore's three-fold definition of music:

> It [human music] has three parts: harmonics, rhythmics, and metrics. Harmonics distinguishes high and low inflection in sounds, as in the *Ant. Exclamaverunt ad te Domine. Ex-* is a low inflection; *-clama-* is a harmonic; *-verunt*, a high inflection. Rhythmics inquires into the relationship of the words, whether the sound hangs well or ill. Rhythmics seems to be very similar to metrics; but rhytmics is a moulded (or sung, *modulata*) composition of words, analysed not by the system of metrics, but by the number of syllables, and it is judged by the discrimination of the ears: such are most Ambrosian Hymns... For metre is quantitative measurement (or quantity with a melody *ratio cum modulatione*); rhytmics is measurement (or melody, *modulatio*) without quantity and is discerned by the number of syllables. Metrics investigates with sound reasoning the measurement (*mensuram* of different metres...[108]

Aurelian refocuses rhythm back to counting syllables. Yet he distinguishes rhythm as not being contained within a metrical system that measures duration quantitatively, but rather simply the number of syllables regardless of their lengths. Both Isidore and Aurelian understand meter as requiring an agreed upon template, e.g. elegiac, heroic, etc. These templates allow for varying syllable counts dependent upon syllable length, while rhythm is syllable count alone, ensuring that the words agree with one another (reminiscent of Isidore's *cohaereat*, i.e., the natural stress of the language). The anonymous 9th century *Scholia enchiriadis* and counterpart treatise *Musica enchiriadis* incorporate these early distinctions and definitions of rhythm by relaxing the role of quantity in music:

[107] As the earliest rhymed poetry has been discovered in Ireland, and subsequent forms found later going south, many scholars have concluded that such poetry began north and travelled south. This logically supports why Bede was the first to distinguish the new form.Jeserich and Curley, *Musica naturalis: speculative music theory and poetics, from Saint Augustine to the late Middle Ages in France*, 257; Norberg and Ziolkowski, *An introduction to the study of medieval Latin versification*, 106

[108] Aurelian, *The discipline of music: Musica disciplina*, Colorado College Music Press Translations 3 (Colorado Springs: Colorado College Music Press, 1968), 11. "Sunt ergo tres: videlicet, armonica, rithmica, metrica. Armonica est quae discernit in sonis acutum et gravem accentum, ut est hic: *Ant. Exclamaverunt ad te Domine. Ex*, gravis accentus, *clama* armonica, *verunt*, acutus accentus est. Rithmica est, quae incursionem requirit verborum, utrum sonus bene an male cohereat. Rithmus namque metris videtur esse consimilis quae est modulata verborum compositio, non metrorum examinata ratione, sed numero sillabarum atque a censura diiudicatur aurium, ut pleraque Ambrosiana carmina... Etenim metrum est ratio cum modulatione, rithmus vero est modulatio sine ratio, et per sillabarum discernitur numerum. Metrica est, que mensuram diversorum probabili ratione cognoscit metrorum..." Aurelian and Lawrence Arthur Gushee, *Musica disciplina*, 21 (American Institute of Musicology, 1975), 67.

> Therefore, to sing rhythmically in this way is to measure duration proper for long and short tones, never extending or contracting here and there more than is proper, but holding out the pitch in conformance with the rule of scansion so that the melody can end in that tempo (*mora*) in which it began. If at different times you wish to change the tempo for the sake of variation, that is, to make the course of the melody around the beginning or end more sustained or more rapid, do it by a factor of two, that is, halve a long duration or double a shorter one.[109]

While the author clearly recommends conforming to the meter of the song, there is also a clear license given to change syllable quantity for variation. Importantly, the author does not instruct a doubling of already long syllables, or shortening of short syllables, but making long syllables short (by halving) or short syllables long (by doubling), essentially relegating quantity to a mere guide in music. The *Scholia enchiriadis* further emphasizes number and counting as the crucial element in melody:

> Therefore, number controls through the proper measurements of pitches whatever is agreeable in well-formed melody. Whatever is admirable in a delightful rhythm or in well-formed melodies or in any rhythmic movements is all produced by number. Pitches certainly pass away quickly; numbers, however, which are altered through the corporeality of voices and the material substance of things in motion, remain.[110]

After these 9th century treatises hinting at a deemphasis of quantity and an increasing emphasis on count, there is little written about rhythmical verse until Alberic of Monte Cassino in the 11th century, at which time "rhythmic poetry became increasingly important in western Europe, [and] the tradition came to exert itself in all types of sacred poetry, in texts for hymns, tropes, sequences, and other liturgical and paraliturgical forms."[111] The *De rithmis*

[109]Raymond Erickson and Claude V. Palisca, eds., *Musica enchiriadis ; and, Scolica enchiriadis*, Music theory translation series (New Haven: Yale University Press, 1995), 51. "Solae in tribus membris ultimae longae, reliquae breves sunt. Sic itaque numerose est canere, longis brevibusque sonis ratas morulas metiri, nec per loca protrahere vel contrahere magis quam oportet, sed infra scandendi legem vocem continere, ut possit melum ea finiri mora, qua coepit. Verum si aliquotiens causa variationis mutare moram velis, id est circa initium aut finem protensiorem vel incitatiorem cursum facere, duplo id feceris, id est ut productam moram in duplo correptiore, seu correptam immutes duplo longiore." Hans Schmid, ed., *Musica et Scolica enchiriadis: una cum aliquibus tractatulis adiunctis*, Recensio nova post Gerbertinam altera ad fidem omnium codicum manuscriptorum, Veröffentlichungen der Musikhistorischen Kommission 3 (München: Verlag der Bayerischen Akademie der Wissenschaften : In Kommission bei der C.H. Beck'schen Verlagsbuchhandlung, 1981), 87.

[110]Erickson and Palisca, *Musica enchiriadis ; and, Scolica enchiriadis*, 69. "Igitur quicquid in modulatione suave est, numerus operatur per ratas dimensiones vocum, quicquid rithmi delectabile prestant sive in modulationibus seu in quibuslibet rithmicis motibus, totum numerus efficit. Et voces quidem celeriter transeunt, numeri autem, qui corporea vocum et motuum materia decolorantur, manent." Schmid, *Musica et Scolica enchiriadis: una cum aliquibus tractatulis adiunctis*, 113-114.

[111]Fassler, "Accent, Meter, and Rhythm in Medieval Treatises "De rithmis"," 169.

of Alberic of Monte Cassino is considered the starting point for the treatise literature on speech rhythms.[112] Alberic distanced himself from writing on chant by considering rhythmic poetry and its relationship to music. Although *De rithmis* is relatively short, it focuses solely on rhythmic poetry and makes a strong claim in the introduction:

> There are some rhythms, in which measure is considered, so far as all of the syllables without consideration of length (*longitudinis*) or brevity (*brevitatis*). Others are such, in which with certainty and definition the number of syllables and also the length and brevity is foreseen/provided for. It can clearly be said: rhythmical and metrical verse are equal.[113]

Jeserich understands this as implying that rhythmical and quantitative verse "differ in their base, but not in their definition by numerical determinations and not in their emphasis on the isometric construction of repeatable units."[114] Both forms require quantification on two levels, that of syllable valuation *and* systemization: "What is crucial with respect to Alberic of Monte Cassino is the broadening of the frame of reference to the repition of larger rhythmical units."[115] The only difference being in the syllable valuation—the additional syllable length restraint in quantitative verse, which obviates consistent line counts. While Alberic does describe the lengthening or emphasizing of penultimate syllables, he importantly does not yet mention rhyme.[116] Although it could be inferred from his description, his writing style would presume such an explicit explanation if this phenomenon was widespread at the time.

Upon several occasions in his treatise, Alberic mentions penultimate syllabes as long by reason of accent ("*producta penultima accentu*" [6], "*penultima accentu producta*" [7], "*penultima uniuscuiusque accentu et sono producta*" [9]). Scholars disagree on interpreting this *producta* as being held for a long duration, or simply bearing the natural accent of the word. Margot Fassler raises this question in her article 'Accent, Meter, and Rhythm in Medieval Treatises "*De rithmis*"':

> When he speaks of "a long syllable by reason of accent", does Alberic mean that the syllable was really read or, more properly, sung with long duration? This question, of great importance to students of medieval music, does not seem to

[112] Jeserich and Curley, *Musica naturalis: speculative music theory and poetics, from Saint Augustine to the late Middle Ages in France*, 258.

[113] Hugh H. Davis, "The 'De rithmis' of Alberic of Monte Cassino: A Critical Edition," *Mediaeval Studies* 28 (January 1, 1966): 208, ISSN: 0076-5872. "Rithorum alii sunt in quibus considertur mensura tantum sillabarum sine omni longitudinis et brevitatis consideratione. Alii sunt in quibus cum certo et determinato numero sillabarum etiam longitudo et brevitas est prospecta. Quod est apertius dicere: rithmi pariter sunt et metra."

[114] Jeserich and Curley, *Musica naturalis: speculative music theory and poetics, from Saint Augustine to the late Middle Ages in France*, 259.

[115] Ibid., 260.

[116] Fassler, "Accent, Meter, and Rhythm in Medieval Treatises "De rithmis"," 171-172.

> have interested either medieval theorists concerned with rhythmical poetry or modern students of this verse.[117]

Davis, author of the sole critical edition of Alberic's *De rithims*, opts for a strongly contextualized reading:

> The formula throughout the treatise is *producta penultima accentu* (the word order is sometimes changed, and a finite verb used instead of the participle)—'the penult held by reason of the word accent'—and the meaning in this passage is the same, namely, that the next-to-the-last syllable in the first half of the verse, that part composed of eight syllables, receives the stress because it is long, or, more precisely from the point of view of medieval rhythmics, it receives the word accent.[118]

Fassler makes an important point about taking interest in this question, but also prompts a follow-up question: if Alberic did not seem interested in this distinction, was this critical to the music of the period? As with the *Scholia enchiriadis*, the importance of the dynamic penultimate in these cases is its accented nature. The distinction between long and short syllables was already disappearing in music and poetry. Thus perhaps the greatest significance of Alberic's *De rithmis*, as Fassler suggests, is uniting music and poetry, which up until his time had been kept separate by theorists and practitioners alike: "The total penetration of the poetic and musical arts by the style described by Alberic of Monte Cassino was the most significant single event of the twelfth century in either of these realms... It [rhythmic poetry] operated through the uniting of words and music under a system of common aesthetic goals."[119] This unification was furthered by John of Garland's *Parisiana poetria*, which also shifted the focus on rhythmic poetry from the grammarians and rhetoricians to music theorists.[120]

Not until the 12th and 13th centuries do treatises emerge considering rhyme as an integral part of rhythm; the anonymous *De rhythmico dictamine* is the first, proclaiming: "Rhythm is rhyming equality of syllables assembled under a given number."[121] This *consonans* is crucial to the new form of verse, which *De rhythmico dictamine* exemplifies through a wide variety examples, both sacred and secular.[122] Not only does *consonans* encompass end rhyme, but a *consonans* of the *clausulae*, also mentioned in *De rhythmico dictamine*:

> Here the musical term *consonantia* is encountered as a designation of the effect that sometimes results from isorhythm, that is, from the repetition of the

[117]Fassler, "Accent, Meter, and Rhythm in Medieval Treatises "De rithmis"," 172.
[118]Davis, "The 'De rithmis' of Alberic of Monte Cassino: A Critical Edition," 220.
[119]Fassler, "Accent, Meter, and Rhythm in Medieval Treatises "De rithmis"," 172.
[120]Ibid., 179.
[121]Giovanni Mari, *I trattati medievali di ritmica latina*, vol. 11 (U. Hoepli, 1899), 383. "Rithmus est consonans paritas sillabarum sub certo numero comprehensarum."
[122]Fassler, "Accent, Meter, and Rhythm in Medieval Treatises "De rithmis"," 179.

> unit defined by a given number of syllables, and sometimes, as *consonantia finalis*, from end-rhyme. Both the concept of "rhythmical" verse documented here and the terminological formulation of the effect based on isorhythm/isometry as *consonantia* established themselves.[123]

Fassler contrasts this discussion of *clausulae* with Alberic's discussion of *membrae*:

> Thus, whereas Alberic was concerned primarily with the rhythmical patterns within individual lines (which he called "*membra*"), this author is very concerned with the relationships between lines within a clausula and by the shapes of four syllables at the least and sixteen at the most.[124]

With *De rhythmico dictamine* and other similar treatises, the larger verse structure and repetition becomes central because end rhyme is pulling the parts together. This new development in rhythmical verse only further equalizes syllables. A variation of the *De rithmico dictamine*, the *Regulae de rithmis* emphasizes syllable count explicitly.[125]

> For rhythm is the spoken arrangement and agreement of consecutive syllables of equal length. — It is said, however, that rhythm, from the greek *rithmos* (number), because it is constituted by a fixed law of numbers. Thus number is observed in this, firstly in diction, afterward in the syllables and *consonances*.[126]

Treatises of this period often refer back to the Greek origin of *rithmos* as 'number' to emphasize the importance of count. The definition in the *Regulae de rithmis* does not even mention duration until later in the treatise. As Fassler writes, the author of the *Regula* seems to say that all syllables in a rhythmic poem are equal, giving the impression that duration is simply not a factor in this style of poetry.[127] Accent is more important to cadence. Subsequent treatises only mention long and short syllables in regards to cadence and accent, otherwise proving unimportant.[128] Fassler concludes that 12th and 13th century treatises describe an environment, in which syllables are generally of equal length save the penultimate, and the accenting is according to the natural stress of the word.[129] This process comes as no surprise when one follows the tradition after the wane of quantitative meters in the 4th century evidenced by Augustine's remarks discussed earlier. As Hans Tischler

[123]Jeserich and Curley, *Musica naturalis: speculative music theory and poetics, from Saint Augustine to the late Middle Ages in France*, 260.

[124]Fassler, "Accent, Meter, and Rhythm in Medieval Treatises "De rithmis"," 175.

[125]Ibid., 177-178.

[126]"Rithmus enim est congrua diccionum ordinatio, consona, continenter sillabarum aequalitate prolata. — Dicitur autem rithmus a graeco rithmos, idest numero, quoniam certa lege numerorum constituendus est. Numerus ergo in ipso notandus est, primo quidem in distinccionibus, postmodum vero in sillabis et consonanciis." Mari, *I trattati medievali di ritmica latina*, 400.

[127]Fassler, "Accent, Meter, and Rhythm in Medieval Treatises "De rithmis"," 178.

[128]Ibid.

[129]Ibid., 186-187.

argues: "it must be presumed that stress remained fundamental in varying degrees to most Western poetry ever since."[130] In fact, the deemphasis on duration and increased attention on numbers and counting simply reflects the evolving culture: "That a period so occupied with numbers would not reflect this tendency in its lyric and epic poetry by metric rhythm is nearly inconceivable."[131]

Tischler is interested in the musicality of songs from the High Middle Ages, and maintains that much of the secular tradition followed the rhythmical format of the Latin songs described by the theorists above:

> For example, the tradition of medieval German song indicates that each text syllable was usually sung to a single melody tone or figure. Exceptions occur occasionally, particularly on penultimate and final verse syllables, less often within a line, but ornaments rarely go beyond three or four tones. Although the *Königstein Songbook*, which dates from the 1470's, is a very late source of medieval songs, works which cannot be designated as minnesongs, their rhythmic approach certainly continues that of the preceding centuries; for German music was very conservative during the Middle Ages.[132]

This rather rigid structure of syllable equality, holding melismas for the cadence, reiterates the decreasing concern of correct stress or duration as introduced by the *Scholia enchiriadis* in the 9th century: "Apparently, however, the medieval poets felt no great compunction about this crime of giving some naturally weak syllables metric stress and leaving some normally strong ones unstressed".[133]

Leo Treitler importantly adds to this discussion the idea of 'poetic license', arguing that we cannot categorize medieval music into "written" and "unwritten" transmission because the actual transmission of medieval music was often itself a very complicated act. Drawing from memory did not imply an exact reproduction, but rather a living *re*interpretation in performance, an argument echoed in the formula and performance scholarship discussed earlier and spearheaded by Zumthor. Treitler synthesizes this scholarship into a middle ground, pointing out that performances were not always a free improvisation; there can be no black and white categorization for these modes.[134] Treitler claims that "from the very beginning of the written tradition reading, remembering, and extemporizing were continuous acts; they were mutually supportive and interdependent."[135] While these songs were not "free rhythm", motives for reproduction varied from intending high similarity to intending high variation, as already stated: "The singers of one tradition may be highly motivated to strive

[130]Hans Tischler, "Rhythm, Meter, and Melodic Organization in Medieval Songs," *Revue belge de Musicologie / Belgisch Tijdschrift voor Muziekwetenschap* 28 (1974): 11.

[131]Ibid., 10.

[132]Ibid., 6.

[133]Ibid., 12.

[134]Leo Treitler, "The "Unwritten" and "Written Transmission" of Medieval Chant and the Start-Up of Musical Notation," *The Journal of Musicology* 10, no. 2 (1992): 132.

[135]Ibid., 135.

for sameness in the reconstruction where the idiom leaves room for choices, while those of another may be motivated toward variety or simply indifferent to the question."[136]

In sum, a clear trend emerges in the period directly preceding the classical MHG *Blütezeit*—–medieval lyric theorists and practitioners deemphasize syllable length in favor of a greater emphasis on syllable count and *consonans* in both secular and non-secular texts alike. This *consonans* is not only that of final consonance, but also the consonance of the entire structure of the song (*clausulae*). In addition, the multiple modes of composition and transmission effectively utilize meter and sound patterning not only for mnemonic purposes, but also as stylistic features. These changes were significant enough to warrant the rise of a new generation of music theorists.

1.4 Justifying the Syllable

While the physicality of MHG voice is difficult to capture today, the sound and rhythm of the MHG voice can be revived. The *Meistersänger* were confronted with a similar problem though fortunately for them, they were not as far removed chronologically. The *Meistersänger* of the German Early Modern period believed they were continuing the famed tradition of the *Minnesänger*. Their solution, and object of greatest focus, was syllable count and placement. Zumthor speaks of quantities and qualities (pitch, frequence, timbre, and volume), yet various Indo-European linguists as early as the Sanskrit grammarian Pāṇini have implicated that the syllable is the fundamental layer of language.[137] The sanskrit word for syllable, *aksara* means precisely 'imperishable' or 'indivisible', as it was the unit of language believed to not be further divisible.[138] It captures a rhythmical unit of speech inclusive of its sound, without concentrating on individual phonemes. Christoph März emphasizes the Meistersänger's imitation of syllable patterning in his article "Der Silben Zall, der Chunsten Grunt. Die gezählte Silbe in Sangspruch und Meistersang".[139] But März places more importance on syllables not as a result of the increasing importance of music as described above, but rather as a method to more closely reproduce the work of the medieval poets. The *Meistersänger* sought out a mathematical approach to prove they were continuing the medieval tradition, and simultaneously developed a system to quantify the success of their work (that of *re*-presentation). One notable early German *Meistersänger* of the 14th century, Hermann von Sachsenheim, relates his art of syllable usage to the famous *Minnesänger* Wolfram von Eschenbach:[140]

[136]Treitler, "The "Unwritten" and "Written Transmission" of Medieval Chant and the Start-Up of Musical Notation," 14, 146.

[137]Zumthor, *Die Stimme und die Poesie in der mittelalterlichen Gesellschaft*, 64; See Rama Nath Sharma and Pāṇini, *The Aṣṭādhyāyī of Pāṇini* (New Delhi: Munshiram Manoharlal Publishers, 1987); John Rupert Firth, "Sounds and prosodies," *Transactions of the Philological Society*, 1948, 127-152

[138]Thomas Egenes, *Introduction to Sanskrit* (Motilal Banarsidass Publ., 1996), 63.

[139]Christoph März, "Der Silben Zall, der Chunsten Grunt. Die gezälte Silbe in Sangspruch und Meistergang.," *Zeitschrift für deutsche Philologie* 119, no. 2000 (2000): 73–84.

[140]Ibid., 79.

> Vonn Eschenbach der eine
> herr Wolffram ist genennet,
> vonn Labern nit der cleyne;
> der beyder kunst ich hann also erkennet
> an rümen, worten, silben wolgemessen.[141]

The idea of "measuring" syllables and rhyme became integral to the *Meistersänger* art. März believes that counting syllables both detracted from the actual art of crafting lyrics and songs, as well as revealed the great desire of the *Meistersänger* to be "correct."[142] Counting syllables ensured they followed the singers before them, the degree of closeness measuring their success as a *Meistersänger*. März adds that because they were influenced so greatly by the seven liberal arts, they needed means to measure or quantify their work. Syllables provided the optimal unit of measurement because they are integral to the metrical tradition. Thus an investigation of the syllable inevitably leads to questions of meter.

Florian Kragl has carried out one of the few analyses of MHG musical notation in comparison to metrical scansion, specifically for Neidhart's *Sommerlied* 23. He observes:

> The correlation of semi-breves and minimae with stressed and unstressed syllables could not be more exact. Semi-breves are always realized in stressed syllables, minimae without exception in unstressed syllables. This yields—of course with the exception of cadence—a strongly alternating meter, which musically is not interpreted any less strongly than long and short note values.[143]

Although little MHG musical notation survives, this insight allows us to understand the importance of syllables to the MHG poetic meter. While Kragl admits that syllable length and syllable duration in MHG does not always correlate as clearly, at least in some manuscripts syllable accents do. Looking at other manuscripts, Kragl sees great variation in how the song could be scanned and thus set to music. Yet Kragl does find one manuscript that sets the words to music in near perfect relation to the meter, but the relationship deteriorates as the text develops.[144]

Klaus Kohrs looks at this relationship as well, coming to a similar conclusion for the melodies of Walther von der Vogelweide: "In principle, the melody sequence corresponds

[141] Reproduced from März, "Der Silben Zall, der Chunsten Grunt. Die gezälte Silbe in Sangspruch und Meistergang.," 79. "Of Eschenbach there is one named Wolfram, who does not speak rarely; both arts I have recognized by rhyme, words, and well-measured syllables."

[142] Ibid., 83.

[143] "Die Korrelation von Semibreves bzw. Minimae mit betonten bzw. unbetonten Silben könnte exakter nicht sein. Semibreves realisieren stets betonte Silben, Minimae ohne Ausnahme unbetonte Silben. Das ergibt—natürlich mit Ausnahme der Kadenzen—ein streng alternierendes Metrum, das im Musikalischen nicht minder streng von langen und kurzen Notenwerten interpretiert wird." Florian Kragl, "wort unde wîse. Formen des sangbaren Verses in der deutschen Literatur des Mittelalters," *Literaturwissenschaftliches Jahrbuch* 52 (2011): 33

[144] Ibid., 65.

to the meter of the text, i.e. the stresses of the text occur on the tones constituting the melody."'[145] But he is sure to point out that this is not always the case, and those other cases must be examined. Kohrs argues that this is where a pre-existing melody takes on its own form and is fitted onto the text, which results in some doubling of tones.[146] Concluding, Kohrs describes three types of musical settings for songs by Walther von der Vogelweide: 1) melodies corresponding very closely to the language and stress accents, 2) melodies not corresponding at all to the language, and 3) an in between, where a melody was clearly modified to better fit the text.[147] Kohrs maintains that metrical form itself is irrelevant for meaning, something I will argue against in the following pages, but is rather only a structured guide for setting to music and that meaning is realized through the meter's combination with language.[148] This was a common belief in the scholarship of the 1970s, claiming that epic literature only becomes such in the moment that epic content is processed through a specific form.[149] I will argue in Chapter 4 that this form, be it music or meter, can actually distinguish genre characteristics independent of content.

In order to detect what Stock refers to as "traces" of a performance in the manuscripts, and to examine broad trends in a varied corpus, this project presents a computational approach to MHG prosody and meter. Much work thus far in the Digital Humanities (DH) has focused on the word as the unit of measurement. This must come as no surprise, as the word is generally considered the starting point for a semantic interpretation. In the interest of moving beyond the word and into the more subtle features of MHG, I suggest the syllable as a rich source of formal information, both aesthetic and stylistic. This element has been neglected until recently in both the DH scholarship and the traditional MHG literary scholarship because these texts are no longer performed and read aloud as they once were.[150] Nevertheless, this project maintains, as do Kragl and Stock, that traces of this voice can be revealed.[151] These traces become all the more powerful and salient when aggregated through computational methods.

The following chapters include: a new syllabification algorithm for MHG (a combination of two powerful algorithms, both of use for universal syllabification), a large-scale analysis of the unique soundscapes of medieval German lyric texts across genres, a prosodic and lexical clustering of MHG verse texts investigating the tension between form and content, and a more restricted analysis of the MHG epic meter (including a supervised machine learning model for prediction of metrical values). This project intentionally grows in the sophistication of the digital methods employed, and thus also necessarily requires a growing tolerance for

[145] "In der Regel entsprechen sich hier Melodieverlauf und Textmetrum, d. h. die Hebungen des Textes fallen mit den melodie-konstitutiven Tönen (Terzenschichtung!) zusammen." Kohrs, "Zum Verhältnis von Sprache und Musik in den Liedern Neidharts von Reuental," 605

[146] Ibid., 619.

[147] Ibid.

[148] Ibid., 605.

[149] Kragl, "wort unde wîse. Formen des sangbaren Verses in der deutschen Literatur des Mittelalters," 41.

[150] As Zumthor argues, we can never truly discover the voice from noisy, heavily mediated manuscripts.

[151] Stock, "Das volle Wort - Sprachklang im späteren Minnesang," 194-195; Kragl, "wort unde wîse. Formen des sangbaren Verses in der deutschen Literatur des Mittelalters," 46-47

treating text as data. Chapter 3 makes very few assumptions about MHG, and the methods employed therein are suitable to any corpus or text structure. Chapter 4 takes liberties by assuming more of MHG verse, as well as its orthography. Chapter 5 makes the most assumptions, creating a model for the most well-known and subscribed scansion theory, yet depending on normalized orthography and implemented only on MHG texts in the *Vierheber* epic meter. However, as Andreas Heusler aptly writes: "Whoever wishes to restrict the study of meter to what can be proven would have to settle for prosodic statistics."[152] Let's not settle.

[152] "Wer die Verslehre auf das Beweisbare einschränken wollte, müßte sie zur prosodischen Statistik erniedrigen." Heusler, *Deutsche Versgeschichte: mit Einschluss des altenglischen und altnordischen Stabreimverses*, 13.

Chapter 2

Syllabification and Noisy Data

To explore the sound and rhythm of the medieval voice on the level of the syllable on a large scale, as suggested in the first chapter, an accurate syllabification algorithm must be developed. This chapter aims to create a rule-based syllabification algorithm for MHG that is accurate for all the varying dialects and orthographies.[1] Syllabification is governed in many languages by established linguistic principles. The principles of onset maximization and legal initials (early formulation by Vennemann[2]) and sonority sequencing (early formulation by Jesperson[3]) govern syllabification in many languages, including MHG to a great degree. If the orthography is true to the phonology of the language, as is the case in MHG, these two principles are extremely accurate on the actual graphemes themselves, if not, such as in languages like English and French, performance can be significantly less accurate.[4] For these cases, these principles remain useful if words are first transcribed to the International Phonetic Alphabet (IPA). What follows is a description of a generalized algorithm combining these two principles implemented programmatically, something not yet attempted for MHG, or any other language for that matter.[5] Following the general algorithm, I enumerate additional rules necessary for a more accurate MHG syllabification. This approach will prove to be particularly effective for low resource languagues such as many historical languages, as they often lack a look-up technique. Moreover, a rule-based algorithm based on phonemes is particularly well-suited to working with non-standardized languages and dialects, which

[1]Parts of this chapter were published in Christopher Hench, "Phonological Soundscapes in Medieval Poetry," *Proceedings of the Joint SIGHUM Workshop on Computational Linguistics for Cultural Heritage, Social Sciences, Humanities and Literature, ACL 2017*, 2017, 46

[2]Theo Vennemann, "On the Theory of Syllabic Phonology," *Linguistische Berichte* 18 (1972): 1–18.

[3]Otto Jespersen, *Lehrbuch der Phonetik;* in collab. with Robarts - University of Toronto (Leipzig, Teubner, 1904).

[4]The issue of standardization and linguistic normalization will be addressed below.

[5]The Sonority Sequencing Principle (SSP), legality principle (LP), and onset maximization were tested separately computationally by Bartlett et al.Susan Bartlett, Grzegorz Kondrak, and Colin Cherry, "On the syllabification of phonemes," in *Proceedings of Human Language Technologies: The 2009 Annual Conference of the North American Chapter of the Association for Computational Linguistics* (Association for Computational Linguistics, 2009), 308–316

characterizes MHG to a high degree.

2.1 Onset Maximization and Legality

Vennemann's 1972 lecture *On the Theory of Syllabic Phonology*, prompted by Noam Chomsky and Morris Halle's 1991 *The Sound Pattern of English*,[6] primarily investigates syllable clusters in Modern Icelandic, but reaches a climax in formulating universal syllabification rules rooted in the strength of a language's consonants.[7] He argues that in every language the inventory of consonants can be scaled by their strength. The strength of a consonant depends on several factors, including its effect on voicing surrounding phonemes and where consonants occur in relation to vowels. Only strong consonants may begin a medial syllable, although these strong consonants differ by langauge. Strong consonants are particularly identifiable as they also appear in word-inital position. Vennemann formulates the law of intials as follows:

> Medial syllable initial clusters should be possible word initial clusters. — I fully realize that this principle is broken over and over again as a consequence of historical change, but I also have examples of how a mess created by historical change is cleaned up by an application of, or enactment of, the same principle.[8]

In plain terms, the consonant cluster beginning any syllable should also be permissable to begin a word in that language. Four years later in 1976, Daniel Kahn's dissertation formulated a similar rule in what became known as *onset maximization*.[9] Kahn investigates English specifically and adheres to Vennemann's law:

> Thus I will accept as a safe working hypothesis the assumption that the set of possible syllable-initial (-final) clusters in English is identical to the set of possible word-initial (-final) clusters.[10]

Yet Kahn notes that this statement still leaves several possibilities if a consonant cluster could be broken up into more than one permissable word-initial grouping. To remedy this, Kahn argues:

> As a first step towards a determination of the actual rules which assign syllabication of intervocalic consonants in English, let us consider slow, over-precise

[6]Noam Chomsky and Morris Halle, *The sound pattern of English*, 1st MIT Press paperback. ed (Cambridge, Mass: MIT Press, 1991).

[7]Vennemann is encouraged by Joan Hooper's work to define rules for universal syllabification. Joan B. Hooper, "The Syllable in Phonological Theory," *Language* 48, no. 3 (September 1972): 525.

[8]Vennemann, "On the Theory of Syllabic Phonology," 11.

[9]Daniel Kahn, "Syllable-based generalizations in English phonology" (PhD diss., MIT, 1976).

[10]Ibid., 41.

> speech, or even the type of speech one might use to imitate a science-fiction movie robot. What we find in this sort of articulation in cases where more than one syllabification is consistent with the general constraints is a strong tendency to syllabify in such a way that initial clusters are of *maximal length*, consistent with the general constraints on word-initial consonant clusters.[11]

These two rules alone, legal initials and onset maximization, form the base of a powerful syllabifier, particularly for languages with a true orthography as noted above.[12] Yet these two rules are still only guiding principles, and are not as accurate as a detailed phonological approach. These rules are more effective when implemented on top of a more precise scheme for syllabification, one provided by the sonority sequencing principle.

2.2 Sonority Sequencing Principle

The Sonority Sequencing Principle (SSP)[13] was formulated by Otto Jesperson, contrasting with other theories of syllabification at the time:

> According to some it depends on the expiration: "a syllable is a group of phonemes, which are spoken with an exhalation" According to others, the concept of a syllable has nothing to do with expiration, rather with the natural strength of the sound, fullness of the sound, and sonority of the phonemes. And finally there are phoneticians who combine both statements and say that there are two types of syllables: expiration syllables and sonority syllables.[14]

Jesperson builds upon the latter group of phoneticians by constructing a hierarchy based on the sonority of a phoneme:

[11] Kahn, "Syllable-based generalizations in English phonology," 41.

[12] As no computer algorithm had yet existed for a syllabification based purely on the *Law of Initials* and *onset maximization*, in cooperation with a colleauge Alex Estes, I wrote a script available for general use at `https://github.com/henchc/LegaliPy`. The code is in Appendix A. In the interest of low resource languages, the script does not require input of the language or permissable onsets, rather determines the permissable onsets from the text to be syllabified. Naturally, the greater the volume of text, the more accurate the syllabifier will be. The program retains those onsets as legal that appear in greater than .02% of all words, thus accounting to a degree for foreign loanwords and typos in the submitted text.

[13] The SSP is a powerful algorithm in its own right, and, like legal initials and onset maximization, has not been made available as free software. To this end, again with Alex Estes, I have written a script available for general use at `https://github.com/henchc/SonoriPy`. The code is in Appendix B. Unlike LegaliPy above, SonoriPy requires more input from the user, specifically a sonority hierarchy. As Jesperson defined, and performed here for MHG, it is suggested to generate a hierarchy of *groups* of phonemes, as opposed to a scale of the individual phonemes themselves.

[14] "Nach einigen beruht sie auf der Expiration: 'eine Silbe ist eine Lautgruppe, die mit einem Ausatmungsdruck (Expirationshub) gesprochen wird.' Nach andern hat der Silbenbegriff nichts mit der Expiration, sondern dagegen mit der natürlichen Schallstärke, Schallfülle, Sonorität der Laute zu tun. Und endlich gibt es Phonetiker, welche die beiden Anschauungen kombinieren und sagen, daß es zwei Arten Silben gibt: Expirationssilben und Sonoritätsilben." Jespersen, *Lehrbuch der Phonetik;* 186

> The fullness of a sound is a result of the same factors that determine its entire form, but in a manner, that the use or non-use of the voice plays the greatest role. Among the voiceless phonemes there are only minor differences, so that I only develop two subdivisions; but on the otherhand within the voiced phonemes there are great differences. Here one can generally say, that the fullness of a sound has a direct relationship to the size of the space through which the moving air has to pass.[15]

Jesperson continues to list phonemes in their order from least sonorous to most sonorous. Importantly, Jesperson groups phonemes first by whether they are voiced or voiceless, then stops, fricatives, nasals, laterals, rhotic, and three levels of vowels (high, medium, and low). Thus it is not a hierarchy of each phoneme individually, but rather a hierarchy of groups of phonemes. A phoneme is judged as "sonorous" based on the degree to which the lips are opened, and sound is allowed to pass through the mouth.[16] Accordingly, the most sonorous phoneme is an open [ɑ]. A syllable break appears *before* a trough in sonority.[17] Importantly, this theory utilizes *relative* sonority, not *absolute* sonority within a word, as there naturally can be several syllables in any given word. The actual working of the legality principle and the SSP will become clear in its application to MHG in what follows.

2.3 Syllabification of Middle High German

A syllabifier for MHG was constructed utilizing the above principles.[18] Syllabification for MHG was implemented in a sequence of steps, each step improving upon the syllabification output of the previous step. The workflow is depicted in Figure 2.1.

[15] "Die Schallfülle eines Lautes ist eine Resultante aus denselben Faktoren, die überhaupt sein ganzes Gepräge bedingen, doch so, daß der Gebrauch oder Nichtgebrauch der Stimme die größte Rolle spielt. Innerhalb der stimmlosen Laute gibt es nur geringe Unterschiede, so daß ich nur zwei Unterabteilungen aufstelle; dagegen gibt es innerhalb der stimmhaften weit größere Unterschiede. Hier kann man im großen und ganzen sagen, daß die Klangfülle im direkten Verhältnis zu der Größe des Raumes steht, den die schwingende Luft zu passieren hat." Jespersen, *Lehrbuch der Phonetik;* 186.

[16] Ibid., 186-187.

[17] Ibid., 187-188.

[18] This syllabification is presented in Hench, "Phonological Soundscapes in Medieval Poetry." The source-code for the computational syllabification of MHG used in this project is available for general use at `https://github.com/henchc/ACL-LaTeCH-CLfL-2017/blob/master/utils/syllabipymhg.py`. The code is also in Appendix C. Syllabification for other languages with this combined approach is also underway. This project has been grouped into a library under the title SyllabiPy, for its use of the Python programming language. I follow primarily the *Lautlehre* (phonetics) in *Mittelhochdeutsche Grammatik* and Richard Wiese's *The Phonology of German*. Hermann Paul et al., *Mittelhochdeutsche Grammatik* (Tübingen: Max Niemeyer, 1982) Richard Wiese, "The Prosodic Structure of German," in *The phonology of German*, The phonology of the world's languages (New York: Oxford University Press, 2000)

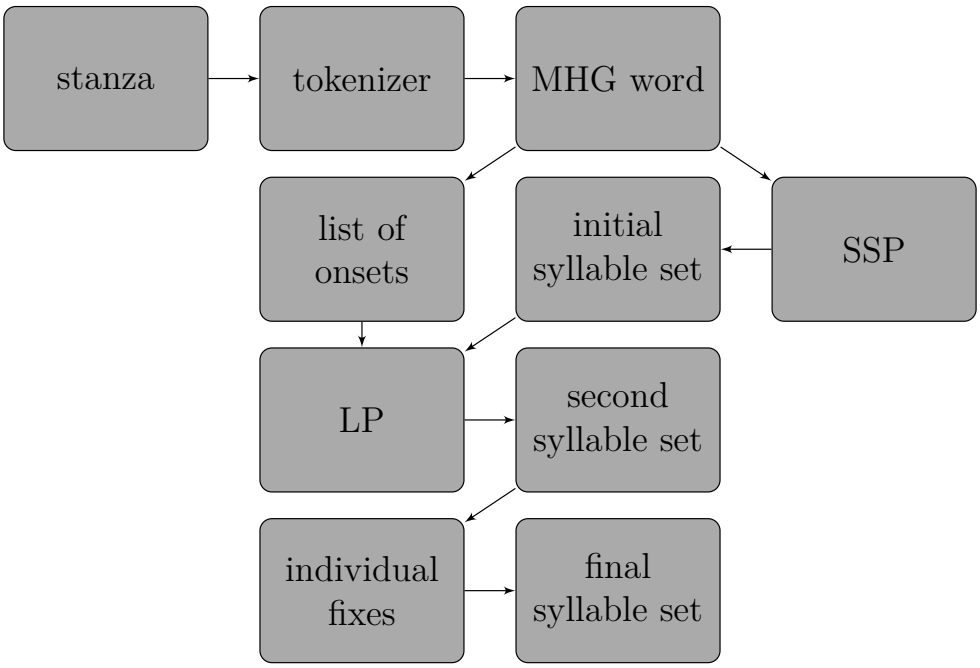

Figure 2.1: Flow chart for MHG syllabification.

First a stanza text or entire corpus is required, only for the purpose of extracting a list of permissible onsets as described in onset maximization above.[19] The text provided for syllabification is first tokenized (a term used in Natural Language Processing (NLP) tasks for splitting a text into words, or units) on white space.[20] After the text has been tokenized, onsets are extracted from the text if a list is not already provided, and the first word becomes the focus of the program. The word *werltlîchen* will serve as an example for understanding the algorithm, as it necessitates each step to yield the correct syllabification.[21] Subsequent steps are only implemented if the syllabification produced by the SSP violates a subsequent rule. Before being syllabified by the SSP, the '*ch*' in *wertlîchen* is converted to a single phoneme.[22] In the SSP, the phoneme hierarchy for MHG is established as:

- vowels (1): a, e, i, o, u, y, â, ä, æ, ê, î, ô, ö, œ, û, ü[23]

[19]For the purposes of the released program SyllabiPy-MHG, a file with permissible onsets is already provided.

[20]While contractions do exist in Middle High German, they are not the focus of this work because contractions can also be accurately syllabified in the same manner.

[21]In fact, most words are syllabified correctly after the first step using the SSP.

[22]For the general MHG algorithm, the phonemes consisting of multiple graphemes converted to a single phoneme include: '*ch*', '*sch*', and '*ph*'. Fortunately for MHG, the orthography is very true to the phonology of the language, and transription to IPA is not necessary. Other language may need further phonetic transcriptions for an accurate syllabification by the SSP.

[23]Note that this allows for varying orthography of the vowels, as all vowels are treated equally, long or

- resonants (2): l, m, n, r, w
- consonants (3): b, c, d, g, h, k, p, q, v, t

While this hierarchy is quite simple, and could certainly be ordered into more levels, a three-level hierarchy resulted in the greatest accuracy for MHG. Phonemes are then assigned a phoneme value corresponding to their placement within the hierarchy; the phonemes of the word *werltlîchen* thus receive the following value assignments:

werl | tlî | (ch)en
2322 123 1 32

After the SSP, the '*ch*' is converted back to two graphemes and *werltlîchen* is now *werl-tlî-chen*. This first output contains the correct number of syllables, but still has several errors (namely the '*t*' and '*ch*'). The syllable set moves on to the legality principle, i.e., every syllable onset must exist in the language as a word onset. It is quite clear that '*tlî*' is not a permissible onset in MHG and must thus be broken up into the greatest possible onset following onset maximization, which is also a legal initial in the language. In this case, the syllable must be '*lî*', and the '*t*' will be pulled to the previous syllable. After this stage, the new syllable set becomes *werlt-lî-chen*.

Metrical Influence

While some may consider this the correct syllabification of *wertlîchen*, I argue that an intervocalic '*ch*' should also be split up due to metrical evidence. Although a more detailed discussion of meter is still to come, it is necessary to briefly note here how meter can provide insight into the correct syllabification of MHG:

daz |ist ein |fremdiu |zec|he[24]
× | x́ × | x́ ×| ˊ– |x̀ ⏜

The syllabification *ze-che* would leave a light syllable to fill an entire foot, which is not permitted. Other syllabifications would break with the natural stress of the language. This is accounted for in a similar fashion as the legal initials and the individual fixes stage.[25] The final syllabification thus yields *werlt-lîc-hen*. This last change is important for my analyses in that it changes an open syllable ('*lî*') to a closed syllable ('*lîc*'), although both are heavy syllables.

A similar situation arises with intervocalic affricates, which may be viewed as either ambisyllabic or biphonemic, for example in MHG '*sitzen*' ('sit'). '*sitzen*' could be syllabified as

short.

[24]Wolfram von Eschenbach Wolfram von Eschenbach et al., *Parzival*, 1 Aufl, Bibliothek deutscher Klassiker (Frankfurt am Main: Deutscher Klassiker Verlag, 1994), 17, l. 5-21 "This is a peculiar arrangement."

[25]The individual fixes include: intervocalic '*ch*', intervocalic '*sch*', a sequence of one long vowel followed by one short vowel, and the suffixes '*lîch*' and '*heit*'

(1) '*si-tzen*' ([zɪ.tsən]), (2) '*sit-zen*' ([zɪt.sən]), or (3) '*sitz-en*' ([zɪts.ən]). Syllabification (1) would yield NHG [zits.ən], which is incorrect. It further does not correspond to manifestations in MHG meter, e.g.:

sie | sêre | solde | let|zen[26]
× | ×́ × | ×́ ×| — |×̀ ⏜

und vil | gar ent|set |zen[27]
| ×́ × | ×́ × | — | ×̀ ⏜

Both lines require a heavy syllable in the double mora (*beschwerte Hebung*) position. Syllabification (3) violates onset maximization because '*ze*' is a permissible onset. This leaves only syllabification (2) as the preferred syllabification. Furthermore, for the algorithm there is no difference between the two interpretations of syllabification (2) — '*sit-zen*' can be interpreted as either ambisyllabic or having two heterosyllabic phonemes. Thus, the program does not need to take a side in this debate.

Individual Fixes

Generally, the division of words with only one intervocalic consonant such as '*ta-ge*' poses no difficulties to the algorithm. Only certain consonant clusters require further information from MHG phonology and must undergo the individual fixes stage. For example, the orthographic sequence of a nasal followed by a velar obstruent, although representing simply a velar nasal in modern German, was in fact still two separate phonemes in MHG.[28] Thus the word *lange* is syllabified as '*lan-ge*' (NHG [lɑŋə] MHG [lɑŋgə]). Some double consonants may also be understood as either ambisyllabic or geminate, e.g. MHG '*vallen/fallen*' (NHG [fɑllən] MHG [fəlːən]). '*vallen*' can either be syllabified as '*val-len*' or as in NHG with the ambisyllabic consonant.

There are also instances where morpheme boundaries interfere with the otherwise normal processes of syllabification. For example, the common MHG suffix *-lich* in *wîplich* results in the syllabification '*wîp-lich*', not '*wî-plich*', despite onset maximization preferring the latter ('*pl*' is a common MHG onset).

Most errors arise from areas of MHG morphology ambiguous to the computer, such as the prefix 'ge-' acting as the perfect participle, as well as functioning as the onset for non-participles. This occurs with the consonant cluster of '*st*'. MHG '*geste-*' could be syllabified as '*ge-ste*' or '*ges-te*'. Naturally, the participle prefix as in MHG '*gestechen*' would prefer '*ge-stec-hen*'. Yet the simple noun '*geste*' shows metrical evidence as prefering '*ges-te*'. This applies to other weak prefixes such as 'be-':

[26]Hartmann and Volker Mertens, *Der arme Heinrich*, Bibliothek des Mittelalters (Cambridge, [England] : Ann Arbor, MI: Chadwyck-Healey ; ProQuest Information / Learning, 2005), 250, l. 361. "[that his death] would cause them great harm."

[27]ibid., 250, l. 362. "and even lose..."

[28]Paul et al., *Mittelhochdeutsche Grammatik*.

Dirre |werlde | ves | te[29]
|x́ × | x́ × | — | x̀ ⌢

ir | stæte | und ir | bes|te[30]
× | x́ × | x́ × | — | x̀ ⌢

Once more, for the double mora to be valid, we must syllabify '*veste*' and '*beste*' as '*ves-te*' and '*bes-te*' respectively. The algorithm cannot solve this problem without further morphological information or stress marking. Fortunately, MHG part-of-speech (POS) taggers are in development, and can soon rectify some of these challenges. For the moment, the most common nouns consisting of the cluster '*st*' preceded by '*e*' are treated individually, and the participle prefix, being more common, is the default syllabification for '*e-st*'.

Accuracy and Advantages

While algorithmic computational syllabification is nothing new, previous methods have not combined syllabification principles in the manner presented here by controlling SSP syllabification with the legality principle (LP).[31] Moreover, this method is particularly suited to languages lacking a standardized orthography. Testing the algorithm across dialects on the new *Referenzkorpus Mittelhochdeutsch* (ReM) corpus[32] yielded an accuracy of 99.4% on a randomly sampled 1,000 words from the entire corpus of diplomatically transcribed texts.[33]

[29]Hartmann and Mertens, *Der arme Heinrich*, 234, l. 97. "The fortification of this world"

[30]ibid., 234, l. 98. "its sturdiness and power"

[31]Early work from Bartlett et al. devised a new method for phonemic translation of English, an important step toward accurate syllabification. Susan Bartlett, Grzegorz Kondrak, and Colin Cherry, "Automatic Syllabification with Structured SVMs for Letter-to-Phoneme Conversion.," in *ACL 2008* (2008), 568–576. Bartlett et al. produced gold standard results of the SSP, LP, and OM, also creating an SVM-HMM model. Bartlett, Kondrak, and Cherry, "On the syllabification of phonemes." Adsett and Marchand test several algorithms across multiple languages concluding that Syllabification by Analogy is most accurate. Connie R Adsett and Yannick Marchand, "A comparison of data-driven automatic syllabification methods," in *International Symposium on String Processing and Information Retrieval* (Springer, 2009), 174–181. Rogova et al. develop an SCRF model for phonetic transcriptions. Kseniya Rogova, Kris Demuynck, and Dirk Van Compernolle, "Automatic syllabification using segmental conditional random fields," *Comput. Linguist. Neth. J* 3 (2013): 34–48 Kondrak et al. seek to improve the gold standard syllabification by including morphological segmentation information. Garrett Nicolai Lei Yao Grzegorz Kondrak, "Morphological Segmentation Can Improve Syllabification," *ACL 2016*, 2016, 99.

[32]Klein, Thomas; Wegera, Klaus-Peter; Dipper, Stefanie; Wich-Reif, Claudia (2016). Referenzkorpus Mittelhochdeutsch (1050–1350), Version 1.0, `https://www.linguistics.ruhr-uni-bochum.de/rem/`. ISLRN 332-536-136-099-5.

[33]Syllabifying diplomatically transcribed texts introduces several orthographic obstacles for computers. Most symbols have a 1-1 signification to a standard grapheme, which for computational purposes, were resolved. The greatest obstacle was resolving 'v' and 'u' notations, as the consonant 'v' is at the other end of the sonority hierarchy than the vowel 'u'. This project assumes a conservative approach by using a corpus of standardized texts from the *Mittelhochdeutsche Begriffsdatenbank* (MHDBDB) to determine permissable environments of the 'v' grapheme Margerete Springeth, Nikolaus Morocutti, and Daniel Schlager, *Mittelhochdeutsche Begriffsdatenbank (MHDBDB). Universität Salzburg.*, `http://www.mhdbdb.sbg.ac.at/`, Accessed: 2016-10-01. If a given environment with a 'v' grapheme in the diplomatic translation does not exist

This algorithm thus offers itself as a useful tool for the syllabification of low-resource languages, particularly those with varied orthography, a significant obstacle for computational text analysis of medieval texts.

Perhaps the most important aspect of this work, and the glue that holds it all together, is the flexibility of the rule-based algorithm. While this algorithm is highly accurate on standardized dialect and orthography for MHG, its true advantage is for application to irregular orthography as demonstrated by its performance on the ReM corpus. Other syllabification methods, particularly dictionary lookup techniques, could not easily handle word variants or differing orthographies. Primarily due to the SSP[34], the algorithm proposed above is accurate across dialects and orthographic variances in the MHG corpus. A survey of dialects and syllabifications is provided below:

- **Hessisch**
  - *hob, hobes* → *hob, ho-bes* (std. mhd. *hof, hofes*)
  - *grebe* → *gre-be* (std. mhd. *græve*)
- **Mittelfränkisch**
  - *maghen, maggen* → *mag-hen, mag-gen* (std. mhd. *machen*)
  - *kracht, nichtel* → *kracht, nich-tel* (std. mhd. *kraft, niftel*)
- **Thüringisch Hessisch**
  - *liden, geliden, adem, verterben* → *li-den, ge-li-den, a-dem, ver-ter-ben* (std. mhd. *liten, geliten, gelitten, âtem, verderben*)
- **Ostmittelhochdeutsch**
  - *ader, adir* → *a-der, a-dir* (std. mhd. *oder*)
  - *quahen* → *qua-hen* (std. mhd. *twahen*)
  - *kârte, gekârt, lârte, gelârt* → *kâr-te, ge-kârt, lâr-te, ge-lârt* (std. mhd. preterites of *kêren, lêren*)

in the standardized environments, the 'v' is converted to a 'u'. All other cases remain, and no other changes are made to the diplomatic transcriptions. Other spelling variations do not undermine the algorithm, as they commonly remain at the same level of the sonority hierarchy. This project therefore errs on the side of undercorrection. Note, however, that one should not compare results to Bartlett et al. (2009) as MHG is significantly easier to syllabify than English due to orthography and vocabulary. These results also include repeated words to demonstrate relevance to its application in the method.

[34]The SSP provides the initial 'rough' syllabification, and as many differences between dialects do not jump between sonority value assignments, syllabification remains stable across dialects

As it will become important later, I would like to introduce here two properties of syllables that characterize their sound: open syllables and closed syllables. An *open syllable* is simply a syllable, which ends in any vowel, it leaves the mouth open with sound coming out. A *closed syllable* is any syllable ending in a consonant, with no sound or air exiting the mouth. I introduce these definitions here to direct attention to the variations above. In *every* case above, the dialect variant of a word is syllabified with *exactly* the same syllable *properties* in order as the normalized, or standard dialect MHG, e.g. *nich-tel* and *nif-tel* both generate two close syllables, *ge-li-den* and *ge-li-ten* both generate two open syllables and one closed syllable. While schwa instances will vary by dialect, other morphological variance can be mitigated with such an abstraction from the lexical unit, helping to unite an analysis across complex orthography and manuscript variance.

2.4 Linguistic Normalization

In an attempt to create a universal language between the varieties of MHG, Karl Lachmann began publishing edited editions of a normalized MHG (*normalisiertes Mittelhochdeutsch*). Scholars such as Andreas Heusler argue that these "Schreibeingriffe" (editorial interventions) have significantly complicated any attempt to seriously analyze true MHG verse.[35] Heusler claims that sentence rhythm, weak syllables, and editing changed so much that we cannot be certain about anything, and for that reason he does not give any statistics in his in-depth discussion of MHG meter.[36] Recent scholars have been less concerned about grammatical corrections, criticizing more the lack of transparency normalization causes, in that words are changed and manuscript mouvance is absent. Yet there are advantages to normalization, in that it allows for students to assimilate the basic principles of the language quicker, and is especially helpful for computational analysis of texts. Nevertheless, normalization is at best a further abstraction from what one may consider the 'text'. However, Joachim Bumke has shown us that even these manuscripts cannot be considered 'true' or the "*Urtext*", viz. "original text".[37] In fact, the normalized abstraction from the manuscript may have a similar distance from the language used in the actual performance as the manuscript itself does. But what issues does this raise for a computational text analysis project?

In many ways, a normalized orthography and grammar actually raise the performance of any analysis on a basic semantic level. Most approaches to content analysis such as clustering and topic modeling suggest first stemming or lemmatizing words. Because most analyses at the root level depend on word frequencies, this preprocessing instructs the computer to not consider words of the same stem or lemma as different. Were it not for a normalized MHG, the computer would understand all the variant forms of e.g. '*machen*' (*maghen*, *maggen*), despite all being the infinitive form, as different *concepts*. It is these *concepts* that are key. Is

[35] Andreas Heusler, *Deutsche Versgeschichte: mit Einschluss des altenglischen und altnordischen Stabreimverses*, 2nd ed., vol. 2, Grundriss der germanischen Philologie (Berlin: W. De Gruyter, 1956), 4.

[36] Ibid., 79.

[37] Joachim Bumke, "Der unfeste Text," *Germanistische Symposien Berichtsbanden* 17 (1996): 118–129.

the concept of *machen* from one MHG variation to the next the same? While semiotics and philosophy will argue that they are not, any computational analysis based on frequencies must, at least initially, in order to have sufficient observations for interpretation.

One motivation of this project is to abstract from the word and rely less on orthography. Nevertheless, as this project attempts a linguistic analysis beginning on the orthographic and phonologic level, we must first consider the effect the use of such normalized texts could have on such an analysis. While I do not aim to dispel all concerns, my main argument is that such textual 'Eingriffe' and normalization may not pose a serious problem if the unit of analysis (here syllables and syllable qualities) is abstracted far enough from the text, and if sufficient data is available. More pointed: by abstracting from graphemes to syllables, the traditional nomalization techniques of normalizing 'Vokallänge' (vowel length), 'Vokalqualität' (vowel quality), 'Auslautverhärtung' (final fortition), and 'Konsonantenschreibung' (consonant orthography), as performed for example by the new *Lyrik des Deutschen Mittelalters* (LDM) online project[38] and other methods of standardization, do not significantly affect the target statistics for the analyses in this project. The main interference for syllable analysis is the handling of schwas. While the small exploration below will address normalization, a later discussion of prosodic clustering will attempt to show the effect of editing practices on a large scale.

Due to the lack of diplomatic transcriptions, to test this hypothesis I begin by using a very small selection from the LDM project, which has already completed much work in transcription and normalization. I restrict myself to the Codex Manesse as a focus, admittedly a less noisier manuscript than others, but also representing a far greater proportion of the MHG lyric corpus than any other manuscript.[39] Within the Codex Manesse, the LDM project has completed all of the strophes attributed to Dietmar von Aist (42), Rubin (68), and Der wilde Alexander (35). Fortunately, these three also provide significant breadth coverage within the MHG lyric corpus itself. Across these texts I compare the diplomatic transcription with the normalized MHG, with special attention to the features I use in my subsequent analyses[40], the example below is from Dietmar (C 7)[41]:

> Ahy nv kvmt vns dů zit der kleinē vogelline ſang
> es grůnet wol dů linde breit zergangē iſt der wintˇ lanc
> nv ſiht man blůmē wol getan an der heide vbent ſi ir ſchin

[38]The LDM project is made available by Manuel Braun, Sonja Glauch, and Florian Kragl at `http://www.ldm-digital.de`

[39]Within MHG manuscripts, dialect variation is often muted, as it becomes written in the dialect of the scribe. While I do not address dialect directly here, and we do not know exactly how *traveling* poets grappled with dialect variation (Rüdiger Schnell argues poets must have adapted at least their rhyme. Schnell, "Prosaauflösung und Geschichtsschreibung im deutschen Spätmittelalter," 215.), I simply take the dialect of the text from the manuscript in which it was written as a "trace" of evidence following the recent trend of MHG scholarship.

[40]It may be helpful to return to this section after reading Chapters 4, 5, and 6, and a better understanding of methodology is acquired.

[41]Unless otherwise noted, for the duration of the section the text is sourced from the LDM project.

des wirt vil manig h˜ze fro deſ ſelbē troſtet ſich dc min

Ahy[42], nu kumt uns diu zît, der kleinen vogellîne sanc.
ez grüenet wol diu linde breit, zergangen ist der winter lanc.
nu siht man bluomen wol getân, an der heide üebent si ir schîn.
des wirt vil manic herze frô, des selben trœstet sich daz mîn.[43]

We see that both stanzas have the exact same syllable count (46), each line has the same number of syllables, and the exact same open and closed ratios, as the sequences are kept the same.[44] No changes needed to be made for grammar, and the only changes made were restricted to individual phonemes, not groups of phonemes. While the summary statistics hint that this is not uncommon, there are certainly instances where the two do not line up as well, or *mouvance* introduces further complications. The above stanza is also trasmitted in manuscript B 7.

Hei nv kvmet v̇ns dv̇ zit der clainen vogellinē ſang
es grv̇net wol dv̇ linde brait zergangen iſt d˜ wint˜ lang
nv ſiht man blv̇men wol getan an d˜ haide v̇bent ſv̇ ir ſchin
des wirt vil manig h˜ ze vro des ſelben troſtet ſich das h˜ ze min

Hei, nu kumet uns diu zît, der kleinen vogellînen sanc.
ez grüenet wol diu linde breit, zergangen ist der winter lanc.
nu siht man bluomen wol getân, an der heide üebent si ir schîn.
des wirt vil manic herze vrô, des selben trœstet sich daz herze mîn.

Both the manuscript and the normalization correspond to a high degree again.[45] While manuscript B does differ slightly, namely 'Ahy' → 'Hei', 'kumt' → 'kumet', 'si' → 'su', and the addition of 'herze' as the second to last word in manuscript B. The abstracted statistics show them to be quite similar: manuscript B has 66 syllables with 33.33 percent open, while manuscript C has 64 syllables with 34.38 percent open. The Levenshtein ratio for the characters between the normalized editions, which have shown to follow the manuscript quite closely, is .964.[46]

A slightly more difficult case is presented in Dietmar's *Man sol die biderben unde die guoten*, surviving in C 9 and B 9:

[42] *MF* 33,15 'Ahî'

[43] "Yipee! Now comes the time for us, during which the small birds sing. The wide linden trees are greening already, the long winter is gone. Now you can see beautiful flowers, which let themselves shine on the heather. Many hearts become happy because of this, and my heart also finds comfort in it."

[44] 64 syllables, 34.38% open syllables, Levenshtein ratio for sound patterning (See Chapter 4) is 1.0. The Levenshtein ratio defined by the Levenshtein distance divided by the alignment length. The Levenshtein distance is the least number of edits to one string of characters to match the other string of characters.V. I. Levenshtein, "Binary Codes Capable of Correcting Deletions, Insertions and Reversals," *Soviet Physics Doklady* 10 (February 1, 1966): 707

[45] 66 syllables, 33.33% open syllables, Levenshtein ratio for sound patterning (See Chapter 4) is 1.0

[46] The Levenshtein ratio between sound sequences is .969.

Man ſol die biderben vn̄ die gv̊ten zallen ziten haben w˜t
ſwer ſich gerv̊met alze vil der hat der beſten mâſſe nihtgegert
io ſol es niemer hov̊feſcher man gemachen allen wiben gv̊t
er iſt ſin ſelbes meiſter niht ſwer ſin alze vil getv̊t

Man sol die biderben unde die guoten zallen zîten haben wert.
swer sich gerüemet alze vil, der hât der besten mâze niht gegert.
jô sol ez niemer hövescher man gemachen allen wîben guot.
er ist sîn selbes meister niht, swer sîn alze vil getuot.

We see from manuscript MHG to normalized MHG only a minor change affecting a syllable analysis.[47] Yet the *MF* edition draws from manuscript B:

Man ſol die biderben vn̄ die gv̊ten ze allen ziten haben liep
ſwer ſich gerv̊met alzevil der kan der beſten maſſe niht
ioch ſol es niem˜ hoveſch˜ man gemachen allen wiben gv̊t
er iſt ſin ſelbes maiſt˜ niht ſwer ſin alze vil getv̊t

Man sol die biderben und die guoten ze allen zîten haben liep.
swer sich gerüemet alze vil, der kan der besten mâze niet.
joch sol ez niemer hövescher man gemachen allen wîben guot.
er ist sîn selbes meister niht, swer sîn alze vil getuot.[48]

No changes except for word boundaries between the manuscript and the normalized MHG[49], yet between manuscript B and C there is larger difference, at least in semantics. The end rhyme for the first two verses is entirely different, though the meaning is related. We see the Levenshtein ratio between normalized MHG editions drop to .915, still high, but quite a difference from the comparison of the stanzas above. Yet my focus here is on the difference in syllable measurements and qualities. A two syllable difference is minimal, and the percent of open syllables also changes only slightly, about two percent.[50]

[47]Both have 68 syllables, though the manuscript has 30.88% open while the normalized has 32.35%, and a Levenshtein ratio for the sound patterning of .985, due *mâsse* to mâze. In the first phase of the second Germanic consonant shift Germanic /t/ shifted to /ss/, and although there is orthographical variation for this phoneme, normalized MHG often denotes this as 'z', or occasionally 'zz'. The long fricative suggests splitting the consonants between syllables, i.e., mâs-se, not mâ-sse. The correct syllabification thus depends on one's view of ambisyllabicity for certain German fricatives, a problem not yet settled. The normlized orthography argues against ambisyllabicity, and that the fricative is short, necessitating a syllabification of mâ-ze.

[48]*MF* 33, 31-34. "Those which are esteemed and good, they should be supported all the time. Whoever brags too much, he doesn't understand the right measure. Yet no courtly man should make it right to all women. He's no longer his own master, if he does too much."

[49]Both have 66 syllables, though the manuscript has 28.79% open while the normalized has 30.30%, and a Levenshtein ratio for the sound patterning of .985, due *mâsse* to mâze.

[50]Perhaps more importantly for the sequential quality, the Levenshtein ratio for the sound sequence at .970 is greater than the stanzas above, implying that a sequential syllable quality sequence is actually more alike between manuscripts A and B for Dietmar 9 than A and B for Dietmar 7.

Finally, as an example of error correction, we have Dietmar 40 C, MF 40, 35, not transmitted in other manuscripts:

Was wiſſet mir der beſte mā
ich habe ime leides niht getan
er froͤt ſich ane ſchulde
dc er in hat vōn mir geſeit
dc iſt mir hüte vn̄ iemer leit
er v˜lüſet mine hulde
mir wirret niht ſin boͤſer kib
wc half dc er toͤrſchē bi mir lag ion en wart ich nie ſin wib[51]

Waz wîzet mir der beste man?
ich habe ime leides niht getân,
er fröut[52] sich âne schulde.
daz er in hat[53] von mir geseit,
daz ist mir hiute unde iemer leit.
er verliuset mîne hulde.
mir wirret niht sîn bœser kîp.
waz half, daz er tœrschen bî mir lac? jô enwart ich nie sîn wîp.[54]

'ion' in the manuscript is normalized to 'jô', as it appears the scribe anticipated the negation of 'en' too soon. While this does change what the original manuscript presents, it does so only slightly.

I focus on syllable count, syllable count per line, open and closed syllable ratio, and open and closed syllabe sequencing, as these are the measures used in the following chapters. As the Table 2.1 shows, an abstraction at the syllable level appears to retain many of the phonological features of the transcription, including correction for errors. What this table does *not* show are statistics for editions combining several manuscripts. Yet this *Leithandschrift* (leading/guiding manuscript/hand) technique affects primarily stanza ordering, though occasionally word choice as well. Furthermore, many would argue that the phenomenon of *mouvance* actually creates new songs in differing manuscripts, and that these should not be considered the same anyway. If anything, a syllable analysis would show a high degree of similarity in sound and rhythm, whether they share the same content or not.

At this time, the data does not yet exist to sufficiently test these comparisons on a large scale, but the recently published ReM corpus presents itself once again as a useful

[51] Fridrich Pfaff and Baden (Germany)., *Die grosse Heidelberger Liederhandschrift, in getreuem textabdruck* (Heidelberg: C. Winter, 1898), 185.

[52] *MF* vröut

[53] *MF* hât

[54] "What does the best man know about me? I have not caused him pain, he is happy without guilt. That he has fallen for me, that is for me today and always sad. He wastes my grace. His mean scolding doesn't bother me.

	stanzas	sound diff.	syll. diff.	Lev. ratio
Dietmar von Aist	42	.79% ± 1.00%	.469 (3)	.988 (.970)
Rubin	68	.82% ± .89%	.535 (4)	.988 (.957)
Der wilde Alexander	35	1.09% ± .99%	.543 (4)	.985 (.943)
lyric total	145	.86% ± .95%	.516	.987
Parzival (D)	500 verses	-	-	.940

Table 2.1: Comparison of diplomatic transcriptions and normalized editions in manuscript C, weighted by syllable count in stanza

source for comparing diplomatic transcriptions and normalized texts. Though containing few texts from the popular literary corpus, it does have *Parzival* from manuscript D, with both a diplomatic and normalized transcription. Thus I undertake the above Levenshtein ratio calculation on the first 500 verses of *Parzival* manuscript D compared to *Parzival* as edited by Albert Leitzmann, who used manuscript D as the *Leithandschrift*.[55] The results are comparable to the lyric above. The mean Levenshtein ratio of the 500 lines is .966. 264 of the 500 verses have the exact same syllable sequence patterning. The Levenshtein ratio of the entire sequence is .965.[56] The actual statistics may be higher, as the one abbreviation difficult to algorithmically simplify is the n̄, indicating for one to fill in the correct ending, which the editor has done. The first ten lines compared are reproduced below:

Swaz da waſ volcheſ inne
more vn̄more
waſ beidiv wip vn̄man
der herre ſchowen began
manegen ſchilt zebrochen
mit ſpern gar dvrchſtochen
der waſ da vil gehangen fvr
an die wende vn̄an die tvr
ſi heten iamer vn̄gvft
in div venſter gein dem lvft

swaz da was volkes inne
moere und moerinne
was beidiu wîp unde man
der herre schouwen began
manegen schilt zebrochen

[55]Albert Leitzmann, ed., *Wolfram von Eschenbach*, Altdeutsche Textbibliothek, Nr. 12-16 (Tübingen: M. Niemeyer, 1960), IV.

[56]This is used for the 10-gram sampling in Chapter 4.

mit spern gar durchstochen
der was dâ vil gehangen vür
an die wende und an die tür
si heten jâmer unde guft
in diu venster gein dem luft[57]

Therefore, although we lack sufficient data to fully test the discrepency between manuscripts and edited editions resulting from the methods employed in the following analyses, I believe it may change little, and as argued by Zumthor and Bumke, the manuscripts are not the exact performance either. Initial results imply that 1) normalization of MHG from the manuscript to the edited edition (used in my subsequent analyses) is not dramatic enough to severely derail a corpus-wide syllable analysis. On average, the variation is only around 2%, which, if implemented correctly, would not have far-reaching effects, and 2) while this claim must be tested further, slight realization of *mouvance* may still keep the general soundscape of any given stanza, in fact, if it is labeled the same stanza, it likely *must* keep a similar soundscape (otherwise the words would vary enough to be considered a different stanza). If it does not, then it must be questioned as to whether these are derived from the same stanza in the first place.

By abstracting to the syllable, this project will adopt a new approach to computational text analysis in hopes to implement a viable, corpus-wide analysis. The method in the following pages will strive to 1) move beyond issues in orthography and pure semantics into large and small scale aesthetic analyses with semantic implications, 2) artificially increase corpus size by analyzing a level of the text below the word, and 3) repeatedly return to close readings during and between analyses in order to point out manifestations on the level of the text and enter the discussion in the relevant scholarship.

2.5 MHG Corpus for Subsequent Analysis

The main corpus for subsequent analysis has been generously made available by the *Mittelhochdeutsche Begriffsdatenbank* (MHDBDB), started by Klaus Schmidt in 1992.[58] Since then, the MHDBDB has gradually broadened not only its corpus of texts, but its toolbox for text analysis. The MHDBDB is currently curated by a group of researchers in the United States, Germany, and Austria, though contributions have been made by many others around the world. The MHDBDB is the largest collection of digitized MHG texts, containing a range of authors, genres, and themes. Though not utilized here, the MHDBDB also offers a range of tools to interact with these texts, including word querying with regular expressions-like

[57]"For whatever the people therein (the town), they were black poeple, both women and men. The lord began to see many broke shields pierced by spears. There were many hanging on the walls and on the doors. There was clamoring and shouting, in the windows came the air!"

[58]Full information about the MHDBDB, its history, corpus, and analytical tools, see its website `http://mhdbdb.sbg.ac.at`

syntax, lemmata data, and a conceptual index, in which words are mapped to concepts. The MHDBDB is unique in its collection of texts, consisting of a variety of genres, but primarily MHG verse texts. It is comprised of mostly published edited texts, many of which are normalized, some remaining in a more dipolomatic transcription. The MHDBDB comprises 659 texts, most of which are also fully lemmatized. Many texts are also supplied with relevant metadata, such as assumed author, approximate date, and genre.[59]

When working with the MHDBDB, we must remember that these are texts *chosen* to be transcribed and added to the database, and if a *Leithandschrift* technique was implemented, other manuscripts for a song or poem are simply not represented. While for this case it may be better to be unrepresented than a mixture of all manuscripts, future work must include as many texts as possible, all holding as true to a manuscript as possible, as the texts in the LDM project do. Because MHDBDB is a mixture of standardized and non-standardized MHG texts, despite these orthographic and dialect differences, it is important to remember that the features of interest are identified based on orthography, not the exact lexicon entry, and thus susceptible to editorial practices. For example, an ulterior goal in the following chapters is to show that the following group of stanzas from the *Nibelungenlied* is measured as similar, if not the same, soundscape:

> Manuscript A[61]
> Uns ist in alten meren wunder vil geseit
> von helden lobeberen von grozzer arebeit
> von vreuden und hochgetziten von weinen und von chlagen
> von chuener recken striten mugt ir wunder hoeren sagen[62]
>
> Manuscript B/C[63]
> Uns ist in alten mæren wvnders vil geseit
> von heleden lobebæren von grozer arebeit
> von freude und hochgeciten von weinen und klagen
> von kuner recken striten muget ir nu wvnder horen sagen

[59]Until recently, the MHDBDB was the only corpus for MHG. In December of 2016, the *Referenzkorpus Mittelhochdeutsch* (ReM) (see above) was published.[60] As a corpus more oriented for linguistic purposes, and a combination of Das Bochumer Mittelhochdeutschkorpus (BoMiKo) and the Korpus der Mittelhochdeutschen Grammatik (MiGraKo), the REM focuses on early MHG texts, primarily prayers, blessings, and scripture. The incredible contribution of the REM is that each text is fully annotated, with diplomatic transcriptions and normalized annotations. Yet while deserving a mention here, it is not used in this project except for reference, due to its low overlap with the verse texts of interest, and those of the MHDBDB.

[61]Michael S. Batts et al., eds., *Das Nibelungenlied. Paralleldruck der Handschriften A, B und C nebst Lesarten der übrigen Handschriften* (Tübingen: Niemeyer, 1971).

[62]We are told in old stories many wonderful things. Of famous heroes, great hardship, of fortunate days and festivals, of tears and lamentations and of the battling of brave warriors, may you now hear about their amazing achievements.

[63]Karl Bartsch and Helmut de Boor, eds., *Das Nibelungenlied*, 20th ed., Deutsche Klassiker des Mittelalters 3 (Wiesbaden: Brockhaus, 1972).

Manuscript C[64]
Uns ist in alten mæren wunders vil geseit
von helden lobebærn von grozzer chůnheit
von frouden hochgeziten von weinen und von klagen
von chůner rechen strite muget ir nu wunder hören sagen

While the orthography and editing practices may differ, the patterning of the syllable sounds and rhythm is very similar. This sort of comparison can be conducted on all verse texts, not restricting to one tradition such as the *Nibelungenlied* above. In this manner, in the following chapters I will attempt to group, categorize, and close-read these subtle features in hopes to draw conclusions previously unnoticed, or to support conclusions quantitatively, which only had rough indications prior.

[64]Ursula Hennig, ed., *Das Nibelungenlied: nach d. Hs. C*, 1. Aufl, Altdeutsche Textbibliothek 83 (Tübingen: Niemeyer, 1977).

Chapter 3

Soundscapes

In the last decades of the 20th century, most applications of computers to literature produced the same statistics won through careful counting by hand previously, albeit at a scale previously inaccessible. Recently, modern large-scale investigations of literary form have led researchers to newer methods that have been broadly labeled as the 'Digital Humanities', a concept yet to be consistently defined, but one aspect of which includes tools for computational text analysis. This particular aspect has been called a revival of Russian Formalism and its systematic approach to language and literary theory. Roman Jakobson and Victor Shklovsky would certainly be in awe at the progress Natural Language Processing (NLP) has made over the recent years, especially concerning its potential for impact on literary analysis. Many efforts in the Digital Humanities for literary texts currently focus on either technical linguistic models or large-scale "distant reading" analyses as popularized by Franco Moretti.[1] Technical contributions often enhance or adapt part of speech (POS) taggers, named entity recognition (NER) taggers, or word lemmatizers.[2] "Distant reading" has been mostly associated with topic modeling or clustering large corpora, literary network theory, and genre formation or identification.[3]

In the next two chapters I investigate MHG soundscapes employing the technical contributions made in the previous chapter for both a small and large scale analysis of MHG poetry. This is done in hopes to reconcile the abstracted distant reading of trends with traditional close reading techniques in order to help better understand the individual text or song. In this chapter I quantify phonological soundscapes by aggregating the syllable properties in a given stanza, and correlate these different auditory experiences with voice.[4]

[1]Franco Moretti, *Distant reading* (London: Verso, 2013).

[2]Delphine Bernhard and Anne-Laure Ligozat, "Es esch fàscht wie Ditsh, oder net? Étiquetage morpho-syntaxique de l'alsacien en passant par l'allemand," *Actes de TALARE 2013: Traitement Automatique des Langues Régionales de France et d'Europe*, 2013, 209–220; Delphine Bernhard and Anne-Laure Ligozat, "Hassle-free POS-tagging for the Alsatian dialects," *Non-Standard Data Sources in Corpus Based-Research*, 2013, 85–92.

[3]Franco Moretti, "Network Theory, Plot Analysis," *New Left Review*, II, no. 68 (April 2011): 80–102.

[4]Tanya Clement's study of applause serves as a methodological contrast for a parallel modern phenomenon.Clement and McLaughlin, "Measured Applause: Toward a Cultural Analysis of Audio Collections"

3.1 Close Reading Soundscapes

Let us first read aloud the following stanzas from Reinmar, cognizant of the fact that each stanza is excerpted from a different song. The diplomatic transcription is included for reference.

> Dêst ein nôt, daz mich ein man
> vor al der werlte twinget, swes er wil
> sol ich, des ich niht enkan,
> beginnen, daz ist mir ein swaerez spil
> Ich hât ie vil staeten muot
> nu muoz ich leben als ein wîp,
> diu minnet und daz angestlîchen tuot.[5]

> Swenne ich sî mit mîner valschen rede betrüge,
> sô het ich sî unreht erkant
> und gevâhe sî mich iemer an deheiner lüge,
> sâ sô schupfe mich zehant
> Und geloube niemer mîner klage,
> dar zuo niht, des ich sage
> dâ vor müeze mich got behüeten alle tage.[6]

> D eſt ein not das mich ein man.
> vor al d˜ w˜lte twinget ſwes er wil.
> ſol ich deſ ich niht enkan.
> beginnen dc iſt mir ein ſweres ſpil.
> ich hat ie vil ſtetē mv̊t.
> nv mv̊s ich lebē als ein wib
> dv̇ mīnet vn̄ das angeſtlichen tv̊t.[7]

Modern poetry readings are indeed a modern form of at least what we believe the *Minnesänger* to have been performing in most cases. Clement takes advantage of existing recordings to emphasize the relationship between the audience and the performer. Lacking such recordings, I elect to investigate this relationship auditory patterns crafted by the poets.

[5] *Des Minnesangs Frühling* (*MF*) 192, 25-214C. "It is distressful that a man may force me before all the world to do as he wishes. Am I to begin things that I cannot, that is for me a difficult game. I've always had steadfast courage. Now I must live as a woman who loves, and does so in fear." All translations are the author's own unless otherwise noted.

[6] *MF* 173, 13-96C, 53b. "If I were to ever deceive her with my false words, then I would have valued her improperly. And if she were to ever catch me in any lie, then she would immediately shake me up and never believe my lamentation, moreover anything I say. May God protect me from that every day."

[7] Pfaff and Baden (Germany)., *Die grosse Heidelberger Liederhandschrift, in getreuem textabdruck*, 197.

> Swēne ich ſi mit miner valſchen rede betruge.
> ſo het ich ſi vn reht erkant.
> vn̄ gevahe ſi mich iemer an deheiner lvge.
> ſa ſo ſchvpfe mich zehant.
> vn̄ gelo͛be niemer miner klage.
> dar zo͛ nicht deſ ich ſage.
> da vor mv͛ſſe mich got behv͛ten alle tage.[8]

It is difficult to not be immediately drawn to the unique soundscape, rhythm inclusive, of each stanza. *Dêst ein nôt* is marked by closed and heavy syllables, which add to the overall consonant-heavy phonemic soundscape that generates a particular sound patterning certainly influencing the possible musical and rhythmical settings. In contrast, *Swenne ich sî* is populated by many open syllables, even in the closed syllable cadences. By virture of the mouth ending a unit of speech in an open position, this produces a very different sound patterning than that exhibited by *Dêst ein nôt*. Its consistent alternation of open and closed syllables forces a more metered, even iambic feel. While the music setting for these songs have not survived, we can gather from the phonology of the stanzas that they must have employed different soundscapes, not only linguistically, but also musically. The great French philosopher Jean-Jacques Rousseau spilled a tremendous amount of ink over phonology and music. I quote in length here from Rousseau's 'Letter on French Music' for his detailed, yet entertaining, description:

> One can conceive of some languages as being more appropriate to Music than others; one can conceive that some would not at all be so. Such might be one that was composed of only mixed sounds, of mute, indistinct, or nasal syllables, few sonorous vowels, many consonants and articulations, and that further lacked other essential conditions of which I shall speak under the subject of meter. Out of curiosity, let us try to find what would result from a Music applied to such a language.
>
> First, the want of vividness in the sound of the vowels would require giving a great deal to that of the notes, and because the language would be indistinct, the Music would be piercing. In the second place, the harshness and the frequency of the consonants would force one to exclude many words, to proceed on others only by elementary intonations, and the Music would be insipid and monotonous, its progress would also be slow and tiresome for the same reason, and if one wanted to press the movement a little, its haste would resemble that of a heavy and angular body rolling along on cobblestones.[9]

[8] Pfaff and Baden (Germany)., *Die grosse Heidelberger Liederhandschrift, in getreuem textabdruck*, 185.

[9] Jean-Jacques Rousseau, John T. Scott, and Jean-Jacques Rousseau, *Essay on the origin of languages and writings related to music*, vol. 7, The collected writings of Rousseau (Hanover, N.H.: University Press of New England, 1998), 144.

Germanic languages, including MHG, certainly fit Rousseau's description as a language phonologically "less appropriate". Although Rousseau's description is surely exaggerated, many writers after him have shared his sentiment that vowels and open syllables are crucial to the composition of sung music and may even correlate with melismatic syllables.[10] Yet Rousseau's hypothetical application can also shed light on the formal play in MHG lyric. Rousseau claims an unfit language would be "force[d]" to exclude many words, and it would become "monotonous", exactly what Cramer observes in the content of MHG lyric before the shift in the 13th century. Motifs are constantly rehashed, and Kuhn claims that nothing new was contributed after the shift. Reading the MHG formal scholarship into Rousseau, it may appear that a vocabulary bound by the musicality of a language (this restriction we also see in consistent rhyme pairs throughout the MHG corpus[11]), certainly in addition to the natural development of the genre, led to formal play in the 13th century in order to break these restrictions and "press the movement a little". When we arrive at Konrad von Würzburg's infamous rhyme poem (*Schlagreimlied 26, I*), in which every word is rhymed, we get Rousseau's sense of exactly "a heavy and angular body rolling along on cobblestones."

> Gar bar lît wît walt,
> kalt snê wê tuot: gluot sî bî mir.
> gras was ê clê spranc
> blanc, bluot guot schein: ein hac pflac ir.
> schoene doene clungen jungen liuten.
> triuten inne minne mêrte:
> sunder wunderbaere swaere wilden
> bilden heide weide rêrte,
> dô frô sâzen die,
> der ger lâzen spil wil hie[12]

Konrad's poem has signicantly more open syllables than Reinmar's heavy and closed verses above (also straying from a typical MHG prosodic characterization), but its heavy syllables are often heavy by a vowel, not a consonant syllable ending. Braun calls this song a "meta-text" because both rhyme and rhythm are clearly thematized, which are central to lyrical poetry in general.[13] Braun claims there cannot be rhythm if rhyme words continually follow

[10]See forthcoming work by Murray Schellenberg, 'Influence of Syllable Structure on Musical Text Setting.' Ross et al. Ross, Choi, and Purves, "Musical intervals in speech" suggest that preferred musical intervals are related to formants in vowel phones.

[11]Christopher Hench, "Alienated Bodies and the lîp/wîp Connection," *Interdisciplinary Journal for Germanic Linguistics and Semiotic Analysis* 21, no. 2 (2016): 165–189.

[12]Konrad 26, I. "The forest is left largely bare, cold snow hurts: May warmth stay with me. There was once grass, clover grew light, the blossoms grew well: A hedge protected them. Beautiful sounds resounded from young people, love bred inside intimacy: especially wonderful heavy wild form heather falling afar, there sat happily those who want to let the game here desirously."

[13]Manuel Braun, "Aufmerksamkeitsverschiebung. Zum Minnesang des 13. Jahrhunderts als Form- und Klangkunst," *Wolfram-Studien*, Veröffentlichungen der Wolfram von Eschenbach-Gesellschaft, 21 (2013): 229.

one another. In end rhyme they usually punctuate the end of a unit, which is the reason why *clausulae* grew in importance for medieval music and poetic practice. Thus Konrad intentionally works against rhythm. He writes a poem in which form outweighs content and makes an attempt at rectifying MHG's typical sonority boundaries. In contrast to MHG, Rousseau sees Italian as having one of the greatest phonemic inventories for this reason:

> It is sweet because its articulations are not very compounded, because in it the grouping of consonants is rare and without roughness, and, since a great number of syllables are formed of vowels alone, the frequent elisions make its pronunciation more flowing. It is sonorous because the majority of its vowels are bright, and because it has no compounded diphthongs, because it has few or no nasal vowels, and because its rare and smooth articulations better distinguish the sound of the syllables, which become clearer and fuller because of it. With regard to the harmony, which depends on number and prosody as much as on sounds, the advantage of the Italian language is manifest on this point; for it must be remarked that what makes a language harmonious and genuinely picturesque depends less on the real force of its components than on the distance there is from soft to loud among the sounds it employs and on the choice that can be made for the portraits to be painted.[14]

Rousseau's 18th century idea of musical language has certainly developed since. A century later, the well-known pianist Charles Kensington Salaman wrote an article for the Musical Association entitled 'On the English Language as a Language for Music', in which he argues against the apparently widely held notion at the time that "our language [English] is unfitted for music", though admitting it is not the "best" suited either.[15] Salaman is prompted to action in this article by 18th century writer Joseph Addison, and both are primarily fixated on opera. Salaman writes:

> It is suggested that the singing voice requires a language whose verse contains a majority of words formed of open syllables—free for vocal utterance. It must be allowed that the minority of words of this class is amongst the inherent difficulties of our language as a language for music.[16]

Yet despite this state of affairs, which Salaman and Addison deem as a handicap, Salaman argues that the great lexical breadth of the English language more than makes up for these shortcomings. Salaman also believes that music requires rough sounds as well, "to express and describe rage, frenzy, revenge, despair, the natural war of the elements, and the unnatural war of human beings."[17]

[14]Rousseau, Scott, and Rousseau, *Essay on the origin of languages and writings related to music*, 148.

[15]Hon. Mem. Acad. Charles Kensington Salaman Esq., "On the English Language as a Language for Music," *Proceedings of the Musical Association* 3, no. 1 (January 1, 1876): 121.

[16]Ibid., 123.

[17]Ibid., 124.

This project does not intend to take any side on the 'musicality' of any language, nor the suitability of MHG for musical setting. Some, if not most, MHG poems were set to music, all ranging in sonority and complexity. However, what is critical to this project is *relative* differences in this sonority, which is where the scholarship on what is known as 'rhythm metrics' has arrived.[18] Since the mid-20th century, language has traditionally been divided into either stress-timed or syllable-timed languages.[19] Because of syllable structure and vowel reduction "some syllables are far more salient than others in stress-timed languages, and [...] all syllables tend to be equally salient in syllable-timed languages. This in turn, creates the impression that there are different types of rhythm."[20] Ramus et al. focus on percentage of vocalic intervals in all speech and duration of consonantal intervals: "A vocalic interval is located between the onset and the offset of a vowel, or of a cluster of vowels. Similarly, a consonantal interval is located between the onset and the offset of a consonant, or of a cluster of consonants. The duration of vocalic and consonantal intervals adds up to the total duration of the sentence."[21]

A second approach emphasizing vocalic alternation, proposed by Grabe and Low in 2002, entails measuring successive accoustic-phonetic intervals with Pairwise Variability Indices (PVI), specficially the $rPVI$ (raw PVI) and $nPVI$ (normalized PVI).[22] These two metrics measure vowel duration and interval duration between vowels.[23] Because vowel duration and interval duration between vowels is heavily tied to syllable structure, Grabe and Low's research support claims made in 1982 by Dasher and Bolinger, who argue that rhythm is more determined by syllable structure and vowel reduction than by their classification as syllable-timed or stress-timed.[24]

These measures of rhythm and prosody have also been integrated into studies of music. How does the prosodic structure of a language affect its musical tradition? By comparing prosodic structure with sung music and poetry, Palmer and Kelly argue that "Sung verse allows the amplification or exaggeration of prosodic patterns present in normal speech..."[25] Their explanation is that both musical and prosodic structures rely on the principle of alternation, as evidenced above by Dasher and Bolinger as well as Grabe and Low. Palmer and Kelly suggest an extreme view that "the rhythms underlying music and language are

[18]Amalia Arvaniti, "Rhythm, Timing and the Timing of Rhythm," *Phonetica* 66, no. 1 (April 2009): 46–63.

[19]D. L. Bolinger, "Pitch accent and sentence rhythm," in *Abe, Kanekiyo, Forms of English: accent, morpheme, order* (Cambridge: Harvard University Press, 1965), 139–180.

[20]F. Ramus, M. Nespor, and J. Mehler, "Correlates of linguistic rhythm in the speech signal," *Cognition* 73, no. 3 (December 17, 1999): 268.

[21]Ibid., 271.

[22]E. Grabe and E. Low, "Durational variability in speech and the rhythm class hypothesis," *Papers in Laboratory Phonology* Vol. 7 (January 1, 2002): 515–546.

[23]Ibid., 8.

[24]Richard Dasher and Dwight Bolinger, "On pre-accentual lengthening," *Journal of the International Phonetic Association* 12, no. 2 (December 1982): 58–71 For more see Aniruddh Patel's *Music, Language, and the Brain*Aniruddh D. Patel, *Music, language, and the brain* (New York: Oxford University Press, 2008)

[25]Carloline Palmer and Michael Kelly, "Linguistic prosody and musical meter in song," *Journal of Memory and Language* 31, no. 4 (August 1992): 539.

identical morphologically, in that they have the same roots but have evolved to serve different roles."[26]

While these measures are not perfect, and are certainly bound to our own perceptions of time and rhythm[27], for the purposes of this project there are two key points: 1) vowel duration and alternation, influenced by syllable structure, are important factors in prosodic rhythm and the composition of music, and 2) NHG, as well as its predecessor MHG[28] are particularly affected by these intervals in comparison to the romance languages due to signficant vowel reduction and a larger set of syllable structures.

If MHG soundscapes can vary so much, as evidenced above, and are correlated with musical possibility, are there thematic relations inherent in these soundscapes? In the Reinmar pairing it is further noteworthy that *Swenne ich sî* is a *Frauenstrophe*, underlining the assumed position of the speaker. Reinmar was not the only one to distinguish speakers through soundscapes:

> Owê,
> sol aber mir iemer mê
> geliuhten dur die naht
> noch wîzer danne ein snê
> ir lîp vil wol geslaht?
> der trouc diu ougen mîn.
> ich wânde, ez solde sîn
> des liehten mânen schîn.
> dô tagte ez.[29]

> Owê,
> sol aber er iemer mê
> den morgen hie betagen,
> als uns diu naht engê,
> daz wir niht durfen klagen:
> owê, nu ist ez tac,
> als er mit klage pflac,
> dô er jungest bî mir lac?

[26]Palmer and Kelly, "Linguistic prosody and musical meter in song," 540.

[27]Arvaniti discusses how these measures are not as accurate for non-prototypical languages, and are further bound by perceptions of time.Arvaniti, "Rhythm, Timing and the Timing of Rhythm," 46

[28]Although vowels had been signficantly reduced since Germanic and Old High German, MHG had not yet reduced as far as NHG.

[29]*MF* 143,22-93C ff. "Alas, shall it ever again shine to me through the night, whiter yet than snow—her body, so well formed? It deceived my eyes: I thought it must be the gleaming of the bright moon. Then the dawn came." Jonathan Saville, *The Medieval Erotic Alba: Structure As Meaning* (iUniverse, 1999), 264 Statistics: 8 light syllables, 43 heavy syllables, 15.69% light. 18 open syllables, 33 closed syllables, 35.29 % open.

dô tagte ez.[30]

Owê,
si kuste âne zal
in dem slâfe mich.
dô vielen hin zetal
ir trehene nider sich.
iedoch getrôste ich sie,
daz si ir weinen lie
und mich alumbe vie.
dô tagte ez.[31]

Owê,
daz er sô dicke sich
bî mir ersehen hât!
als er endahte mich,
sô wolt er sunder wât
mîn arme schouwen blôz.
ez was ein wunder grôz,
daz in des nie verdrôz.
dô tagte ez.[32]

Heinrich distinguishes the last two stanzas by soundscape. As the sexual tension grows from the beginning of the song, the contrasting soundscapes support the differing maculine and feminine perspectives. Both figures shift the responsibility for their actions onto the other. In the last two stanzas, the male voice highlights the kissing and embracing desired by her, while the female voice reacts to the insatiability of the man. The punctuated female response underlines her sentiment of how she sarcastically describes the situation.

Unfortunately, because we can no longer measure speech intervals of native MHG speakers as rhythm metrics scholars might, we have to use only features from the linguistic evidence that remains. Syllables provide a level of measurement above the phoneme, and an oppor-

[30] "Alas, shall he ever again stay here in the morning?—so that when the night leaves us we shall not have to lament, 'Alas, now it is day,' as he lamented the last time he lay by me. Then the dawn came." Saville, *The Medieval Erotic Alba: Structure As Meaning*, 264 Statistics: 10 light syllables, 42 heavy syllables, 19.23% light. 20 open syllables, 32 closed syllables, 38.46 % open.

[31] "Alas, she kissed me numberless times as I lay sleeping. Then her tears fell down and down. But I comforted her so that she left off her weeping and embraced me all around. Then the dawn came." ibid. Statistics: 17 light syllables, 32 heavy syllables, 34.69% light. 29 open syllables, 20 closed syllables, 59.18 % open.

[32] "Alas, that he has so often looked upon me! When he took the covers off me he wanted to look at poor me naked, without clothing. It was a great wonder that he never grew bored with this. Then the dawn came." ibid., 264-265 Statistics: 6 light syllables, 42 heavy syllables, 12.5% light. 13 open syllables, 35 closed syllables, 27.08 % open.

tunity to understand these soundscapes. Because MHG poets often structured their poems by stanza, I sum the sonority of a stanza rather than an entire song. Thus the percentage of open syllables in a stanza is calculated. A large difference appears in open and closed syllables, and as the long vowels are marked, the difference in light and heavy syllables is also evident. It is no coincidence that Heinrich uses this sonority difference to highlight the *Wechsellied* (alternating song) characteristics of the poem. This sound is a feature that largely goes unnoticed by modern readers, as these songs are no longer performed or read aloud. Braun, describing Ulrich von Liechtenstein's *Wizzet frouwe wol getan*, similarly shows a clear distinction between male and female voices in form: "The man, who speaks the first stanza, demostrates his abilities by using only a single rhyme. In contrast, the woman in the second stanza does not rhyme at all. This makes her speech appear artless and inferior to the masculine monorhyme."[33] Thomas Cramer illuminates the same example, emphasizing that such formal mastery would not have been understood by the average audience member, especially not in a purely oral tradition.[34]

3.2 Formal Debates

How can a syllable analysis of MHG shed light on MHG lyric? As this project is foremost interested in sound, a description of the current state of scholarship on *Klanglandschaften* (soundscapes) in MHG lyric is necessary. Although an engagement of the scholarship with *Minnesang*'s formal play has been ongoing for centuries, the idea that medieval German poets intensely played with sound through rhyme, assonance, and meter, complicated by performative elements, has only been a focus of the scholarship in the past decade. Within these discussions are three main debates: 1) how are we to extract formal elements from a noisy manuscript transmission?, 2) how did subgenres within lyric (mainly *Minnesang*, *Sangspruchdichtung*, *Spruchdichtung*, and *Leich*) differ in respect to their reception?, and 3) how did this formal play change throughout the MHG period? While these debates continue today, their implications for the analysis in this chapter will prove valuable.

Transmission

The modern line in the scholarship concerning the reception of medieval German lyric is that it was a mixed medium, both song and spoken performances by the poets themselves as well as messengers to a private audience, but that there was also an evident readership, small though it may be. This complicated transmission environment provided for an even more difficult manuscript tradition. Many different versions of songs circulated, and the

[33]"Der Mann, der die erste Strophe spricht, demonstriert seine Könnerschaft indem er nur einen emzigen Reim verwendet. Die Frau tut in der zweiten Strophe das Gegenteil und reimt überhaupt nicht. Damit erscheint ihre Rede als kunstlos und der männlichen monorimen Rede unterlegen." Braun, "Aufmerksamkeitsverschiebung. Zum Minnesang des 13. Jahrhunderts als Form- und Klangkunst," 223

[34]Cramer, *Waz hilfet âne sinne kunst?: Lyrik im 13. Jahrhundert Studien zu ihrer Ästhetik*, 16.

mediation between performance and scribe is uncertain.[35] To complicate matters further, Karl Lachmann set forth an editing tradition in the 19th century seeking to reveal the *Urtext* (original text), providing for the unfortunate case that few modern editions of MHG texts follow any manuscript directly. Given this situation, how are scholars to attempt a formal analysis, not to mention a computational formal analysis?

Recent investigations favor looking for 'traces' of the work as a whole in these texts, including traces of the performance. Markus Stock raises the question not *whether* we can somehow reconstruct a performance, but rather "which traces of a 'conceptual' performance can be found in the song."[36] Stock reasons that while our goal should not be to rediscover the exact performance of *Minnesang*, it should rather be to look for traces of this performance within the noise that is the manuscript tradition. Stock argues that these artfully crafted songs must have been written down prior to a performance, and that some aspect of this preconceived act had to have been transmitted through the manuscripts:

> With the artful songs of the late *Minnesang* in its formal play it is unimaginable that they were not written down before the performance. Traces of this prior act of writing must be recognizable in the text of the song transmitted in the manuscript. I am thus not concerned with traces of a 'real' performance, rather with traces of the conception of how a song was supposed to sound. An opportunity to handle this methodological problem presents itself here, that the vocal genre of *Minnesang* only survives in written form: the surviving text of the manuscript must preserve traces of the conception of the song designed for an oral performance, likely written down before the performance, as traces of the performance itself.[37]

Like Stock, Florian Kragl also argues that despite a smaller corpus, these texts can still leave behind "Spuren" of what could have been, and it is these, for which we must search.[38] He calls for a new "normative descriptive language for the formal phenomena, which can above all accomodate the uncertainties."[39] In Hugo Kuhn's search for a phenomenology of medieval German texts in his book *Text und Theorie*, he suggests this problem of sifting out the original existence of an art from its heavily mediated transimission as one of the greatest

[35]Particularly regarding the MHG word 'dihten' and whether it implied a dictation format.

[36]Stock, "Das volle Wort - Sprachklang im späteren Minnesang," 194.

[37]"Bei den kunstvollen Liedern des späteren Minnesangs in seiner formalistischen Spielart ist es unvorstellbar, daß sie vor dem Vortrag nicht niedergeschrieben wurden. Spuren dieser dem Vortrag vorgängigen Niederschrift müssen sich im Text des handschriftlich überlieferten Liedes wiederfinden. Es geht mir also nicht um Spuren einer 'realen' Aufführungssituation, sondern um Spuren einer Vorstellung davon, wie das Lied klingen sollte. Hier deutet sich eine Möglichkeit an, mit dem methodischen Problem umzugehen, daß die Vokalgattung *Minnesang* nur schriftlich überliefert ist: Der überlieferte Text der Handschrift mußte eher Spuren einer dem Vortrag vorgängigen, wahrscheinlich schriftlich niedergelegten Konzeption des für den mündlichen Vortrag bestimmten Liedes bewahren als Spuren eines Vortrags selbst." ibid., 194-195.

[38]Kragl, "wort unde wîse. Formen des sangbaren Verses in der deutschen Literatur des Mittelalters," 46-47.

[39]Ibid., 42.

problems posed to MHG scholars.[40] Recalling Benjamin's aura of an artwork, Kuhn argues that the originality of an artwork, even if copied, can never actually be reproduced.[41] Yet Kuhn is not so pessamistic as to believe that there can be no recovery of any relationship to the original:

> But also its existence, its reality can find the individual work only in the real ingenuity of the same universal forms (in contrast to the subjective finality of the 'original'), that is, however; its given form of existence is the—transmission! Because the existence of the work is, like every individual thing, forcibly enclosed within the reality of the individual forms.[42]

This leads Kuhn to question the real usage ("Gebrauch") of this artwork, and the person making use of it, and the totality, or 'universality' of an artwork. Kuhn argues that in the medieval German context, even if everything 'real' is incorporated into an artwork, and it is the 'original' in all senses of the word, it still does not encompass a "complete whole".[43] It was never actually *more* complete or *more* original at any point in its history than in the transmission, with which we are left; all forms are still open. Thus Kuhn, like Stock and Kragl, while softening slightly on the use of manuscript transmission, believes that we can use these incomplete transmissions of medieval texts (because in his view, *all* possible transmissions since inception were incomplete) not to reconstruct the original, or ever claim that we are approaching an original, but simply to make our representation "deutlicher" (clearer).[44] The aim of this project is to aggregate these "traces" on a large scale, in hopes to make our idea of MHG soundscapes "deutlicher". As demonstrated in the previous chapter, discrepancies between the texts in my corpus and manuscripts, as measured in the following methodologies, would not hinder such an analysis.

Subgenres

Within the MHG lyric there are certainly thematic differences between poems about love (*Minnesang*) and poems about political and religious subjects. Yet most manuscripts, including the large collections A, B, and C, do not distinguish subgenres within lyric, and place differing themes side by side, likely only distinguishing by author, or possibly date.[45]

[40] Kuhn, *Text und Theorie*, 29.

[41] Walter Benjamin, *Das Kunstwerk im Zeitalter seiner technischen Reproduzierbarkeit: drei Studien zur Kunstsoziologie*, 2nd ed., Edition Suhrkamp 28 (Frankfurt am Main: Surhkamp, 1968); Kuhn, *Text und Theorie*, 30

[42] "Aber auch seine Existenz, seine Gegebenheit kann das einzelne Werk nur in der realen Geschichtlichkeit derselben universalen Formen (im Gegensatz zur subjektiven Endgültigkeit des 'Originals') finden, d.h. aber; seine gegebene Existenzform ist die - Überlieferung! Denn die Existenz des Werks ist, wie die jedes Einzeldings, zwangsläufig in die Realität der Einzelformen eingeschlossen." ibid., 33.

[43] Ibid., 35.

[44] Ibid., 37.

[45] Brunner and Tervooren, "Einleitung: Zur Situation der Sangspruch- und Meistergesangsforschung," 5.

Because it is generally agreed that *Minnesang* was set to music, it is the latter category of political and religious poems, in which it has been proposed to further split the genre into those set to music and those intended to simply be read aloud, likely to a particular rhythm but without accompaniment (referred to as either *Sangspruchdichtung* or *Spruchdichtung*).[46] Nevertheless, there are two axis, upon which any given poem for the MHG lyric corpus can be placed to any degree — 1) the degree of thematic occupation with *minne*, and 2) the degree of musicality or rhythm assumed to have accompanied a poem.

In a collection of essays by Helmut Tervooren, *Schoeniu wort mit süezeme sange*, he focuses on these genre differences within the MHG lyric itself.[47] In the early 19th century, Karl Simrock was the first to note possible subgenres within MHG lyric.[48] The early years of the debates in the 19th century primarily focussed on differences in form, in that *Spruchdichtung* was generally one longer stanza, while songs, including *Minnesang*, were several stanzas. In the late 19th century, these theses were refuted, and it was questioned by some whether a difference within the MHG lyric corpus even existed.[49] Beyond formal differences, there are obvious content differences between *Minnelieder* and political songs.[50] In addition to the two axis just mentioned, Hermann Schneider suggests an additional dimension to distinguish poems—social environment, i.e., the intended audience.[51] The debate, to a degree, rages on today as to what exactly the difference between *Lied* and *Spruch* is, and whether there are further divisions within each. The scholarship has also identified *Spruchtöne* through a "unity" of songs and tones within a corpus.[52] It has also begun to take into consideration song composition outside the Germanic language realm, specifically the Provençal tradition and Latin love poetry. Tervooren finishes his literature review with a challenge: "desirable would be further investigations that clear up what I believe to be an ill-suited conception that there is a metrical-formal difference between *Lied* and *Spruch* (anacrusis, filling of feet, coincidence, etc."[53]

In an effort to quantify these subgenres and provide a response to Tervooren (though not a response Tervooren would find attractive), Johannes Rettelbach compares 60 songs

[46]Helmut Tervooren, ""Spruch" und "Lied"," in *Schoeniu wort mit süezeme sange: philologische Schriften*, ed. Susanne Fritsch-Staar and Johannes Spicker, Philologische Studien und Quellen 159 (Berlin: E. Schmidt, 2000), 60.

[47]Ibid.

[48]Karl Joseph Simrock and Wilhelm Wackernagel, *Gedichte Walthers von der Vogelweide* (Berlin: Vereinsbuchhandlung, 1833), 175.

[49]Johann Rathay, "Über den Unterschied zwischen Lied und Spruch bei den Lyrikern des 12. und 13. Jahrhunderts" (PhD diss., Leopoldstädter real- und obergymnasiums, 1875), 28-30.

[50]Gustav Roethe, *Die Gedichte Reinmars von Zweter*, Nachdr. der Ausg. Leipzig, 1887 (Amsterdam: Rodopi, 1967).

[51]Hermann Schneider, "Mittelhochdeutsche Spruchdichtung," in *Mittelhochdeutsche Spruchdichtung*, Wege der Forschung 154 (Darmstadt, 1972), 134–145.

[52]Friedrich Maurer, *Die politischen Lieder Walthers von der Vogelweide*, 2nd ed. (Tübingen: Niemeyer, 1964).

[53]"Erwünscht wären weiter Untersuchungen, welche die meines Erachtens nicht zutreffende Vorstellung ausräumten, zwischen Lied und Spruch besteht metrisch-formale Unterschiede (Auftakt, Füllung, Fügung usw.)." Tervooren, ""Spruch" und "Lied"," 72.

of *Minnesang* with 45 songs of *Sangspruchdichtung*, focussing on the total length of a song according to accented syllables and rhyme.[54] His results: "the average length in *Minnesang* is 35.4 feet, the average length in *Spruchdichtung* is 60.3 feet; *Spruchdichtung* is thus 70% longer."[55] Rhyme, however, did not prove to be as significantly different. Furthermore, long songs in *Minnesang* are generally limited to 2-3 stanzas in length, while *Sangspruchdichtung* offers fewer restrictions.[56] Anacrusis is also much more regulated in *Minnesang*.[57] While Rettelbach's analysis is a great start to such an investigation, the dimensions of MHG lyric are too many to restrict an analysis to so few formal features, and a sample size of 105 songs from such a varied corpus should hardly be understood as the rule. Nevertheless, Rettelbach does support many of the observations brought to light in the early years of the subgenre debate.

Recently, Jens Haustein has ventured beyond distinguishing subgenres to identifying poems uncharacteristic in either form or content to their most evidently related counterparts, posing the question: "Just how much rhyme can a clear thought tolerate?"[58] Haustein is interested in what happens when formal features of *Sangspruchdichtung* appear in Minnesang, and vice versa, essetially mismatching form and content. This is similar to Kragl's research, already discussed.[59] He looks specifically at Marner *Ton VI*, Konrad von Würzburg *Spruchton 23*, and Frauenlob's *Goldenen Ton*. While not going into too much detail due to the length of the collected articles, he shows that even without leaving the expected genre, some poets intentionally mixed up form to experiment.

Early-, High-, and Late MHG Lyric

In addition to the subgenre distinctions, there were also many chronological developments in MHG lyric through its relativey short flourishing. Cramer was the first to observe a shift in MHG lyric, arguing that rhetorical play intensified in the beginning of the 13th century, when the poems became more about how the poets crafted them than about the actual content of the poem: "the world, in which the poet was engaged, became formulaic language or the opposite thereof: the linguistic formulation created its own world, language created itself and constituted its own space."[60] Cramer discusses how such poets sung about a world generated

[54]Johannes Rettelbach, "Minnelied und Sangspruch: Formale Differenzen und Interferenzen bei der Tonkonstitution im 13. Jahrhundert," in *Sangspruchdichtung. Gattungskonstitution und Gattungsinterferenzen im europäischen Kontext* (Berlin: De Gruyter, 2007), 154.

[55]Ibid.

[56]Ibid., 155.

[57]Ibid.

[58]Jens Haustein, "Grenzgänger Formexperimente in der Sangspruchdichtung des Marner, Konrads von Würzburg und Frauenlobs," in *Sangspruchdichtung um 1300: Akten der Tagung in Basel vom 7. bis 9. November 2013*, ed. Gert Hübner, Dorothea Klein, and Universität Basel, Spolia Berolinensia 33 (Hildesheim: Weidmann, 2015), 250-251.

[59]Ibid., 251.

[60]Cramer, *Waz hilfet âne sinne kunst?: Lyrik im 13. Jahrhundert Studien zu ihrer Ästhetik*, 163.

by language, one which represented the world.[61] He justifies this with an example from late MHG poet Konrad von Kilchberg, in which he uses language to "generate wilfulness" and its "own sound and space for associations."[62] Yet Cramer does not necessarily see this as a positive development, but he rather calls this discovery of the "creation" ability of language by the 13th century poets a "crisis".[63] More recent scholarship from Marcus Stock and Manuel Braun disagree with Cramer's and Kuhn's assessment that the shift from a focus on content to a focus on form in the 13th century led to fewer and less important contributions to the genre and witnessed less innovation. Stock argues specifically against Kuhn's interpretation of Gottfried von Neifen's song as not contributing anything new, arguing that "an excess of artful language, as it appears, stands opposite a minimum of content."[64] Braun argues against those who believe early *Minnesang* in the 12th century was not as innovative or form-centric as the 13th century, which is heavily represented by Cramer, witnessing a language "crisis". In fact, Cramer believes this led to the demise of the form: "Language can only talk about its own emptying of sense (*Sinnentleerung*) for a limited time. Exactly this is the reason that the exposure of the formulaic lyrical language cannot remain poetically productive or bring forth a new song (*niuwen sanc*)."*niuwen sanc* hervorbringt."[65]

Call for Further Work

The current scholars researching sound, Kragl, Braun, and Stock, have all in one way or another called for a new approach to formal analysis. Kragl calls for a new formal analysis that addresses the crucial aspects of meter, music, rhyme, and sound, but do not simply revive debates from the 19th and 20th centuries that do not incorporate current research on mediality, performance, and manuscript traditions. Kragl's method is to distinguish (but not analyze separately) the medieval form from its content: "With 'forms of sung verse in the German literature of the Middle Ages' the secondly named 'form-concept' — as a distinct concept opposite 'content' — is missing."[66] Traditionally, for medieval Germanists, form has meant verse structure, rhyme, and perhaps melody, if available. Kragl is searching for a tool to speak about, and differentiate between, different manuscripts that permit entirely different scansions, and thus likely entirely different musical compositions, yet are stil the same song.[67] Kragl asks an important question: Can content coerce the poem into a specific form (i.e., can "Heldenepikstoffe" (the material (semantic content) of a heroic epic) coerce the poet to compose in "Heldenepikstrophen" (heroic epic stanzas)?[68]

[61]Cramer, *Waz hilfet âne sinne kunst?: Lyrik im 13. Jahrhundert Studien zu ihrer Ästhetik*, 165.
[62]Ibid., 174-175.
[63]Ibid., 177.
[64]Stock, "Das volle Wort - Sprachklang im späteren Minnesang," 188.
[65]Cramer, *Waz hilfet âne sinne kunst?: Lyrik im 13. Jahrhundert Studien zu ihrer Ästhetik*, 188.
[66]Kragl, "wort unde wîse. Formen des sangbaren Verses in der deutschen Literatur des Mittelalters," 33.
[67]Ibid., 72.
[68]Ibid., 41.

Braun has noticed a shift of attention, or "Aufmerksamkeitsverschiebung", in many disciplines that arrived somewhat delayed in medieval German. Specifically, despite the focus on manuscript *mouvance*, performance, and orality, while still integral to the literature, scholarship must move past this barrier. Returning to the form has become a new possibility. He correctly identifies this problem as stemming from the exhaustive formal scholarship conducted by the 20th century philologists, but as others, emphasizes that this scholarship needs to be remolded with our modern approaches: "Nobody would deny that *Minnesang* was above all a formal art, a formal art to the highest degree."[69] Braun also demands that the scholarship develop new ways to discuss form, instead of simply identifying a period of innovation.[70] Differing from Kragl, Braun agrees with Jakobson in that form and content should not be, and cannot be, separated, as words are constituents of form, and vice versa.[71] He is interested in where the balances are found between the two, or one is prioritized over the other.

Stock is surprised that given the recent work on the performance and presentation of MHG songs, there has been relatively little work on sound.[72] If we believe these poems were primarily read aloud, despite the extent to which they may have been read, any insight into the sound used in them would be a crucial aspect to the performance. In his article 'Das volle Wort - Sprachklang im späteren Minnesang', Stock focuses on the sounds of words, and how repetition of words is not only a repetition of a concept, but the repetition of the word's sound. The repetition of sounds, be they words or not, reinforces concepts and their connection to that sound.[73] It is not about hearing and recognizing these patterns, rather that "in the sound dimension the individual rhetorical figure is blurred and a complete acoustic impression emerges."[74]

The syllabification of MHG opens up a world of possibilities to investigate these formal aspects. This chapter explores the possiblities of a large-scale computational formal analysis for characterizing MHG soundscapes, while the next chapter takes this one step further into an analysis of prosodic patterning and the tension between form and content.

3.3 Distant Reading Soundscapes

With the above two *Frauenstrophe* (female stanzas) examples in a closed, heavy syllable environment, one may wonder if these soundscapes carry specific content—i.e, do these formal features correlate with themes or voices? To determine this relationship, I take all

[69]Braun, "Aufmerksamkeitsverschiebung. Zum Minnesang des 13. Jahrhunderts als Form- und Klangkunst," 204.

[70]Ibid., 206.

[71]Ibid., 210.

[72]Stock, "Das volle Wort - Sprachklang im späteren Minnesang," 185.

[73]Ibid., 195.

[74]Ibid., 200.

	stanzas
count	7856
median syllables per stanza	81
median lines per stanza	9
median syllables per line	8
median percent open syllables	32.88%

Table 3.1: Summary statistics for soundscape stanza analysis

lyric poetry from the MHDBDB corpus[75], neglecting subgenres for the moment. I first split each text into stanzas. While debates still continue over which stanzas belong to which poems, I will not attempt to make any sequential arguments based on stanzas in this chapter, but consider each stanza in the corpus a separate unit, regardless of author or song. The MHDBDB texts are divided into songs and stanzas, but also provide majiscule markings, where observed in the edition or manuscript. For the purposes of this analysis, I restrict myself to stanzas > 2 lines and < 50 lines. While some songs have extended stanzas, especially in the *Spruchdichtung* tradition, considering a unit as > 50 lines distracts from the focus of my study, i.e., a smaller unified group creating a clear soundscape. I run all subsequent analysis on the stanza divisions as well as majiscule divisions for a comparison of smaller units. I also leave out authors and songs that have multiple editions of the same poem from the same manuscript, as well as songs in the early 15th century, leaving lyrical poetry in the 12th, 13th, and early 14th centuries.

The soundscape for a stanza is quantified simply by calculating the percentage of open syllables, i.e., syllabes ending in a vowel. Reinmar's first stanza from the beginning of the chapter has 9 open syllables and 50 closed syllables, and is thus assigned the value of 15.25%. Reinmar's second stanza has 42 open syllables and 31 closed syllables and is assigned 57.53%. The method in this chapter therefore only considers the aggregate soundscape affect of a stanza, and does not incorporate sequential information, rhyme, or any other formal features.[76] Relevant summary statistics for this subset of the MHDBDB corpus are reproduced in Table 3.1.

To see the extent to which MHG poets utilized the spectrum of soundscapes available to them, the variance was calculated for each text's soundscape inventory. Text attribution debates aside, it is evident that different poets had varying degrees of soundscape variance within their corpus. A selected list of variances are shown in Table 3.2. The table provides four noteworthy observations: 1) 'Namenlose Lieder' (anonymous songs), presumably written by different poets, logically has a higher variance, 2) Ulrich von Liechtenstein's poetry and his *Frauendienst* have nearly the exact same variance, 3) Konrad von Würzburg has

[75]The texts are all those designated as 'Lyrik' in the MHDBDB, with duplicates removed. A list of the texts for this analysis can be found in Appendix D.

[76]I will attempt to account for these features in a more sophisticated model in the next chapter. Nevertheless, such a simplistic sound-centric approach can provide for more interpretable, salient results.

	variance
Heinrich von Stretelingen	0.094953
Günther von dem Vorste	0.087832
Namenlose Lieder (Minnesangs Frühling)	0.077619
Konrad von Altstetten	0.069000
...	...
Albrecht von Johansdorf (Minnesangs Frühling)	0.058233
...	...
Ulrich von Liechtenstein	0.045749
Frauendienst Lieder (Bechstein)	0.045000
...	...
Dietmar von Eist (Minnesangs Frühling)	0.038921
...	...
Konrad von Würzburg Lyrik	0.011005
...	...

Table 3.2: Variance in texts' soundscapes

a suspiciously low variance, given the scholarship's agreement on his innovation of form[77], and 4) while the poet with the highest variance in this case is certainly partly due to his few leftover songs (only 11 stanzas), Heinrich von Stretelingen's awareness of soundscapes is clear from his most notable song:

> Nahtegal, guot vogellîn,
> mîner frouwen solt du singen in ir ôre dar,
> Sît si hât daz herze mîn
> und ich âne fröide und âne hôhgemüete var.
> Sî daz niht wunder,
> son weiz ich frömder dinge niht,
> daz man darunder
> hie bisunder
> dike frô mich siht.
> Deilidurei faledirannurei
> lîdundei faladaritturei![78]

[77]Konrad von Würzburg will be revisited in a more thorough analysis later in this chapter and in the next two chapters.

[78]Karl Bartsch, *Die Schweizer Minnesänger* (Frauenfeld: Huber, 1964), 106-107. "Nightingale, good little bird, you should sing into the ears of my lady, since she has my hear and I was without joy and without optimism. If it's no wonder that I don't know anything wonderful, that one often sees me particularly happy. Deilidurei faledirannurei lîdundei faladaritturei!"

Frowe, bluomen unde klê
unde heide, diu so wunneklîche grüene lît,
Die wen muoten unde mê,
daz diu vogellîn wol singen suozze widerstrît.
Des fröit sich sêre
mîn gemüete, dâz si sint fröiderîch.
al dur ir êre
singe ich mêre,
sît si ist minneklich.
Deilidurei faledirannurei
lîdundei faladaritturei![79]

Süezze minne, hilf enzît,
daz diu sælderîche erkenne mîne nôt!
Sît daz mîn trôst an dir lît,
so füege, daz ir süezzer munt durliuhtig rôt
Der senden quâle
in kurzen zîten werde gewar.
schiuz dîn strâle
zeinem mâle,
du weist wol selbe war.
Deilidurei faledirannurei
lîdundei faladaritturei![80]

Not only does Heinrich experiment with the real soundscape of singing birds, but he also emphasizes his own singing through a marked change in soundscape introduced by the nightingale itself.[81] In the first stanza, Heinrich instructs the nightingale to begin its song by

[79]Bartsch, *Die Schweizer Minnesänger*, 106-107. "Lady, flowers, and clovers, and heather, which lay so wonderfully green, and who want very much that the little bird sings well and sweet in response. They are pleased by this very much, my senses, that they are joyful. All through her honor, I sing more, since she is dear. Deilidurei faledirannurei lîdundei faladaritturei!'

[80]ibid. "Sweet lady love, help me timely, so that the exhilerating lady notices my suffering! Because any help for me is in your power, so do, so that her sweet, brilliant red mouth recognizes quickly the agonizing longing. Shoot your arrows quickly, you know where to. Deilidurei faledirannurei lîdundei faladaritturei!"

[81]The image of the nightingale is well-known in *Minnesang* as often representing the singer himself, not only in song, but the world's appreciation for it. This symbolism especially took hold through the songs of Walther von der Vogelweide (See Hugo Kuhn and Christoph Cormeau, *Minnelieder Walthers von der Vogelweide: ein Kommentar*, Untersuchungen zur deutschen Literaturgeschichte 33 (Tübingen: Niemeyer, 1982)). While the connection between singer and nightingale is most strong in the texts, Ingrid Bennewitz shows that the nightingale can just as easily transform into a chicken, farmer, and a number of other birds. Ingrid Bennewitz, "Von Nachtigallen, Krähen, Hühnern, und Sägern: Überlegungen zu Aufführung und Sängerrollen im Minnesang, speziell bei Neidhart," in *Edition und Interpretation: neue Forschungsparadigmen zur mittelhochdeutschen Lyrik: Festschrift für Helmut Tervooren*, ed. Helmut Tervooren and Johannes Spicker (Stuttgart: Hirzel, 2000), 73–85

singing into his lady's ear because she has his heart and he lives on without happiness. The nightingale begins to sing in its punctuated, yet open syllable manner at the end of the first stanza, and the second stanza begins this new open syllable environment describing spring. A longer, more metered verse arises in *frowe, bluomen, unde klê, unde heide*. It builds upon the end syllable rhyme of the first stanza, continuing the bird's song and reinforcing the interconnectedness of the natural landscape and the soundscape created by the nightingale. It is no wonder that the bird's *singing* opens up a much more sonorous sequence of syllables: *Dei-li-du-rei fa-le-di-ran-nu-rei lî-dun-dei fa-la-da-rit-tu-rei!*, with only three closed syllables in the 19 syllable sequence. Heinrich clearly recognizes the stereotypical requisites for musical language, nearly imitating a romance language, as imagined by Rousseau.[82] The natural landscape in Heinrich's second stanza, in all its green splendor, *desires* the song of the nightingale, much like his lady *desires*, or in his mind *requires*, his song.[83] To show this similarity, Heinrich fits the soundscape of this stanza into the soundscape created by the nightingale. Kragl writes about how the nightingale can stand in to create an onomatopoeic environment through vocabulary, but we see here that Heinrich utlizies the nightingale to not only increase the musicality and vocabulary of his song, but also to lend legitimacy to himself as as singer and poet as he reimagines himself in the position of the bird.[84]

percent open	stanzas
$< 20\%$	272
$< 25\%$	803
$< 30\%$	1582
$< 35\%$	2055
$< 40\%$	1592
$< 45\%$	974
$\geq 45\%$	516
total	7794

Table 3.3: Counts of stanzas per soundscape bucket

To look for specific thematic and voice relationships within the soundscapes themselves, I calculate the percent of open syllables in each stanza, and group stanzas into seven groups at 5% intervals roughly following a normal distribution shown in Figure 3.1 based on this percentage. The counts for each group are given in Table 3.3. Each stanza in the corpus is then lemmatized and the lemmata are re-assigned to its corresponding bucket.[85] Finally,

[82]Ingrid Bennewitz points out that Neidhart similarly makes the success of *Minnesang* dependent upon its acoustic perception. Bennewitz, "Von Nachtigallen, Krähen, Hühnern, und Sägern: Überlegungen zu Aufführung und Sängerrollen im Minnesang, speziell bei Neidhart," 78

[83]Günther Schweikle, "Heinrich von Stretelingen," in *Die deutsche Literatur des Mittelalters, Verfasserlexikon*, Zweite, völlig neu bearbeitete Auflage, vol. 3 (Berlin ; New York: W. de Gruyter, 1978).

[84]Kragl, "wort unde wîse. Formen des sangbaren Verses in der deutschen Literatur des Mittelalters," 50.

[85]Lemmata in the MHDBDB corpus are annotated by hand for each text included in this analysis.

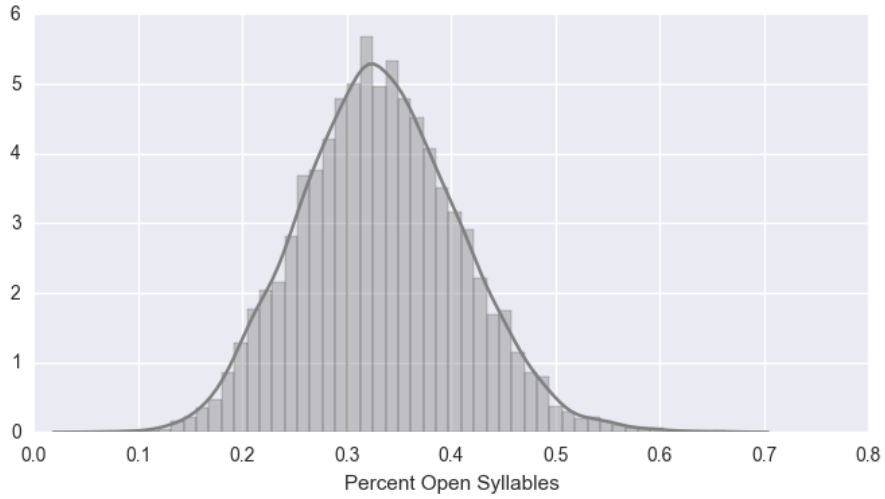

Figure 3.1: Distribution of soundscapes in the MHDBDB lyric corpus

normalized lemmata frequencies are calculated for each bucket and are examined across buckets.[86]

Lemmatization changes the inflected or declined word to its linguistic root. This popular NLP technique collapses words under the same concept. As lemmatizing determines a word's lemma, it also conveniently (and perhaps controversially) normalizes orthography. The first line of the *Nibelungenlied* is shown with its corresponding lemmata below:

uns → *wir*
ist → *sîn*
in → *in*
alten → *alt*
maeren → *mære*
wunders → *wunder*
vil → *vil*
geseit → *sagen*

All word forms are reduced to the same lemma. When counting, this helps the computer understand *geseit* and *sagen* as equal, at least semantically. Therefore, while we are admittedly changing the text, lemmatization is strictly for semantic, interpretive purposes and is *not* used in soundscape calcuations. The formal soundscapes have already been calculated on the stanza as it appears in the edition or manuscript. After each of the seven groups

[86]Lemmata frequencies are normalized by the length of the bucket. Each bucket is bootstrap resampled 1,000 times with replacement to generate 95% confidence intervals.

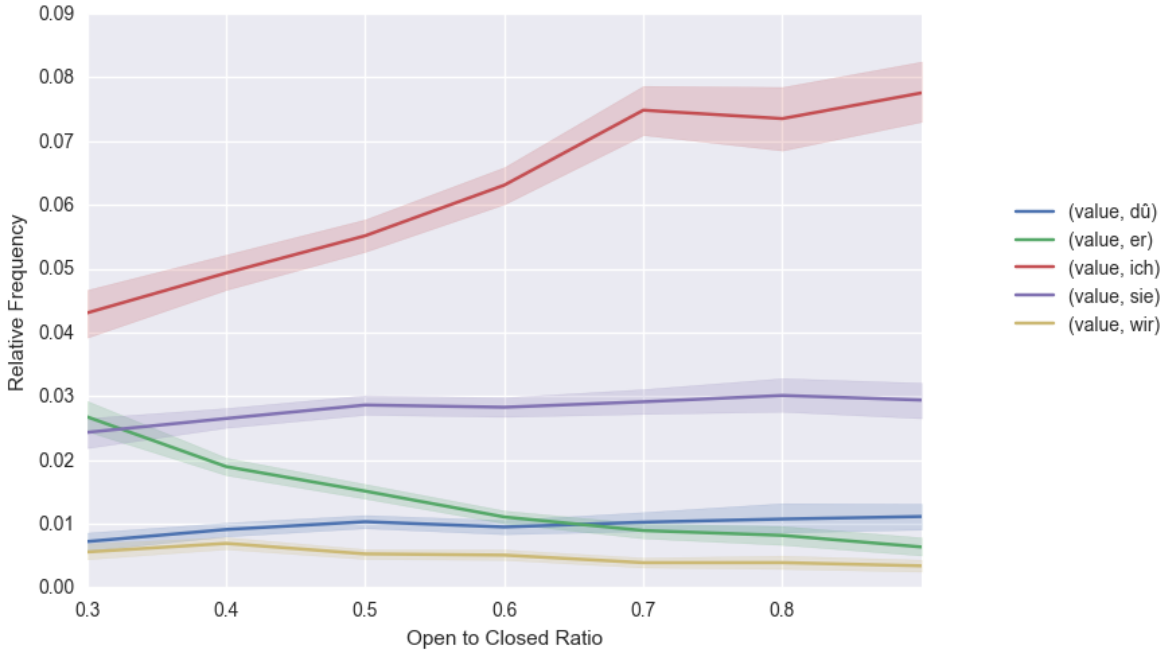

Figure 3.2: Lemmatized pronouns by soundscape grouping, stanza divisions

is converted into lemmata, I take simple normalized word frequencies (a word's share of all words for that text grouping) to look for trends.[87]

Results

The clearest result of this analysis concerns the most common words in the corpus—pronouns and articles. Figure 3.2 shows the relative frequency of the lemmatized pronouns for each soundscape grouping by stanzas[88]:

Among the pronouns, there is clear movement in 'ich' and 'er'. To reiterate, 'ich' in this analysis includes all inflected forms—'ich', 'mich' (me, acc.), 'mir' (me, dat.), etc. Figure 3.2 implies that with a greater share of open syllables in any given stanza, the share of all words referring to a first person subject 'I' increases. The opposite holds true for 'he', while 'she' remains relatively constant across buckets. Most interestingly, 'ich' is a monosyllabic closed syllable, as is all of its forms save the possesive pronoun in the feminine accusative declension. To be growing in its share inside of a soundscape with a greater share of open syllables implies that it must be counteracting itself with more open syllables to offset its

[87]While not performed here, a common task in NLP is to remove 'stopwords', frequent words deemed to have little meaning, and thus difficult to interpret and the cause of further noise.

[88]Unless otherwise noted, majiscule division trends were similar. For all subsequent charts in this section, shaded regions depict 95% confidence intervals determined by bootstrap resampling 1,000 samples.

Figure 3.3: Lemmatized articles by soundscape grouping, stanza divisions

inherent closed quality.[89] Admittedly, 'er' and 'sie' are more complicated cases, as MHG can refer to non-human objects using the masculine and feminine pronouns. Furthermore, 'er' poses a particular problem for lemmatization—'in', the third person masculine accusative pronoun, is also the same word for 'in', the preposition, and 'sie', or often 'sî', can also be the conjunctive form of 'sîn', or 'to be'. Thus from these two pronouns, we may only assume that the third person masculine pronoun in the nominative case likely decreases, and we cannot be certain of the rest.[90] As a sanity check for this analysis, I also chart the trends of each grammatical gender's article, with the hypothesis that due to the endings attributed to each gender, feminine articles would increase as masculine articles decrease, and neuter articles remain somewhat stable.[91] Figure 3.3 confirms this hypothesis.

These results should not be surprising in their general trend, but perhaps are surprising in their clarity. Several tests were conducted to further ensure and explore the accuracy of these results. The most obvious potential complication from the above analysis is that the size, content, and style of these texts vary significantly, and larger texts will inevitably have a larger influence on the analysis. Not only will larger texts have more influence, but especially if that poet tends to write songs in a certain soundscape environment or on a particular subject. The first measure taken to gauge these influences was to remove the

[89]Some of this is due to the first person ending '-e', but not enough to continue growing in an open environment.

[90]A more accurate lemmatization and disambiguation would shed more light on this area.

[91]Following *Mittelhochdeutsche Grammatik* Paul et al., *Mittelhochdeutsche Grammatik*, all masculine definite article inflections are closed syllables, only a nominative nominal adjective ending is open. Feminine neuter and accusative article inflections are open syllables, as is the pronominal nominative adjective ending. All neuter definite article inflections are closed syllables, both the nominative and dative nominal adjective endings are open syllables.

ratio open-closed	stanza count
$< .4$	42
$< .5$	83
$< .6$	119
$< .7$	123
$< .8$	61
$> .8$	66
total	494

Table 3.4: Summary statistics for soundscape stanza analysis, Walther von der Vogelweide

largest texts, and large texts, in which the poet wrote primarily in certain soundscapes (either due to poetic style or other conditions). Neidhart and Frauenlob, for example, have many more stanzas in the closed soundscapes, while Reinmar der Alte and Heinrich von Veldeke on the other hand, have many more soundscapes in an open environment. Walther von der Vogelweide's stanzas are more evenly distributed, but the large share his stanzas take up may also have an unfair influence on the results. Thus the exact same analysis was carried out without these five texts.[92] The results are shown in Figure 3.4 and Figure 3.5. Despite significantly reducing corpus size, and removing the most influential texts, the trends remain. In a similar manner, the same analysis was run leaving out a different text for each iteration, with the trends remaining stable. In Figure 3.6 and Figure 3.7 the shaded regions depict the maximum and minimum from the heldout iterations. Thus it may be concluded that neither a single text, nor an influential text, have significantly influenced the results. As an additional peace of mind, I run the same analysis on Walther's corpus alone, as his corpus is the only one with enough size and breadth in soundscape to yield any kind of confidence intervals, the results are shown in Figures 3.8 and 3.9. While the trends are not as pronounced in Walther's corpus alone, the change from a .6 to a .7 ratio is still very distinct.

Returning to the pronoun analysis, having presented two *Frauenstrophen* with closed soundscapes, it may not be surprising that 'er' and all its inflections distinctly decrease with an increase of the percentage of open syllables. While I have not separately annotated the *Frauenstrophen*, soundscapes with many open syllables rarely identify a third person singular masculine entity, while the rate of feminine references remains relatively stable (as well as the first person plural and second person singular entities). Katharina Boll's thorough study *Also redete ein vrowe schoene* identifies two characteristics marking a *Frauenstrophe*: an

[92]Taking a small sample of 75 texts from the most varied collections, all those with a variance above .02 as shown in Table 3.2, and running the same analysis yields similar results. Thus, when considering only those poets with the most varied collections of soundscapes, the trend appears to hold, though as seen before, the shift between .6 and .7 is the starkest contrast.

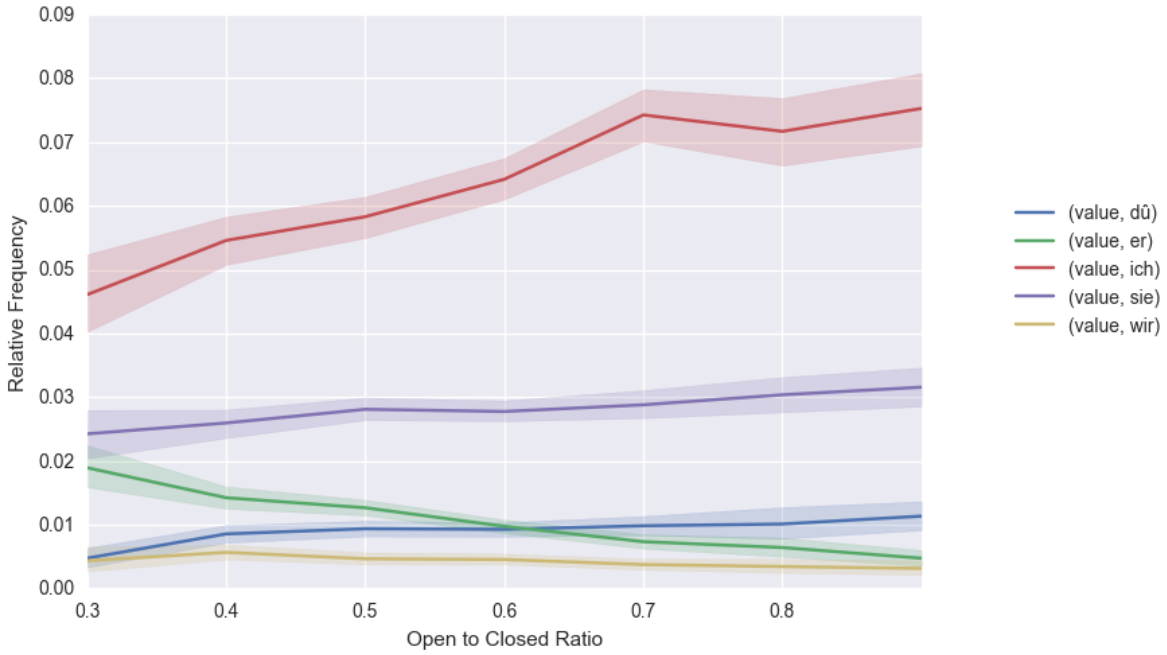

Figure 3.4: Lemmatized pronouns by soundscape grouping, stanza divisions, without NEI, FR3, MRA, MVL, WVV

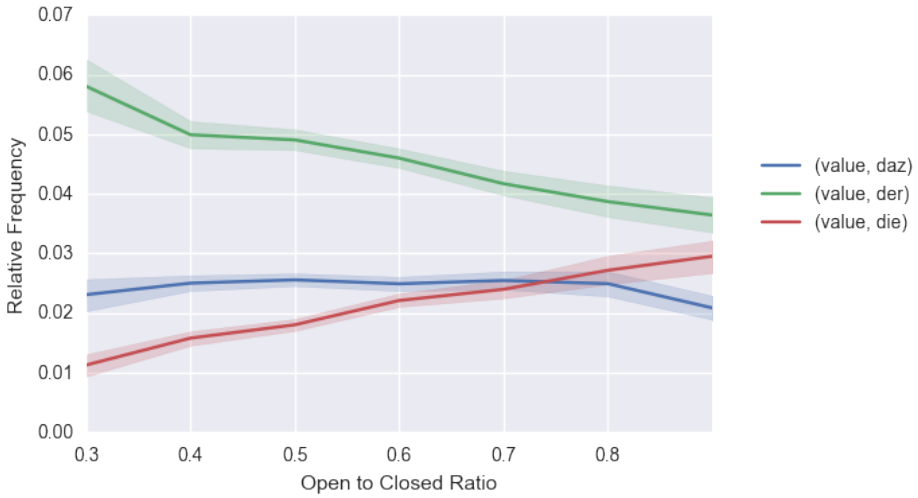

Figure 3.5: Lemmatized articles by soundscape grouping, stanza divisions, without NEI, FR3, MRA, MVL, WVV

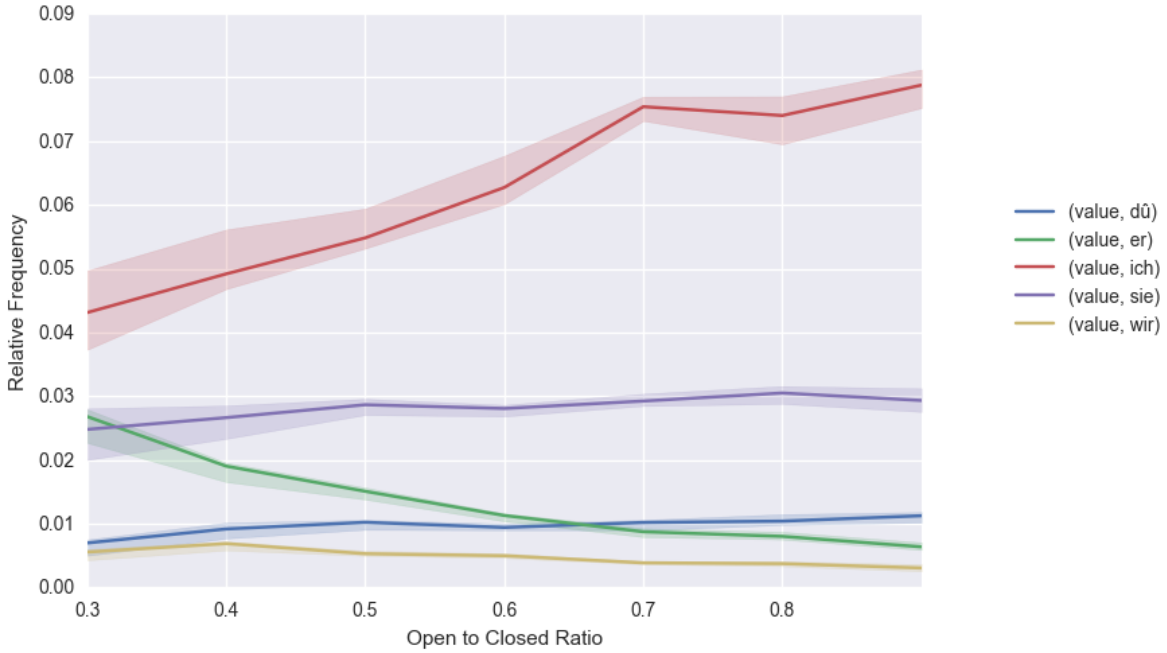

Figure 3.6: Lemmatized pronouns by soundscape grouping, stanza divisions, one held out

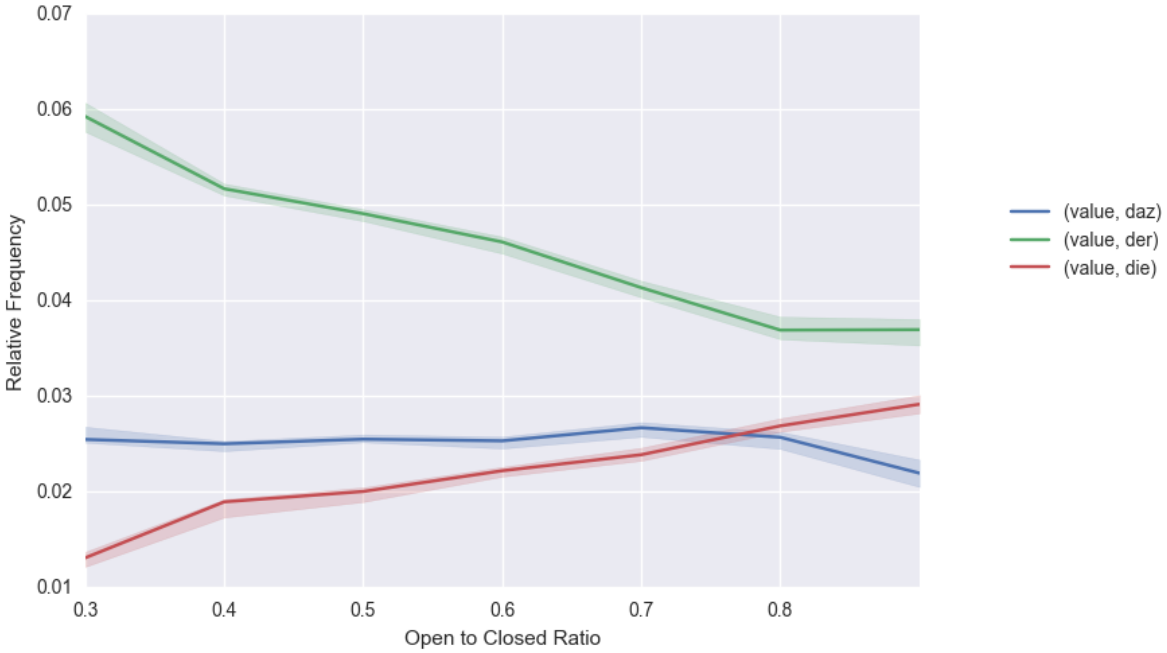

Figure 3.7: Lemmatized articles by soundscape grouping, stanza divisions, one held out

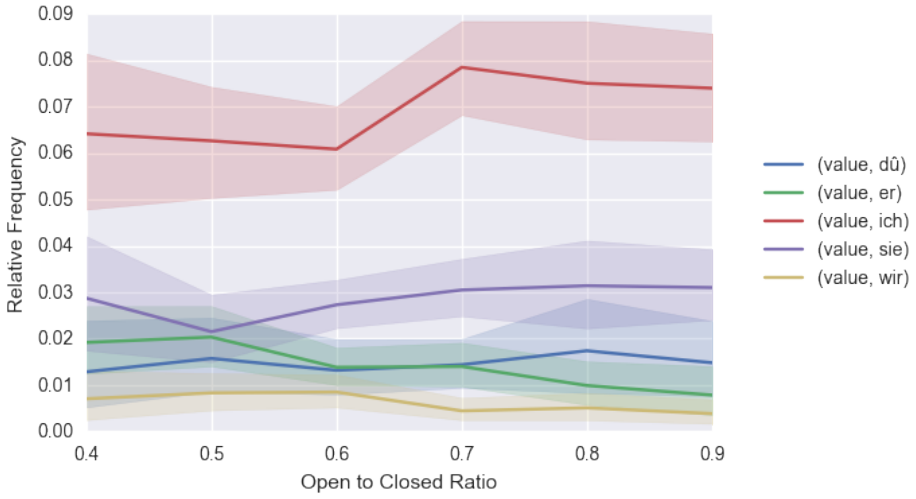

Figure 3.8: Lemmatized pronouns by soundscape grouping, stanza divisions, for Walther von der Vogelweide

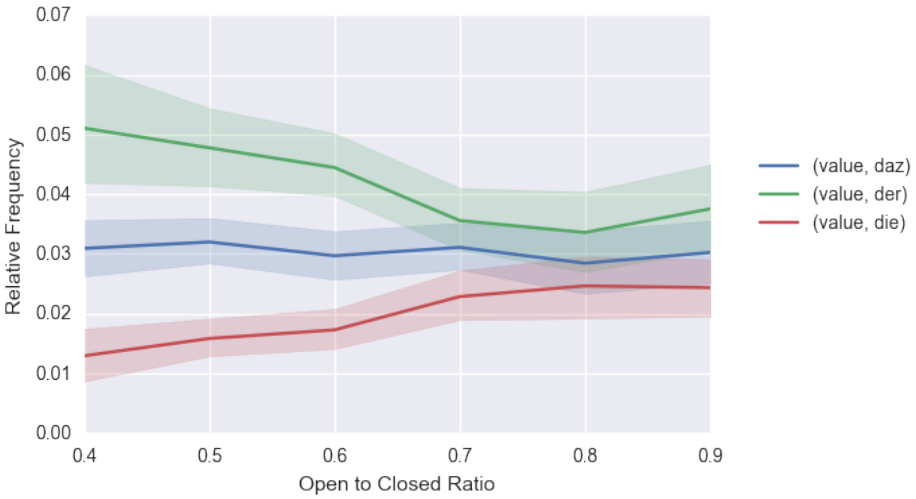

Figure 3.9: Lemmatized articles by soundscape grouping, stanza divisions, for Walther von der Vogelweide

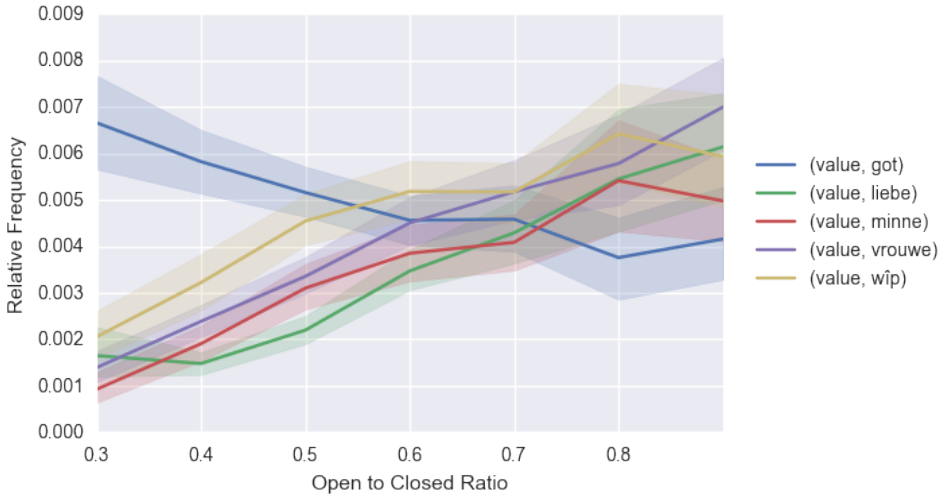

Figure 3.10: Lemmatized *Leitworte* by soundscape grouping, stanza divisions

'Inquit' forumula or apostrophization of the male figure.[93] Boll later writes explicitly that these are internal pronominal references.[94] This 'Inquit' formula, "sprach diu vrouwe", is not captured in the preceding analysis, but the common internal pronoun reference ('er' and its inflections) is captured quite clearly, and is consistent with the two preceding examples from Reinmar and Heinrich. While I hesitate to make any stronger claims in this area (to argue that these pronouns are exclusively associated with humans, as opposed to German's ability to use gendered pronouns for things, is not warranted), this line should be further investigated.[95]

As alluded to above, these soundscape results may also be tied to specific subgenres, which, if true, would contribute significantly to the *Minnesang* and *Sangspruchdichtung* scholarship. While the distinction between *Minnesang* and *Sangspruchdichtung* is still tenuous at best, I chart in Figure 3.10 a few common lemmata for *Minnesang* ('liebe', 'minne', 'vrouwe', 'wîp'), and one counter-lemma for *Sangspruchdichtung*, 'got'. One of the main reasons that selecting a *Leitwort* for *Sangspruchdichtung* is difficult is because the themes are so varied, but religion distinguishes much of this different genre (although as the scholarship suggests, a mixing of religion in *Minnesang* is certainly not rare).

We see a significant trend upward for *liebe*, *minne*, *vrouwe*, and *wîp*, and a significant trend

[93]Katharina Boll, *Also redete ein vrowe schoene: Untersuchungen zu Konstitution und Funktion der Frauenrede im Minnesang des 12. Jahrhunderts*, Würzburger Beiträge zur deutschen Philologie, Bd. 31 (Würzburg: Königshausen & Neumann, 2007), 114.

[94]Ibid., 117.

[95]An interesting extension of this would be to construct a supervised classifier for the gender of MHG voice. This would entail annotating stanzas of MHG lyric, which has been done for the most part by the scholarship. A model could then be constructed to predict the gender for any given stanza. This model would yield the most important features.

downward for *got*. While this cannot be considered the final word, this analysis suggests that both form and content (evidently highly interrelated), or more precisely the soundscape, may contribute to this subgenre distinction. A more sophisticated attempt at a formal distinction will be made in the next chapter.[96]

Soundscapes and the Lyrical 'I'

The forms of 'ich' are unambiguous and unmistakable, and its presence in distinct soundscapes is clear. What consequences does this have for the configuration of the first person in MHG lyric? Unfortunately, as discussed in the first chapter, this is an extremely complicated subject. For the purposes of this project, I adopt a stance between Haferland and Dirk-Müller, recognizing that an increased frequency of the first person pronoun simply implies a stronger sound-based presence of a first person speaker (be it the poet or an assumed role). Bringing Stock's research into the discussion, we recognize that repeating the first person pronoun is not only repeating a semantic concept, but the actual sounds of 'ich', 'mir', 'mîn', etc. It will help to illustrate this trend with an example.

Gottfried von Neifen belongs to the group of *Minnesänger* involved in shifting the genre toward more formalist ambitions. In Stock's article 'Das volle Wort', he investigates Gottfried's use of sound through word repetition, specifically through the repetition of 'fröide' and 'liebe' in the sense of Vickie Ziegler's '*Leitwort*'.[97] While the *Leitwort* is generally understood to have semantic importance due to its high frequency relative to the rest of the corpus or song, what if the omnipresent lyrical 'I' in *Minnesang*, this heavily controversial figure, is elevated to *Leitwort* via its unique soundscape presence in light of the preceding analysis?

> **Ich** solt aber dur die süezen
> grüezen meien walt heid ouwe
> und der kleinen vogel süezez singen,
> ieze eht **mir** an ir gelingen
> trût **mîn** trôst, **mîs** herzen frouwe,
> daz si **mînen** kumber wolde büezen:
> seht, sô wurde **ich** noch an fröiden rîche.
> truter lip, nu tuot genædecliche:

[96] Although constructing a classifier for MHG lyrical poetry is beyond the scope of this project, the preliminary results presented here suggest that this would be a promising endeavor. A supervised classifier would require annotating the genre of MHG lyric and extracting useful features from the model. The annotation process would need to be documented in detail due to the intricacies of the genre debates.

[97] Vickie L. Ziegler, *The leitword in Minnesang: stylistic analysis and textual criticism*, The Penn State series in German literature (University Park: Pennsylvania State University Press, 1975).

rôter munt, du maht **mîn** leit verdringen.[98]

Süeziu minne, sît dîn minne
sinne krenket zallen stunden,
wie sold **ich** dan iemer frô belîben?
lieber lîp vor allen lîben,
heilet **mir mîns** herzen wunden,
daz **mîn** fröide lige an dem gewinne.
tuot ir daz, sô wirde **ich** froidebære.
sælic wîp, nu scheidet **mich** von swære.
Minne, dû maht **mir mîn** leit vertrîben.[99]

Ich hân fröide von ir eine:
seine troestet **mich** ir güete.
dâ von muoz **mir** spilndiu fröide swinden.
lieze sie **mich** gnâde vinden,
seht, so fröite **mîn** gemüete,
daz **mir** wurde ir rôter kus, ein kleine;
so wær **ich** vil manger sorgen âne.
triutelehter lîp, **ich** lebe in wâne
daz **ich** fröide von iu müge enpfinden.[100]

Ich ſolt aber dvr die ſv͒ſſen
gru͒ſſen meigē walt heide[101] o͒we.
vn̄ der kleinen vogel ſv͒ſſes ſingen.
lieſſe eht mir an ir gelingē.
trut min troſt miſ herzen frowe.

[98] Text from Kraus Kraus, *Gottfried von Neifen.* 35.71% open syllables. "I should but sing sweetly through the sweet, welcoming May forest, heather, water, and to the small birds. May it only reach her as comfort, my dear, lady of my heart, that she would wish to relieve my suffering: see, so will I be rich in joy. My dear, now do so graciously: red mouth, you drive away my suffering."

[99] 39.29% open syllables. "Sweet love, since your love weakens the senses at all hours, how should I then ever be left happy? My dear above all others, heal me of my heart's wounds, that my happiness is won. If this is done, so will I be joyful. Fortunate lady, part me now from grief. Love, you drive away my suffering."

[100] 49.41% open syllables. "I had joy from her alone: her goodness comforted me, from which the playful joy must vanish. If she allowed me to find grace, you see, so would my disposition be joyed, that I would receive her red kiss, a small one; so would I be without many worries. My dear, I live in hope that I may feel joy from you."

[101] Elided in edited edition.

dc ſi minen kvmber wolde bv̊ſſen.
ſeht ſo wrde ich fro̊ideriche.
truter lip nv tv̊ genedekliche.
roter mvnt dv maht min leit verdringen.[102]

Sv̊ſſe minne ſit din minne
sinne krenket zallen ſtvnden.
wie ſolde ich iemer danne fro beliben.
lieber lip vor allē liben.
heilent mir mins herzen wnden.
dc min fro̊ide lige an dem gewinne.
tv̊t ir dc ſo wirde ich fro̊debere.
ſelig wib nv ſcheident mich von ſwere.
mīne dv maht mir min leit vertriben.[103]

Ich han fro̊ide von ir eine.
ſi eine tro̊ſtet mich ir gv̊te.
da von mv̊s mir ſpilndv̇ fro̊ide ſwinden.
lieſſe ſi mich genade vinden.
ſeht ſo fro̊ite min gemv̊te.
dc mir wrde ir roter kvs ein kleine.
ſo wer ich vil manger ſorgen ane.
trútelehter lip ich lebe in wane.
das ich fro̊ide von iv mvge enpfinden.[104]

Formal play is clearly on display in Gottfried's song through rhyme, alliteration, and soundscape. There is a ABCCBADD rhyme scheme with an additional rhyme from the first to the second line. Each stanza, in content, manages to convey a similar message: the singer can derive joy from her alone, and she and love have the power to end his suffering and impart that joy. Yet the urgency and depression appears to increase with each stanza, as the singer is reduced to the single hope that she will relieve his suffing. Despite the fact that the actual word 'fröide' increases in frequency, the stanzas actually become less joyful. In the soundscapes, the percentage of open syllables gradually increases in each stanza, taking on an entirely new soundscape in the last stanza, while retaining nearly the same number of syllables per line and presumably the same metrical scansion. While we clearly

[102]Statistics: 34 open syllables, 49 closed syllables, 40.96 % open.
[103]Statistics: 36 open syllables, 50 closed syllables, 41.86 % open.
[104]Statistics: 44 open syllables, 43 closed syllables, 50.57 % open. Pfaff and Baden (Germany)., *Die grosse Heidelberger Liederhandschrift, in getreuem textabdruck*, 104

see the increased presence of the 'Leitwort' 'fröide', it is easy to overlook the simultaneous increased presence of the first person (following the trend in Figure 3.2). Not only does this subject increase in count, but also in space and time intervals, both crucial aspects of the performance. In the first two stanzas the subject appears in more concentrated sections, while Gottfried leaves himself (or his fictive role) out of consecutive verses to describe nature, a bird, and love. Not so in the final stanza, in which Gottfried's subject asserts himself in every verse, refocusing the audience's attention on the subject before them, not only through these self-references, but also by intensifying the soundscape. Each self-reference throughout the entire poem is itself a closed syllable for which Gottfried must compensate further as the song progresses. The clear increased presence of the word 'fröide' is certainly one solution. Notably, only few of these appear as the first person verb ending '-e'. Stock hesitantly borrows from Kuhn's terminology, that the 'fröide' 'Leitwort' in Gottfried's complementary KLD song 3 through its repeated employment in formal moves becomes "objective".[105] The objective 'fröide' becomes the central theme of the poem, superseding the singer himself. Yet in his poem above, Gottfried appears to struggle with the objectivisation of 'fröide' in order to maintain the singer as the central reference, in fact it is the subject, to which this increased presence of the sound (and 'reality') of 'fröide' is constantly subjected. Although as we have now come to learn, it is not only the sound of 'fröide', rather the sonorous sound of open syllables, with which the subject is temporally and sequentially surrounded, and which may simply serve as a counter-weight to the increased presence of the ego-centric subject, even if this repetition of 'fröide' does not correspond to an increase in real joy.

[105]Stock, "Das volle Wort - Sprachklang im späteren Minnesang," 188; Hugo Kuhn, *Minnesangs Wende*, vol. 1, Hermaea n.F. (Tübingen: Niemeyer, 1967).

Chapter 4

A MHG Text Morphology

While computer-aided literary analysis has been active for decades, the omnipresence and ever increasing accessibility to powerful computing environments has only recently allowed many large-scale analyses to be carried out on personal computers or cloud servers with programming languages that have become easier to learn.[1] Franco Moretti is the closest to what we may call the pioneer of this field, the "Digital Humanities", despite the fact that his hallmark book *Graphs, maps, trees: abstract models for a literary history* does not actually use technology at all. He does, however, introduce quantitative analyses as a new approach to literary history. In *GMT*, Moretti is primarily concerned with literary form and genre, and how large-scale analysis can enhance our understanding of literary movements. His chapter on "Graphs" illustrates trends in the novel and how they vary in different countries, with audience and market as dependent variables. "Maps" visualizes the plots of novels spatially, and emphasizes this spatial aspect in contrast to the temporal. "Trees" is arguably Moretti's most interesting, but also most criticized chapter, in which he attempts to relate literary form to theories of evolution: "evolutionary trees constitute morphological diagrams, where history is systematically correlated with form."[2] He supports this claim by arguing that literary devices and genre, considered formal elements, shape literary history—not texts.[3] As with a tree's branches, "divergence prepares the ground for convergence, which unleashes further divergence: this seems to be the typical pattern."[4]

The most notable critique of Moretti's work is delivered by Christopher Prendergast's 2005 article in the *New Left Review*.[5] Much of Prendergast's criticism centers on what he

[1]Computers have been used even for medieval German since the 1960s. The true power of today's computing rests in its capacity for *large-scale* analysis. Many earlier studies of medieval texts are limited in scope due to text digitization or computational power, or both. For early German computer-aided analysis, see Rudolf Hirschmann and Winfried Lenders, "Computer-assisted study of early German: The Mannheim symposium of 1973," *Computers and the Humanities* 8, no. 3 (1974): 179–181; Rudolf Hirschmann, "A survey of computer-aided research in early German," *Computers and the Humanities* 8, no. 5 (1974): 279–284

[2]Franco Moretti, *Graphs, maps, trees: abstract models for a literary history* (Verso, 2005), 69.

[3]Ibid., 76.

[4]Ibid., 80.

[5]Christopher Prendergast, "Evolution and Literary History," *New Left Review*, 2nd ser., no. 34 (August

claims is Moretti's fallacy of *petitio principii*. Although not explicitly stated, Prendergast attacks Moretti's scientific approach for not being scientific enough, believing that if Moretti is going to use scientific theory, he should apply it as any scientist would. Specifically, Prendergast argues that Moretti's treatment of Arthur Conan Doyle has two faults. First, there is internal bias in his variables. Moretti uses features of "clue-giving" to come to the conclusion that clues are the decisive factor.[6] Secondly, Moretti's argument rests on consumer preferences, which Moretti could know nothing about—how do we know what readers "like?"[7] Moreover, which readers are we speaking of—contemporary or modern?[8] Prendergast then comes to his main concern: "Most of the difficulties in Moretti's approach to literary history spring, I believe, from placing a very large bet on bringing the laws of nature and the laws of culture far closer than they are normally thought to be."[9] Literature does not have genes, and does not strictly follow the principles of nature, reducing Moretti's theory to what Prendergast claims is merely an analogy.[10] Further criticism, and Moretti's response, have all pointed toward the need of cognitive science to substantiate Moretti's claims, and that better technology and data can only make the picture clearer.[11]

A collection of Moretti's essays written in the early 2000s culminated in his 2013 *Distant Reading*. Moretti believes that to further literary analysis beyond what it has already ventured, and beyond interpreting what already has been, a greater distance from the text must be made.[12] He understands one problem with literary studies to be its dependence on an "extremely small canon", and thus while creating this distance we also need to add quantity through breadth beyond the canon.[13] "Distant reading: where distance, let me repeat it, is a condition of knowledge: it allows you to focus on units that are much smaller or much larger than the text: devices, themes, tropes—or genres and systems."[14] In Moretti's application, he is interested in comparative morphology, but a morphology of *texts*, not *words*——what he calls "Quantitative Formalism."[15] He justifies this formalism in that "form is precisely the repeatable element of literature."[16] Moretti importantly admits that not all literature can be fit into a model, and "truth be told, I would be very disappointed if all of literature turned out to follow the laws of the novel'..."[17] Moretti's 'Style, Inc.' expresses his personal justification for his approach:

2005): 40–62.

[6]Prendergast, "Evolution and Literary History," 49.

[7]Ibid., 51.

[8]Ibid.

[9]Ibid., 56.

[10]Ibid., 57, 59.

[11]Franco Moretti, "The End of the Beginning," *New Left Review*, 2nd ser., no. 41 (October 2006): 71–86; Jonathan Goodwin and John Holbo, eds., *Reading Graphs, Maps & Trees: Responses to Franco Moretti* (Anderson, SC: Parlor Press, 2011).

[12]Moretti, *Distant reading*, 48.

[13]Ibid.

[14]Ibid., 48–49.

[15]Ibid., 65.

[16]Ibid., 86–87.

[17]Ibid., 111.

> For me, formal analysis is the great accomplishment of literary study, and is therefore also what any new approach—quantitative, digital, evolutionary, whatever must prove itself against: prove that it can do formal analysis, better than we already do. Or at least: equally well, in a different key.[18]

If a new approach cannot demonstrably improve our analyes of literature, it must be able to analyze in a *different key*. Methods of the digital humanities can certainly not solve all of our literary problems, and they may not even handle them better than existing methods today. After all, digital humanities analyses can only carry out what we instruct, just much more efficiently.

Later in the same essay, Moretti writes: "Here is a modest example of what quantitative stylistics could do: take those units of language that are so frequent we hardly notice them and show how powerfully they contribute to the construction of meaning."[19] Tools of the digital humanities can aggregate intricacies we only experience subconsciously, and allow us to interpret these through statistical methods. This was undoubtedly the case in the previous chapter, detecting properties of sound on a level that is difficult for a reader to comprehend. Borrowing Moretti's terminology, in this chapter I take syllables as frequent "units of language" and change the unit of analysis from the morphology of *syllables*, as in the previous chapter, to the morphology of *texts* by incorporating patterning features on a large-scale. For this task I will also expand my corpus for analysis to include both epic and lyric poetry.

4.1 Soundscape Patterning within MHG Verse

> Ignore for a moment (as an experiment!) the substance contained in language—the biological linguistic substance, for example in the case of love poetry with all its consequences on the senses and values—, there is still something leftover from the linguistic imprint, especially in literary productions: a free configuration of sounds and accents (which relate back to earlier stages, but also beyond these stages, such as is the case for animals and melodies and rhythmic movements). The sounds, the stress and melody accents, even the conceptual and semantic substance of the language combine freely into sound groups, rhythmic groups, conceptual and semantic sets, which admittedly carry the content, but themselves are free from the intended substance.[20]

[18] Moretti, *Distant reading*, 204.

[19] Ibid., 206–207.

[20] Sieht man einmal (als Experiment!) in der Sprache vom Inhalt ab—dem biologischen Sprach-Inhalt z.B. im Fall der Liebesdichtung samt seinen Weiterungen in alle Bereiche des Sinnes und der Werte—, so bleibt von jeder sprachlichen Prägung, besonders von jeder dichterischen, doch etwas übrig: eine freie Konfiguration von Lauten und Akzenten (die sich auf früher Stufe, aber auch länger hinaus, wie bei den Tieren mit Melodien und mit Bewegungsrhythmen verbinden können). Die Klänge, die Druck- und Melodie-Akzente, ja sogar auch die Vorstellungs- und Bedeutungsinhalte der Sprache schließen sich frei zu Klanggruppen, rhythmischen

While the preceding chapter illuminates the aesthetics of MHG soundscapes in the aggregate, it does not fully account for soundscape *patterning*. Syllable alternation, count, and rhyme, were all neglected in favor of the aggregate sound weight of a soundscape in any given stanza. However, including these aspects may better relate poems to one another on a larger scale and across several dimensions.[21] The immediate thought when it comes to prosodic pattering in MHG verse is meter. MHG meter is complex, and no single theory has been agreed upon in the scholarship. Especially in MHG lyric, many verses permit multiple scansions, and because even the extant neumes do not reveal universal rhythmical features, deciding upon the correct scansion in some cases in not possible.[22] Although it is possible that late 13th century poems were composed without music, the musical aspect of the MHG lyric further complicates any attempt to determine rhythmical patterning.[23] While an attempt at a scansion model for the epic poetry is more feasible for these reasons (and is made in the following chapter), in this chapter I rely solely on what can be agreed upon phonologically and paleographically—syllabification, syllable properties, word boundaries, and rhyme. I will demonstrate that these different combinations, though seemingly sparse as features, allowed for poets to both be individually expressive, while also drawing particular connections between their work and the work of their contemporaries.[24]

Toward a 'Text Morphology'

Manuscript traditions, orthography, and dialects all stand as severe obstacles to any large scale computational comparison between the different manifestations of MHG. When counting things, the computer cannot easily recognize two variant spellings of the same word as being the same word, not to mention that different dialects certainly understood even the same lexical word slightly differently. The few attempts to unite the corpus' formal or metrical features, such as Andreas Heusler's *Deutsche Versgeschichte*, have not gone uncrit-

Gruppen, Vorstellungs- und Bedeutungskreisen zusammen, die zwar den Inhalt tragen, selbst aber frei vom gemeinten Inhalt sind. Kuhn, *Text und Theorie*, 50

[21] As Braun notes, even experimenting with line length alone can have an immense effect. Braun, "Aufmerksamkeitsverschiebung. Zum Minnesang des 13. Jahrhunderts als Form- und Klangkunst," 227

[22] At least the actual rhythm cannot be discerned. For example, MHG lyric does not require four stressed syllables per verse. This allows for a *klingend* (ringing) or *weiblich-voll* (feminine full) cadence to be ambiguous. Herbert Bögl, *Abriss der mittelhochdeutschen Metrik: mit einem Übungsteil* (Hildesheim; New York: G. Olms, 2006), 37

[23] Ibid., 36.

[24] A further logical place to take this string of thought is in differentiating between prose and poetry. Anttila and Heuser investigated this distinction in their 2016 article "Phonological and Metrical Variation across Genres."Arto Anttila and Ryan Heuser, "Phonological and metrical variation across genres," in *Proceedings of the Annual Meetings on Phonology*, vol. 3 (2016) Their approach is similar to what I undertake in this chapter and the next. Antilla and Heuser build a model that marks phonologial features of English and Finnish to determine how often four constraints of poetry and prose are violated in each. Similarly, I build a model of phonological and metrical patterning for MHG. I can then extract the model's confidence in assigning these values, and subsequently determine what language is more typical of this specific type of poetry.

icized.[25] But what do we know about MHG prosody that is universal across dialects and orthography, can be quantified sequentially, and relates to the soundscapes identified in the previous chapter? I suggest the following features for a baseline 'text morphology' of MHG soundscapes:

- **Syllables**: We know the language of MHG was structured in syllables regardless of the dialect or orthography. As underlined in Chapter 2, the use of the SSP and LP for syllabification does not bias a dialect or orthography, but is accurate for the syllabification of most variants of MHG.

- **Syllable properties**: We can be confident in the phonological existence of vowels and consonants. Although vowel sound varies from dialect to dialect, and are recorded differently by scribes, the presence of vowels in MHG is undisputed.[26] To mark whether a syllable is open or closed, therefore, is a task in which we can be confident.

- **Syllables per word**: Most manuscripts delineated word boundaries.[27] This is an important formal feature for MHG because of the language's stress-initial tendency.[28]

- **Rhyme**[29]: MHG poetry knows many different forms of rhyme, which developed over the different periods. Heusler notes that starting in 1190 pure rhyme became standard, and it was employed consistently until about 1300.[30] Differentiating between masculine and feminine rhyme can further distinguish verses. Masculine rhymes end on a stressed syllable, feminine rhymes have a stress on the penultimate syllable.[31] Although rare in epic meter, especially in the love lyric rhyme can extend three or four syllables, which is called *gleitend* (gliding) rhyme.[32] A *rührend* (touching) rhyme is a normal rhyme pair, which additionally share the preceding consonant. Lastly, there

[25]Heusler himself resisted any statistical analysis of MHG verse. Heusler, *Deutsche Versgeschichte: mit Einschluss des altenglischen und altnordischen Stabreimverses*

[26]I follow here primarily the *Lautlehre* (phonetics) in *Mittelhochdeutsche Grammatik* as well as Richard Wiese's *The Phonology of German*. Paul et al., *Mittelhochdeutsche Grammatik* Wiese, "The Prosodic Structure of German" I also draw general on general phonology arguments based on prosody from Alan Prince. Alan Prince, "Quantitative consequences of rhythmic organization," *Chicago Linguistic Society* 26, no. 2 (1990): 355–398

[27]Manuscripts will occasionally separate prefixes from the words to which they belong, e.g. *ver binden* (to connect/tie).

[28]This is likely the reason for separating unstressed prefixes. Recent scholarship has further justified different stress patterns for different types of words. Palmer and Kelly, "Linguistic prosody and musical meter in song"

[29]Andreas Heusler characterizes rhyme as a "wohlgefälliges Lautspiel" (pleasing sound play), with a "gesanglich" (vocal) effect in comparison to other formal features such as alliteration. Heusler, *Deutsche Versgeschichte: mit Einschluss des altenglischen und altnordischen Stabreimverses*, 11

[30]Ibid., 101.

[31]It is also noteworthy that unlike in NHG, in MHG bisyllabic rhyming words with the stress on the penultimate syllable are considered masculine if the penultimate syllable is short (e.g. *leben* and *geben*). Bögl, *Abriss der mittelhochdeutschen Metrik: mit einem Übungsteil*, 19

[32]Ibid.

> are pure and impure rhymes, dependent upon whether the participating vowels and consonants are identical, or just phonetically similar. There are also different rhyme structures. In the case of epic meter, rhyme was always pair end rhyme. The lyric had a wide range of options: *Endreim* (end rhyme), *Binnenreim*[33] (internal rhyme), *Zäsurreim* (cesura rhyme), *Anfangsreim* (beginning rhyme), *Dreireim* (triple rhyme), *Reimhäufung* (groups of rhymes), *Kreuzreim* (cross rhyme), *Schweifreim* (tail rhyme), *Umarmender Reim* (embracing rhyme), and *Kornreim* (ababxc).[34]

Unfortunately, the manuscript evidence does not allow too many additional assumptions to be made. We know that long vowels, and thus heavy syllables, did exist, but cannot be sure how the authors intended them in verse, nor in what exact environments long vowels were realized. The standardization of MHG has attempted to identify these heavy syllables, but editorial markings (usually with a circumflex) are anything but consistent. Vowel length was also dependent on the period and dialect of MHG. While metrical theories exist, as already alluded to, these theories are heavily reliant on orthography and structural consistency, and thus open to interpretation especially in the lyric corpus.

While these features appear few in number, I suggest taking advantage of their sequencing by turning to a relatively simple, yet powerful NLP (Natural Language Processing) technique in calculating the cosine similarity between texts' *tfidf* (Term Frequency Inverse Document Frequency) inventories using n-grams of these prosodic feature sequences. To break this down, I will first explain the general methods I adopt in their traditional NLP use.

I first calculate *tfidf* values. The first two letters in the acronym *tf*, for 'term frequency', mean exactly that. This step tallies the frequency of all the words in any given text. If one were to tally all the words in the MHG *Parzival*, for example, the frequencies would look like the 'freq.' column in Table 4.1. Because the absolute count will vary by text, the data need to be normalized, a process that transforms a term's raw frequency to its proportion of all words in that text. This changes the first column of raw frequency into the second column, the *tf*. This is the exact method used in the previous chapter for each lemmatized soundscape bucket.

While this can be helpful in comparing term usage across texts, as before, it does not automatically reveal the most informative words about a particular text, considering that 'der' and 'daz' appear in all texts with likely a similar relative frequency (very high for MHG). This is where the *idf*, or 'Inverse Document Frequency' is needed. The inverse document frequency is used to weigh down the importance of words appearing frequently across the entire corpus, not just in one document.[35] The *idf* is calculated by taking the natural log of

[33] *Binnenreim* includes *Innenreim* (rhyme in the middle and end), *Mittelreim* (sequential lines with rhyming words on the inside), *Mittenreim* (end rhyme followed by middle rhyme), *Schlagreim* (neighboring words)

[34] Bögl, *Abriss der mittelhochdeutschen Metrik: mit einem Übungsteil*, 20-21.

[35] For more on *idf* and its justification see Karen Sparck Jones, "A statistical interpretation of term specificity and its application in retrieval," *Journal of Documentation* 28, no. 1 (January 1, 1972): 11–21 and Stephen Robertson, "Understanding inverse document frequency: on theoretical arguments for IDF," *Journal of Documentation* 60, no. 5 (October 1, 2004): 503–520

word	freq.	*tf*	*idf*	*tfidf*
der (the)	3432	.026	.006	1.56×10^{-4}
daz (the)	2628	.020	.116	2.32×10^{-3}
er (he/it)	2529	.019	.063	1.20×10^{-3}
ir (she/their/its)	2416	.018	.089	1.16×10^{-3}
und (and)	2213	.017	.047	7.99×10^{-4}
ich (I)	2036	.016	.114	1.82×10^{-3}
si (she/it)	1848	.014	.109	1.53×10^{-3}
in (in/him)	1724	.013	.015	1.95×10^{-4}
den (the)	1530	.012	.032	3.84×10^{-4}
die (the)	1521	.012	.032	3.84×10^{-4}

Table 4.1: Word frequency for the MHG *Parzival*

the total number of texts divided by the number of texts containing a given term:

$$idf(t, d, D) = log\left(\frac{\mid D \mid}{\mid \{d \subset D : t \subset d\} \mid}\right) \tag{4.1}$$

Where t is a specific term, d is a specific document, and D is all of the documents.

If a term such as 'der' appears in every document, the quotient of all the documents divided by the documents in which 'der' appears is 1. The natural log of 1 is 0. Therefore the *idf* value for 'der' in a corpus of MHG texts is likely close to 0. The *idf* will increase for every one fewer documents the term appears in. The *idf* values are calculated in the third column of Table 4.1. 'ich', 'daz', and 'si' have the highest *idf* values (though still relatively low), which should not be surprising in such a large corpus of diverse texts, many of which were not written with any first person perspectives.

To obtain the *tfidf* for a specific word in a specific document, we simply multiply $tf * idf$, which theoretically yields the relative importance of that term to that specific document, considering that term's frequency throughout an entire corpus:

$$tfidf(t, d, D) = f_{t,d} * log\left(\frac{\mid D \mid}{\mid \{d \subset D : t \subset d\} \mid}\right) \tag{4.2}$$

Where f is frequency, t is the specific term, d is the specific document, and D is all of the documents.

This will reduce (though not eliminate) the importance of words like 'der', as the *idf* value will be close to 0, and when multiplied by the *tf*, even if the *tf* is high, will yield a low product. What a *tfidf* value represents is then theoretically the importance of a word to a document considering its frequency in the corpus. The *tfidf* values are in column four of Table 4.1. The pronouns have the highest *idf* values as well as 'daz', likely due to orthographic variation of

the neuter article in the editing.[36] The *tfidf* reveals that the term with the highest *tfidf* in *Parzival* is logically 'Gâwân', which appears countless times in *Parzival*, and only in a few of the Arthurian romances in the rest of the corpus, yielding a high *tf* and *idf*, and subsequently *tfidf* value.

The *tfidf* must then calculated for every term in every document to construct a (sparse) matrix of *tfidf* feature values for all documents. To further mitigate the influence of text length we use the L2 normalization function.[37]. All document feature matrices together represent a vector space filled with documents, each document being its own vector dependent upon the *tfidf* feature values for every word. To calculate the semantic similarity between documents, the angle between these two vectors is calculated. Depending on the *tfidf* values for two documents, if they are similar, the lines will run similar paths, and the angle between the two lines will be small. If they are vastly different, then the angle will be large. The scale runs from -1 to 1. The measurement of this angle is called the *cosine similarity*. As these documents find themselves in a high-dimensional Euclidean space, the formula for calculating this angle difference is derived from the formula for a Euclidean dot product:

$$cosineSimilarity(\mathbf{A}, \mathbf{B}) = \frac{\mathbf{A} \bullet \mathbf{B}}{||\mathbf{A}||||\mathbf{B}||} = \frac{\sum_{i=1}^{n} A_i B_i}{\sqrt{\sum_{i=1}^{n} A_i^2}\sqrt{\sum_{i=1}^{n} B_i^2}} \tag{4.3}$$

This is how traditional text processing has approached measuring the semantic similarity between documents in a corpus. As my intent is to investigate soundscapes and form in MHG texts, I do not wish to use words as terms in my analysis, although it will be implemented later for the sake of comparison in an attempt to disambiguate form and content. For guidance, I turn to biology and new methods for clustering DNA sequences, taking n-gram samples from DNA strands and constructing a *tfidf* matrix.[38] For MHG, the question then becomes how can we *extract* formal groups by *abstracting* enough to not be considering the lexical terms themselves.[39] Instead of constructing a *tfidf* matrix of lexical terms or nucleotides, I code the text into string sequences with the four features described above.

Ein ritter sô gelêret was,

[36]This will not affect the syllable level, as daz = das = dc, etc.

[37]For more on normalization of *tfidf* see Amit Singhal, Chris Buckley, and Mandar Mitra, "Pivoted Document Length Normalization," in *Proceedings of the 19th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval*, SIGIR '96 (New York, NY, USA: ACM, 1996), 21–29

[38]Z. Volkovich et al., "The method of N-grams in large-scale clustering of DNA texts," *Pattern Recognition* 38, no. 11 (November 2005): 1902–1912; Andrija Tomović, Predrag Janičić, and Vlado Kešelj, "N-Gram-based classification and unsupervised hierarchical clustering of genome sequences," *Computer methods and programs in biomedicine* 81, no. 2 (2006): 137–153; S. R. Maetschke et al., "A visual framework for sequence analysis using n-grams and spectral rearrangement," *Bioinformatics* 26, no. 6 (March 15, 2010): 737–744 It is important to note that this is not related to Moretti's implementation of genealogical methods to literature, as critiqued by Prendergast, rather simply a method to fully encompass encoded features.

[39]A comparison of formal vs. lexical methods is carried out in the section below.

daz er an den buochen las, (*Der Arme Heinrich* l. 1-2)[40]

C CC O OOC C X
C C C C OC C 1[41]

Where 'C' is a closed syllable (ends in a consonant), and 'O' an open syllable (ends in a vowel). Word boundaries are included as features to account for the stress-initial tendency of MHG. Numbers at the end of a line mark end-rhyme; the number is how many lines back the rhyme was seen and an 'X' stands for the beginning of a rhyme pair (it was not seen in the past lines). In order to compare across all verses and include line transitions, the above must be joined between all lines in a text, creating one long string of features for each text[42]:

C CC O OOC C X C C C C OC C 1 […]

An entire text is then encoded as one long sequence—the 'morphological text', as Moretti would have it. The question then becomes, what will constitute a 'term' for the *tfidf* matrix method described earlier? To generate 'terms' and account for order and sequencing within the text, I take n-gram samples of this sequence. An n-gram is a sequence of n items in any sequence. The following is an example of a 3-gram:

$\underbrace{\text{A very large}}_{\text{3-gram}}$ dog is running down the street.

["A very large", "very large dog", "large dog is", "dog is running", "is running down", "running down the", "down the street"]

The advantage of taking n-grams is readily apparent—n-gram sequences capture subsequences of any sequence. A DNA sequence of nucleotides in the guiding literature may look something like the following:

[40]There was a knight so learned, that he read in the books, Hartmann von Aue, *Der arme Heinrich*, Bibliothek des Mittelalters 6 (Deutscher Klassiker Verlag, 2004)

[41]For rhyme, I code the number of lines to the most recent end rhyme (pure rhyme) word. If the last syllable begins a rhyme sequence, it is marked with an 'X'. While this rules out many forms of rhyme, it does include the most common form in MHG lyric, and sequences of open and closed syllables will indirectly account for other rhyme to a degree because in order to rhyme, the syllables must have the same quality.

[42]Taking after Anttila and Heuser, to productively compare phonological and metrical features we must consider windows or fragments of syllables or words so as to not allow line breaks too large of an influence. Anttila and Heuser write: "After pre-processing the texts, we divided them into fragments in which each fragment has exactly five words, with no punctuation, in order to guarantee that any phonological or metrical difference between prose and verse that might emerge would have nothing to do with line length, but only with the local phonological and metrical arrangement of words."Anttila and Heuser, "Phonological and metrical variation across genres," 2

$\underbrace{\text{ACA}}_{\text{3-gram}}\text{AGATGCCATTGTCCCCC}$

[“ACA”, “CAA”, “AAG”, “AGA”, …]

A 3-gram of these prosodic features for the first two verses of *Der arme Heinrich* (above) would start:

[“C-C”, “-CC”, “CC-”, “C-O”, “-O-”, “O-O” …]

However, to successfully differentiate between texts we need to take n-grams with an $n > 3$. I select an n of 10, as this allows for sequences of around seven syllables to be compared across texts.[43] But it is important to remember that an n-gram collection will only move one character to the right before taking another 10-gram observation, so the entire text's sequence will be accounted for.

A concrete example will serve to illustrate this method. A match using the above feature codings between Wolfram's *Parzival* and Gottfried's *Tristan* is shown below:

Ist zwîvel herzen nâchgebûr[44] (*Parzival* l. 1)[45]
C-OC-CC-COC-X
Reduced to 10-grams = C-OC-CC-CO, -OC-CC-COC, OC-CC-COC, C-CC-COC-X

von sinen schulden ungemach[46] (*Tristan* l. 769)[47]
C-OC-CC-COC-X
Reduced to 10-grams = C-OC-CC-CO, -OC-CC-COC, OC-CC-COC, C-CC-COC-X

This match implies that the number of words and syllables per word are the same, the syllable quality ordering is the same, and importantly, that the first rhyme is a leading

[43] Although most choices of n will actually return similar relative results, the lower n results in a higher degree of similarity between every text as the sequence inventories are not unique enough. With every increase of n, this similarity inevitably decreases (if $n = 20$, there will be fewer and fewer texts that share sequences of 20 with the above coding guidelines). An n of 10 admittedly provided the most interpretable and scalable results, but as already stated, other values of n do not significantly change the similarity rankings, only the similarity scores (recall what the cosine similarity value is actually measuring). The larger the n also places additional bias on longer texts. While the *tfidf* method will mitigate the influence of any one large text, larger texts still have more opportunities to implement different sequences, and the results will reflect this. An n of 10, however, reduces this bias greatly.

[44] ‘If the heart lives with doubt,’

[45] Eschenbach et al., *Parzival*.

[46] ‘(which had) suffered due to him’

[47] Gottfried, Friedrich Ranke, and Rüdiger Krohn, *Tristan*, 2nd ed., Universal-Bibliothek, 4471-4472 (Stuttgart: Reclam, 1981).

rhyme and the second rhyme refers to the previous verse.[48] Because the sequence matches for 13 features including word boundaries, this will create three additional matches in the *tf* inventories when the 10-grams are taken (while simultaneously increasing the *idf* value of these sequences by a smaller factor). Although these verses also happen to share the same scansion in the Heusler tradition, we cannot assume that every prosodic match would carry the same scansion, as more than phonology plays a role in MHG scansion theory.[49]

There are computational methods (the technicalities of which will not be discussed here) to calculate the Longest Common Substring (LCS), or longest match, of two texts. For my purposes, this would be the LCS of two feature sequences. Taking *Parzival* and *Willehalm* as examples, the LCS for these two texts is 38 features long:

O-C-1O-C-C-C-C-CC-C-XC-C-C-OC-CO-C-1C-

The matching positions in the text are marked between hyphens:

Parzival: 180, 20-23:[50]
der tac gein dem âben—de zôch,
dô kom er an ein wazzer snel,
daz was von sînem duzze hel,
ez — gâben die velse ein ander.[51]

Willehalm: 306, 2-5:[52]
diu stuont uf, mit zuht — si sprach,
e daz sich schiet der vürsten rat:
<swer zuht mit triuwen hinne hat,
der — ruoche hoeren miniu wort.[53]

Because of the high similarity in the formal measure for these texts (demonstrated below) and the n-gram sampling, it is safe to assume that longer matches often occur between the two. Conveniently, this LCS displays a near exact match in prosodic formal features with the medial two lines additionally matching in the Heusler scansion tradition, but containing very different meaning. While not necessarily 'opposite' in meaning, as Kuhn reminds us:

[48] In the computational implementation, it also possible for 'X' to be a rogue syllable without a rhyme in the previous nine lines.

[49] I will address the relationship between this method and the Heusler scansion in the following chapter.

[50] Eschenbach et al., *Parzival*.

[51] "The day drew into the evening, he came to a wild stream, roaring it tossed about, slamming into the rocks."

[52] Wolfram et al., *Willehalm*, Deutscher Klassiker-Verlag im Taschenbuch 39 (Frankfurt am Main: Dt. Klassiker-Verl, 2009).

[53] "She stood up, with marked words she spoke, before the council adjourned: 'Whomever is befitted by civility and loyalty, he shall listen to my word"'

sequence	percent of all sequences
C-C-C-C-C-	0.6201%
-C-C-C-C-C	0.4032%
C-XC-C-C-C	0.3100%
C-1C-C-C-C	0.2960%
C-C-C-OC-C	0.2868%
-C-C-C-OC-	0.2754%
-C-XC-C-C-	0.2726%
-C-C-OC-C-	0.2659%
C-C-C-C-OC	0.2570%
C-C-XC-C-C	0.2439%

Table 4.2: Verse sequence frequency

"Who would not quickly see that the same sound scheme with different words and with a different symbolic incantation can have the opposite meaning?"[54]

After 10-gram subsets of the feature strings were collected from each text, a *tfidf* matrix was constructed to identify the feature strings most charactersitic of certain texts and groups of texts

Results

This method was implemented on 595 verse texts from the MHDBDB corpus.[55] Across MHG verse, the most common sequences are reproduced in Table 4.2. Monosyllabic, closed syllable words dominate the MHG corpus, and rhyme pairs are the most common form of rhyme. The fifth most common sequence implies that mixing in a bisyllabic open and closed syllable word is common as well.

These results become more interesting when we look at the *tfidf* values on the level of an individual text. This network is reproduced below in Figure 4.1:

[54]Wer sieht nicht unmittelbar ein, daß dasselbe Klangschema bei anderem Wortinhalt und mit einer anderen symbolischen Zauberformel auch genau das Gegenteil bedeuten könnte? Kuhn, *Text und Theorie*, 52

[55]A full list is given in Appendix E.

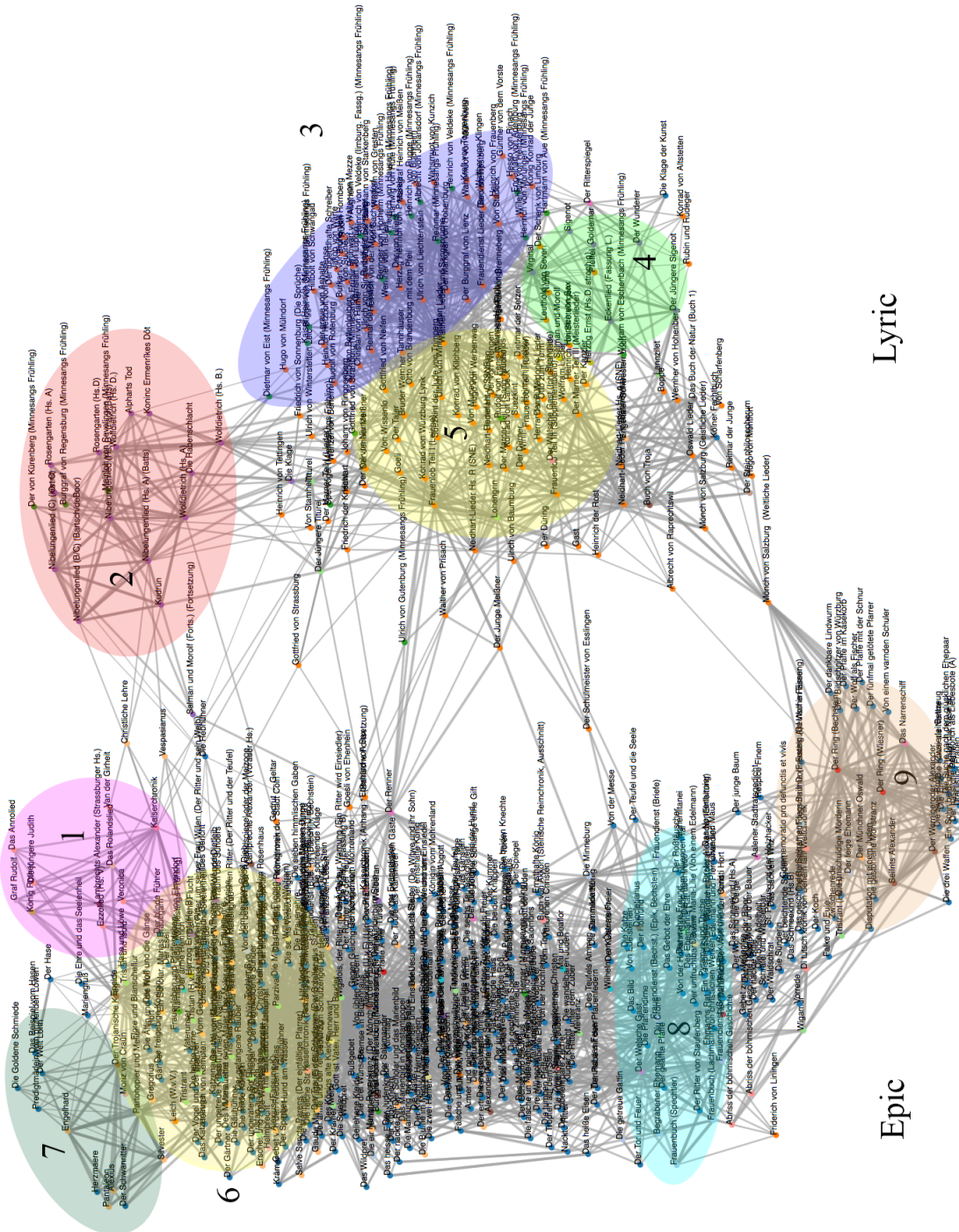

Figure 4.1: Textual network based on formal features

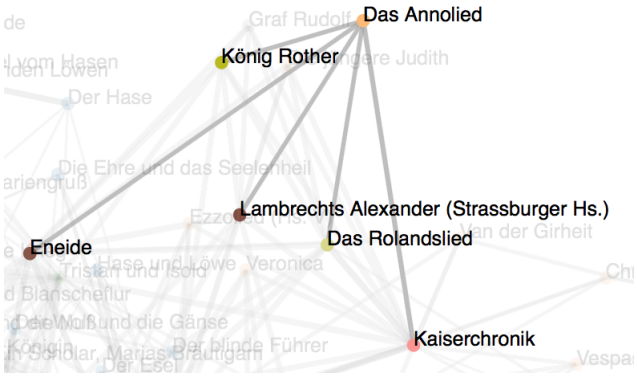

Figure 4.2: Connections for *Das Annolied*

To create this visualization, after the *tfidf* matrix is constructed, the cosine similarity is calculated for each text in the verse corpus to every other text in the verse corpus. For each text, a connection is drawn to its five most similar texts, and the shading of the connection is proportional to the strength of the similarity. Thus one text may have as little as five connections, or as many as there are texts in the corpus (in theory), naturally creating clusters around the most representative texts, i.e., those texts with a higher density of mutual prosidic feature sequence 10-grams. A common subsequent task would be to implement an unsupervised clustering algorithm in order to hard assign texts to specific groups, and then extract the distinct features of these groups. In the interest of flexibility and interpretation, I intentionally choose against this technique, as there are no hard categories in MHG, and to foster a more organic reading of the MHG verse corpus. My reading leads us roughly chronologically through form developments in the MHG period, beginning at the top with texts from the early and mid-12th century, branching out into lyric and epic, and resolving in late 14th and early 15th century lyric and epics.

The top of the network contains the oldest texts, which already split into two distinct forms. *Das Annolied* and *Das Rolandslied* comprise the base of Group 1, shown in Figure 4.2: The *tfidf* values for these texts reveal that there is greater variability in prosidic formal sequences, which will become clearer in a comparison of *tfidf* values for other groups, and that the bisyllabic open and closed syllable sequence is more characteristic of this text group than the rhyme sequence, which does not appear in the top five *tfidf* values for *Das Annolied* in Figure 4.4. Notably, *Das Annolied* was written in the early 12th century near Siegburg in close proximity to French writers, and *Das Rolandslied* is a translation of the Old French *Chanson de Roland*. These texts funnel into the early courtly epics of *Tristan*, *Erec*, and *Iwein* in Group 5, to which I will return. In an entirely different grouping, we see the Germanic heroic epics in the tradition of the *Nibelungenlied* appear in Group 2. Although theoretically authored contemporaneously, the Germanic epics develop a distinct form, referred to as the *Nibelungenstrophe* (the Nibelung strophe). It's *Langzeile* (long line) style distinguishes it significantly from the rhyme pair tradition:

sequence	*tfidf*
-C-C-C-C-C	0.042792
C-OC-C-C-C	0.042338
C-C-C-C-C-	0.042296
O-C-C-C-C-	0.041430
-OC-C-C-C-	0.040294

Table 4.3: Top five *tfidf* values for *Das Annolied*

> The difference between rhyme pair and strophic form concerns not only the organization of the linguistic material (and thus the style), but above all the manner of presentation: Rhyme pair texts are, as far as we are concerned, principally for the presentation of spoken word, strophic texts on the other hand are suited to song (with instrumental accompaniment). Thus it is a very clear formal gesture when the author of a heroic piece decides to implement a rhyme pair form, going against tradition.[56]

This belief is still supported today.[57] We therefore see a logical split between the lyric and epic poetry, with the *Nibelungenstrophe* beginning the grouping of what is believed to have been primarily sung poetry.

It is no coincidence that we see suspected, though heavily disputed, *Nibelungenlied* author Der von Kürenberg, as well as fellow early *Minnesänger* Meinloh von Sevelingen and Burgraf von Regensburg all appear in this *Nibelungenstrophe* cluster. In fact, across the entire corpus, the *Nibelungenlied* manuscripts are all the most similar texts to Der von Kürenberg's poetry[58]:

The unusual four syllable word with all open syllables is prevalent in Der von Kürenberg's few existent songs—'gezamete', 'gevidere', 'lügenaere', always bookended by a closed syllable. In fact, Der von Kürenberg's most famous song 'Ich zôch mir einen valken' includes two verses with such a construction, as he connects the inner verses leaving the outer verses connected by end rhyme:

> Ich zôch mir einen valken mêre danne ein jâr.

[56]"Der Unterschied zwischen Reimpaar- und Strophen- Form betrifft nicht nur die Organisation des Sprachmaterials (und damit den Stil), sondern vor allem auch die Art des Vortrags: Texte in Reimpaaren sind, soweit wir sehen, grundsätzlich für den Vortrag mit Sprechstimme, Texte in Strophen hingegen für Gesangsvortrag (mit Instrumentbegleitung) bestimmt. So ist es eine sehr deutliche Formgeste, wenn sich der Verfasser einer Heldendichtung gegen den Hauptstrom der Tradition für die Reimpaar-Form entscheidet." Joachim Heinzle, *Einführung in die mittelhochdeutsche Dietrichepik*, De Gruyter Studienbuch (Berlin ; New York: W. de Gruyter, 1999), 64

[57]Kragl, "wort unde wîse. Formen des sangbaren Verses in der deutschen Literatur des Mittelalters," 51.

[58]This observation already underlines the significance of this method in measuring prosodic form separate from content, as Der von Kürenberg's poetry had nothing to do with the *Nibelungen* saga thematically.

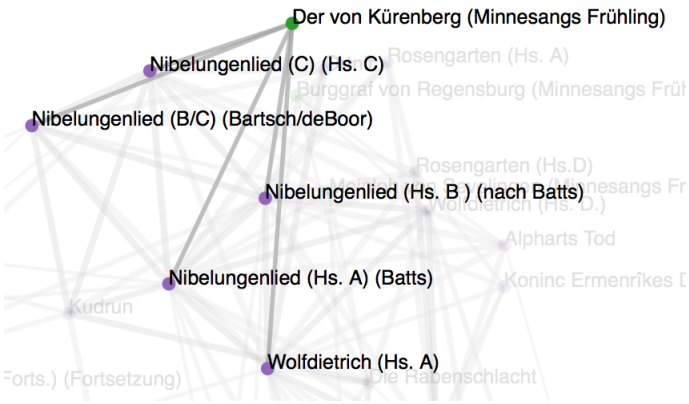

Figure 4.3: Connections for Der von Kürenberg

sequence	*tfidf*
C-OOOO-C-C	0.091087
-C-C-OO-C-	0.089309
-C-OO-C-C-	0.085931
C-C-OOOO-C	0.085273
-C-C-XC-CO	0.077646

Table 4.4: Top five *tfidf* values for Der von Kürenberg

dô ich in gezamete, als ich in wolte hân,
und ich im sîn gevidere mit golde wol bewant,
er huop sich ûf vil hôhe und vlouc in anderiu lant.[59]

The *Nibelungenlied* is marked by sequences ending with a bisyllabic open closed syllable, as shown in Table 4.5. Due to the longer lines of the *Nibelungenstophe*, sequences without end rhyme become more prevalent. Yet both the *Anvers* (first half of the line) and *Abvers* (second half of the line) of the *Nibelungenstrophe* have more specific characteristics. Nearly two thirds of the occurences of the sequence with the highest *tfidf* ('-C-C-C-OC-') occur in the *Abvers*, while two thirds of the occurences of the sequence with the third highest *tfidf* ('C-C-C-OC-C') occur in the *Anvers*.[60] This relationship between verse or metrical structure and phonological sequence will be further explored in the next chapter.

It is worth highlighting here a further advantage of the algorithm employed, evidenced by the clustering and similarity of the *Nibelungenlied* manuscripts. Despite different editions and

[59] "I raised a falcon for more than year. After I tamed it, as I wanted to have it, and I bestowed its feathers with gold, it picked itself up, took off, and flew into a foreign land."

[60] 365 of the 549 occurences of '-C-C-C-OC-' occur in the last half of the verse, and 360 of the 544 occurences of 'C-C-C-OC-C' occur in the first half of the verse.

sequence	*tfidf*
-C-C-C-OC-	0.110758
-C-C-OC-C-	0.109377
C-C-C-OC-C	0.105543
C-C-C-C-C-	0.099511
C-C-C-C-OC	0.081710

Table 4.5: Top five *tfidf* values for *Das Nibelungenlied* (B/C, Bartsch/deBoor)

text	similarity to *NL* (B/C, Bartsch/deBoor)
Nibelungenlied C (Hennig)	.983
Nibelungenlied B (Batts)	.961
Nibelungenlied A (Batts)	.952
Kudrun	.879
Wolfdietrich	.877

Table 4.6: Five most similar texts to *Nibelungenlied* B/C, Bartsch/deBoor

dialects of the text from the manuscripts, this representation of form and sound understands the *Nibelungenlied*'s formal essence, finding a distinct highest similarity among the four editions shown in Table 4.6. It appears that the "Texteingriffe" (text interventions) by Lachmann and others may not influence a large-scale "text morphology" approach rooted in abstracted formal features. This approach certainly does not attain the 'true' form of a poem, but the generalization of it can still place it in its correct position *relative* to the rest of the corpus. Importantly, the manuscript is not the actual performance either. Lachman 'correcting' a manuscript in the end may not be much more of a change than that which any scribe would make to the actual poem or performance. Joachim Bumke reminds us that everything is mediated. If the "text morphology" approach does not change the poem beyond recognition, this analysis can still be used to draw *relative* conclusions.[61]

One of the earliest *Minnesänger*, Dietmar von Eist, leads away from the *Nibelungenstrophe* into a distinct group of lyric poetry. Dietmar distinguishes himself in his use of bisyllabic closed and open syllables propelled by 'her-ze' in eight of the 23 sequences, a reverse of the *Nibelungen* cluster, although preserving the bisyllabic open closed syllable sequence with a lower *tfidf*, and moving beyond the couplets with the addition of ABAB rhyme, seen in Table 4.7. Dietmar was one of the last poets of the first *Donauländische* phase of *Minnesang*. After the influence of the troubadours arrived in Germanic lands, *Minnesang* began using

[61] Bumke, "Der unfeste Text" This may also hold true for the problem of *mouvance*—determining which stanzas belong to which song, stanza ordering, and general variation between manuscripts, addressed in the first chapter.

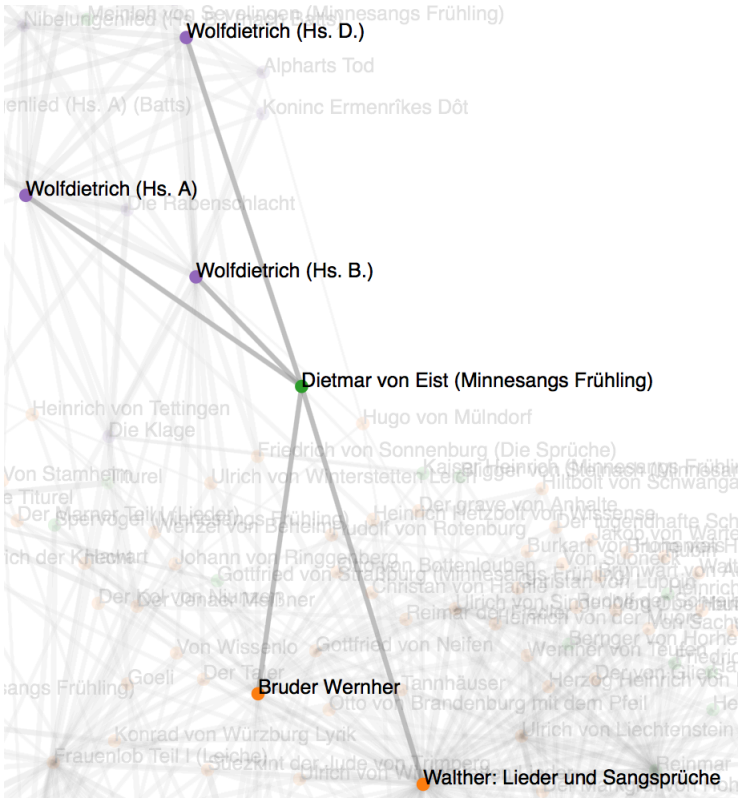

Figure 4.4: Connections for Dietmar von Eist

sequence	*tfidf*
C-C-C-CO-C	0.109651
-C-C-CO-C-	0.100787
C-C-C-C-C-	0.099609
-C-2C-C-C-	0.097086
C-C-C-OC-C	0.095398

Table 4.7: Top five *tfidf* values for Dietmar von Eist

multiple stanzas, and the *Langzeile* began to disappear.[62]

The lyric grouping is divided into several subgroups. Along the right side in Group 3 is primarily early and classical *Minnesang*, most of which is connected to Reinmar der Alte, who, unlike Walther von der Vogelweide and others, composed nearly exclusively *Minnesang*. Reinmar's highest *tfidf* sequences are in Table 4.8. Reinmar's use of monosyllabic closed syllables increases over Dietmar's, as well as the importance of ABAB rhyme coupled with

[62]Rettelbach, "Minnelied und Sangspruch: Formale Differenzen und Interferenzen bei der Tonkonstitution im 13. Jahrhundert," 153.

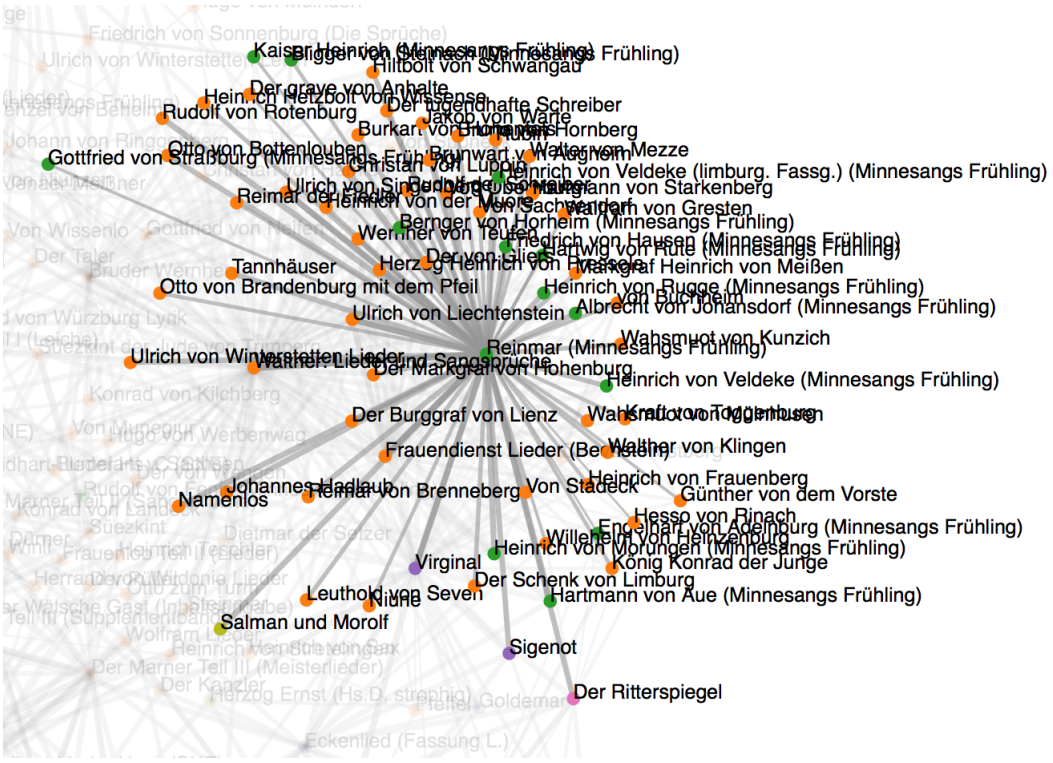

Figure 4.5: Connections for Reinmar der Alte

sequence	*tfidf*
C-C-C-C-C-	0.232212
-C-C-C-C-C	0.190334
-C-2C-C-C-	0.133343
C-2C-C-C-C	0.113482
C-C-2C-C-C	0.112257

Table 4.8: Top five *tfidf* values for Reinmar der Alte

monosyllabic, closed syllable words concluding the rhyme, and beginning the next line.

Below *Minnesang* in Group 4 is clearly the 'Berner Ton' (Bern sound), named not for the place of origin, but rather for the specific meter and form used for certain heroic epics about Dietrich von Bern. *Virginal* is closest to the *Minnesang* grouping, while *Das Eckenlied*, *Sigenot*, *Goldemar*, and *Der Wunderer* begin to distance themselves from the love poetry.

The 'Berner Ton' is a particularly artful form, comprised of 13 lines with very specific accent counts, cadence types, and rhymes for each line.[63] It is, in a way, an evolved form of

[63]Heinzle, *Einführung in die mittelhochdeutsche Dietrichepik*, 100.

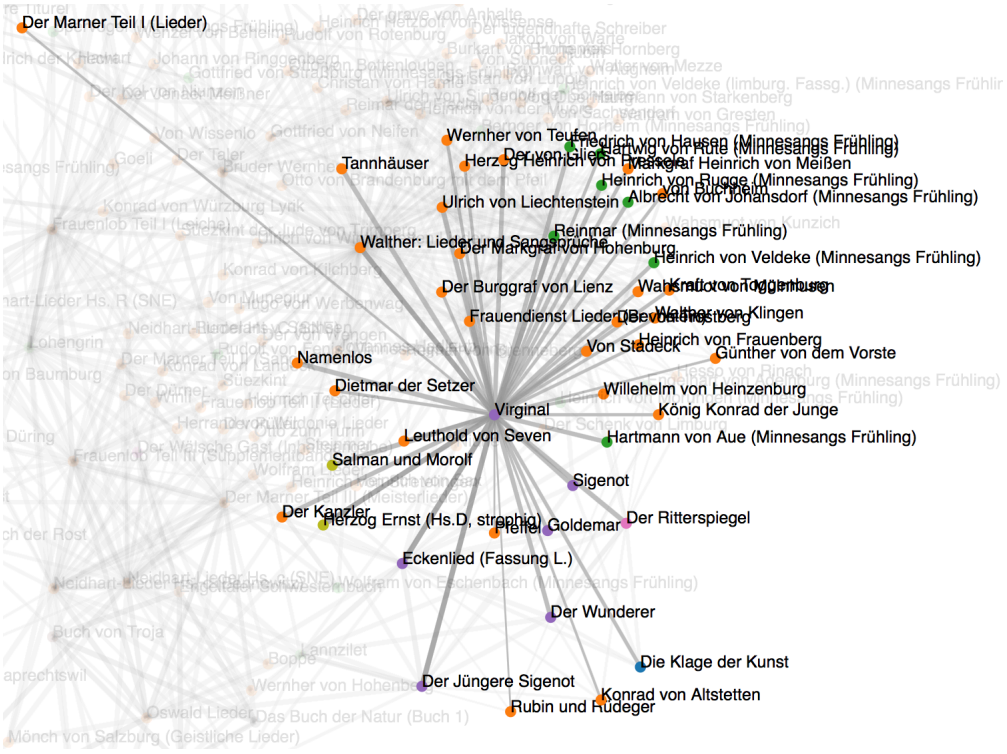

Figure 4.6: Connections for *Virginal*

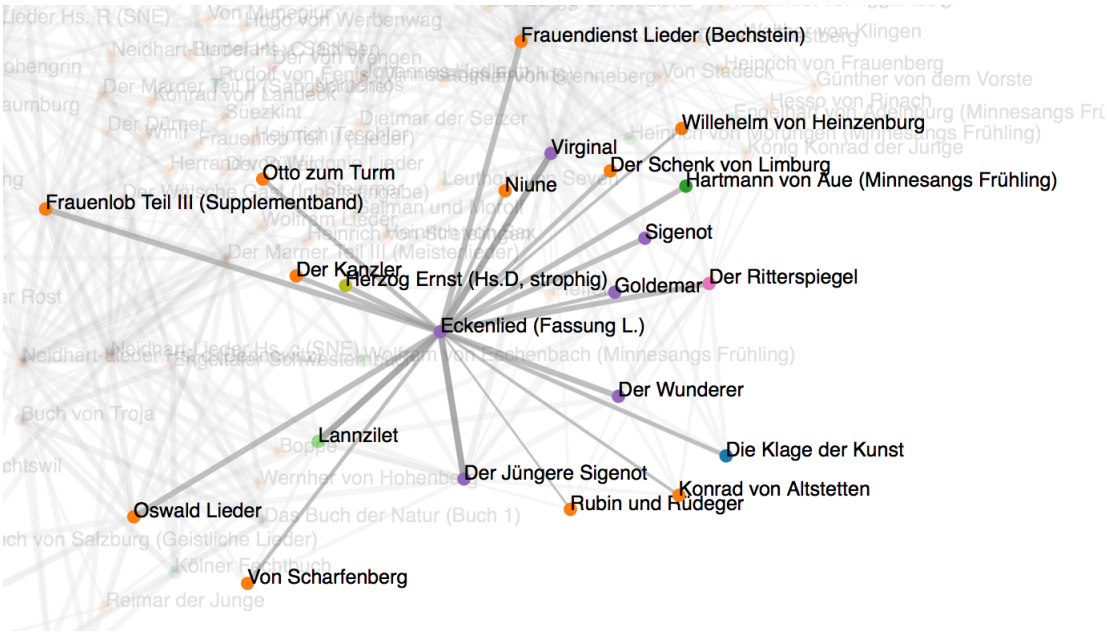

Figure 4.7: Connections for *Das Eckenlied*

the *Nibelungenstrophe* (both in form and content) combined with the rhyme-pair technique of courtly epics and a few other heroic epics. As discussed in Chapter 1, the move away from rhyme pairs was thought to be more suitable to musical accompaniment, and the placement of the 'Berner Ton' texts in the middle of the MHG lyric would appear to support this hypothesis, with *Virginal* closer to the style of the *Minnesänger*, and *Das Eckenlied* closer to *Meistergesang*. Given the plot of *Virginal*, its relationship to *Minnesang* can come as no surprise:

> Sî sprach <vil werder dieterîch,
> nu sînt nâch sorgen vröuden rîch,
> sît iuch vrou saelde minnet.
> durch die ir schaden hânt genomen,
> zuo den vrouwen sülnt ir komen:
> nement daz ir dâ gewinnet.
> dâ sehent ir manegen rôten munt
> smieren ûz liehten wangen,
> lachen ûz ir herzen grunt
> (dâ wert ir schône enpfangen),
> von sîden manec rîch gewant:
> dar varent und nement verdienten lôn,
> den hât ervohten iuwer hant.>[64]

Keeping the 'Berner Ton', which itself parallels much of *Minnesang* form, the author of *Virginal* retains the *inquit* formula, and employs much *Minnesang* imagery ('rôten munt', 'liehten wangen', earlier we see 'wengelîn und mundel rôt', and later 'rôten mundelîn'), which is either less prevalent or non-existent in the other Dietrich epics.

The side of the lyric poetry closer to the epic poetry in Group 5 appears to be a greater mix of form, though much more *Leich* (lai) and *Sangspruchdichtung* (sung word) than classical *Minnesang*. These texts are likely pulled toward the epic texts by a content-driven form differential, which I will discuss below. Bruder Wernher is the representative at the top of the group, true to his "altmodischen Kennzeichen" (old-fashioned features) through use of the *Langzeile*, which is balanced by his *Kanzone* (canzone) constructions.[65] The core representative of Group 5 is Frauenlob's *Leich* and in the middle of these lyric groups is Walther von der Vogelweide, encompassing the broad spectrum of lyrical form and genre.

[64]She spoke "very worthy Dietrich, now joy is abundant after worries, since the lady loves you well, through whom you have taken much harm, you should come to the lades: take what you win there. There you will see many red mouths smiling with light cheeks, laughing from their heart (there you'll be well received), dressed elegantly in silk: go there and take the reward you deserve, won by your own hand!" Albrecht von Kemenaten and Julius Zupitza, "Virginal," in *Dietrichs Abenteuer* (Berlin: Weidmann, 1870), 119:1-13

[65]Rettelbach, "Minnelied und Sangspruch: Formale Differenzen und Interferenzen bei der Tonkonstitution im 13. Jahrhundert," 153.

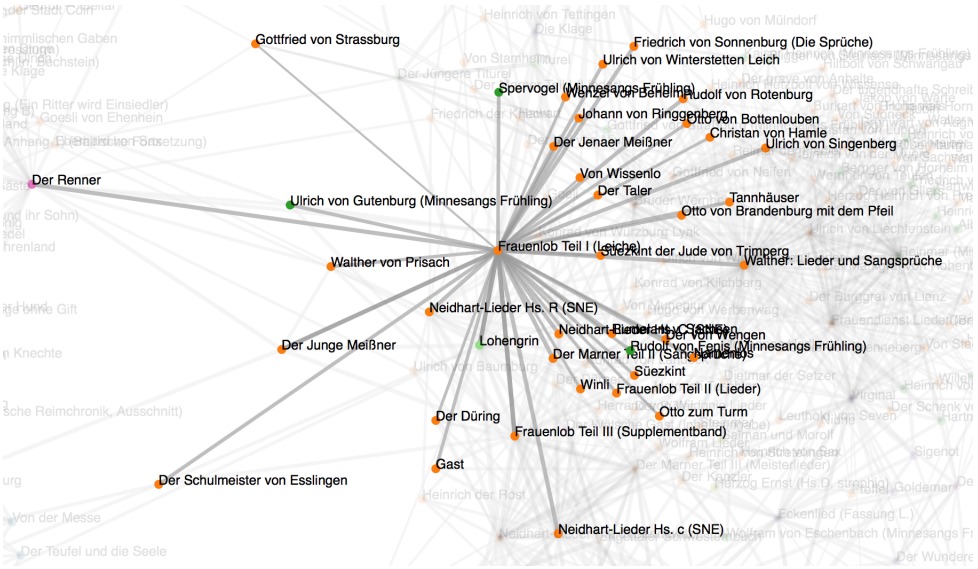

Figure 4.8: Connections for Frauenlob's *Leich*

sequence	*tfidf*
C-C-C-C-C-	0.292145
-C-C-C-C-C	0.230379
C-C-C-CC-C	0.138196
C-C-C-OC-C	0.130636
-C-C-C-C-O	0.120531

Table 4.9: Top five *tfidf* values for Frauenlob's *Leich*

On the other side of the visualization, the early MHG texts with French influence lead into the general rhyme pair epic poetry, beginning with the early works such as Gottfried's *Tristan* and Heinrich von Veldeke's *Eneide*. Fittingly, Heusler claims Hartmann and Gottfrid were heavily influenced by Veldeke, particularly concerning overfilling lines and feet: "His *Eneit* touted the new courtly form through its new courtly substance. The French rhyme pair really resounded in his hears, and had at the least made him sensitive to overfilling."[66] The classic courtly epics of *Erec*, *Iwein*, and *Parzival* are all in the upper part of this cluster in Group 6, as is Wolfram's heroic epic *Willehalm* fashioned in the form of a courtly epic, which maintains strong connections to the Dietrich epics *Biterolf und Dietleib*, *Dietrichs Flucht*, and *Laurin*.

[66] "Seine Eneit warb durch den neuhöfischen Inhalt für die neuhöfische Form. Ihm haben die welschen Reimpaare wirklich im Ohre geklungen und ihn wenigstens nach der einen Seite, gegen die schwere Füllung, empfindlich gemacht." Heusler, *Deutsche Versgeschichte: mit Einschluss des altenglischen und altnordischen Stabreimverses*, 101

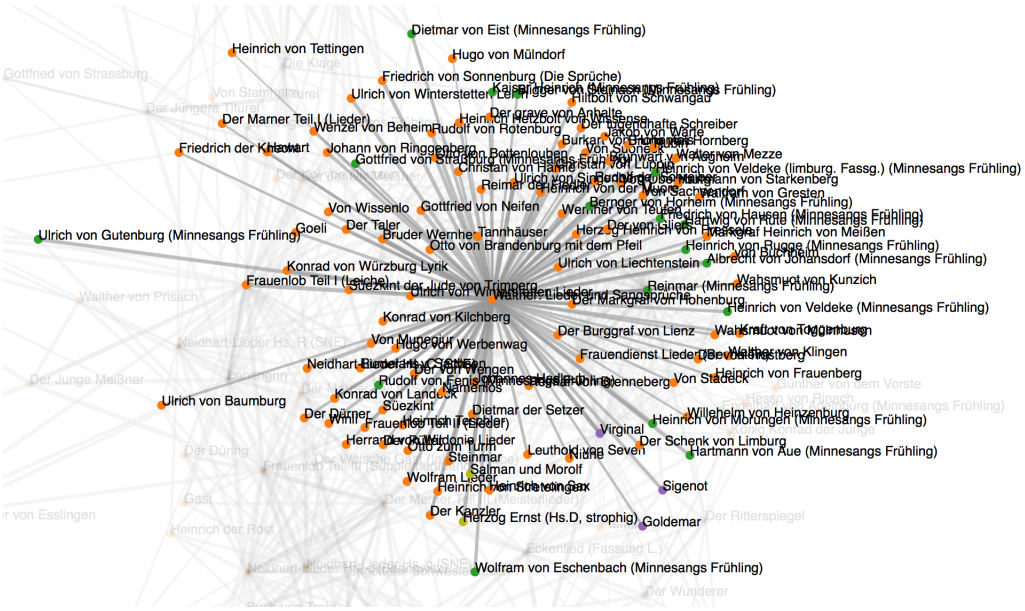

Figure 4.9: Connections for Walther von der Vogelweide

Striking is the cluster of texts in Group 7 above the epic poetry, all attributed to Konrad von Würzburg. *Engelhard*, *Alexius*, *Die Goldene Schmiede*, *Heinrich von Kempten*, *Herzmaere*, *Das Turnier von Nantes*, *Partonopier und Meliur*, *Pantaleon*, *Silvester*, *Der Schwanritter*, *Der Trojanische Krieg*, and *Der Welt Lohn* all group very close to one another—despite the fact they are otherwise non-descript *Vierheber*[67]. What distinguishes Konrad von Würzburg's *Vierheber* from the rest? Heusler argues that Konrad is unique in his use of *klingende*, or ringing, cadences in the style of Otfrid's Old High German:

> So much is certain, Konrad von Würzburg's verse is free of female full cadence. He accordingly made the narrower, clearer choice. His $- \acute{-} - \times$ endings are all ringing, just as with Otfrid. [...] Konrad is in this respect old-fashioned when it comes to sound effect; and he is not the end of the line. Did his affection for sound play and rhyme artistry draw him to a form that brings rhyme into closer proximity?[68]

[67] A term used for the MHG epic meter, in which each line usually has four stressed syllables (*Hebungen*, or 'lifts')

[68] So viel steht wohl fest, daß Konrad von Würzburg von weiblich vollen Versen frei ist. Er hat auch hierin die engere und klarere Wahl getroffen. Seine $- \acute{-} - \times$ Schlüsse sind sämtlich klingend, so gut wie bei Otfrid. [...] Konrad ist in diesem für die Schallwirkung bedeutenden Punkte altertümlich; hierin ist er nicht der Abschluß der Reihe. Zogen ihn Gefallen am Klangspiel und Reimgewandtheit zu der Form, die die Reime in kürzeren Abständen bringt? Heusler, *Deutsche Versgeschichte: mit Einschluss des altenglischen und altnordischen Stabreimverses*, 132, 135

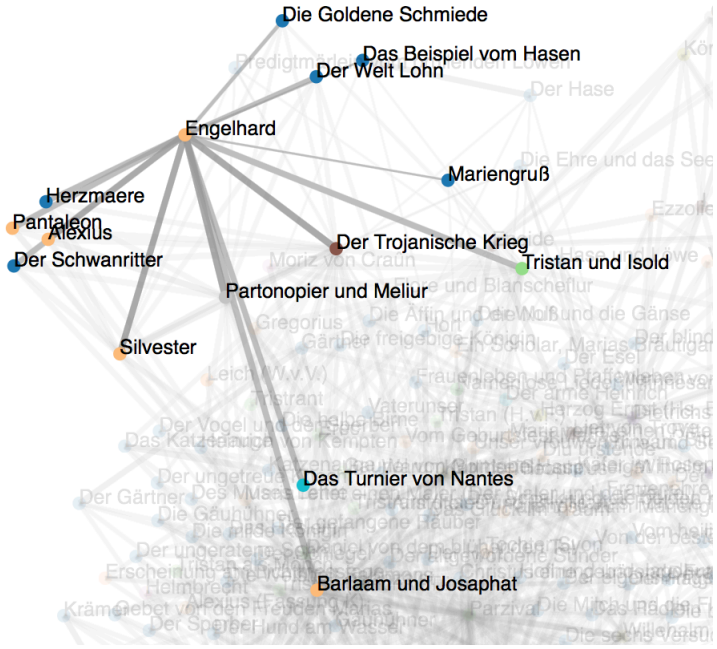

Figure 4.10: Connections for Konrad's *Engelhard*

Heusler poses a compelling question, one that is applicable throughout Konrad's works. Did his interest in sound, and perhaps its setting to music, bring him to shorten the duration between his formal play? According to Zumthor's reading, this could potentially realize "a certain density", by which the sound itself takes over and generates new meaning.[69] In fact, it can become so dense, that the performance becomes a complete sound experience, which strives to free itself from the language.[70] The simple monosyllabic alternation of closed syllables, so integral to *Alexander* and other MHG courly epics, is not as frequent as other sequences in Konrad's work, even 'C-C-C-OC-C', a bisyllabic open-closed word somewhere in the middle of a foot is more characteristic.[71] The "old fashioned" quality of Konrad's work is also illuminated by its position toward the top of the visualization.

Below the large grouping of *Vierheber* are three more noteworthy texts in Group 8: Thomasîn von Zirclaere's *Der Welsche Gast* and Ulrich von Liechtenstein's *Frauendienst* and *Frauenbuch*. The importance of monosyllabic closed syllable words increases in the bottom half of the epic grouping. As has been pointed out by Ranke and Heusler, there appear to be few if any *klingende* cadences in *Der Welsche Gast*, and at first glance this appears to be the case for *Frauendienst* as well, tied to the more serious discovery here of an increasing number of monosyllabic, closed sequences. No open-syllable sequences appear in the top five *tfidf* values for either text, implying major consequences for the soundscape.

[69] Zumthor and Engelhardt, "The text and the voice," 86.
[70] Zumthor, *Die Stimme und die Poesie in der mittelalterlichen Gesellschaft*, 73-74.
[71] This discussion is continued in the metric investigation in the next chapter.

sequence	*tfidf*
-C-XC-C-C-	0.108341
C-XC-C-C-C	0.107090
C-1C-C-C-C	0.099839
C-C-C-OC-C	0.090290
C-C-C-C-C-	0.084167

Table 4.10: Top five *tfidf* values for Konrad's *Engelhard*

Figure 4.11: Connections for Ulrich von Eschenbach's *Alexander*

sequence	*tfidf*
C-C-C-C-C-	0.175153
C-1C-C-C-C	0.132938
C-XC-C-C-C	0.118543
-C-XC-C-C-	0.114662
-C-C-C-C-C	0.108502

Table 4.11: Top five *tfidf* values for Ulrich von Eschenbach's *Alexander*

sequence	*tfidf*
C-C-C-C-C-	0.253562
C-1C-C-C-C	0.154692
-C-C-C-C-C	0.153595
-C-1C-C-C-	0.152189
-C-XC-C-C-	0.147248

Table 4.12: Top five *tfidf* values for Thomasîn von Ziclaere's *Der Welsche Gast*

sequence	*tfidf*
C-C-C-C-C-	0.289783
-C-C-C-C-C	0.195649
-C-XC-C-C-	0.149504
-C-1C-C-C-	0.148114
C-XC-C-C-C	0.147530

Table 4.13: Top five *tfidf* values for Ulrich von Liechtenstein's *Frauendienst*

This and other issues related to meter will be picked up in the following chapter, though it suffices here to state that early philologists believed that Thomasîn and Ulrich did not make use of the *klingende* cadence, the opposite case of Konrad von Würzburg above, and thus logically appear on the other end of the epic formal clustering.[72]

This downward direction follows chronologically into very late MHG and early ENHG. Cramer argues that the importance of language grew and that this was especially true in later *Minnesang*, when language overtook content as the main focus of these songs. In fact, he argues that language became so powerful that songs could nearly be generated by a

[72]Heusler, *Deutsche Versgeschichte: mit Einschluss des altenglischen und altnordischen Stabreimverses*, 132-133; Friedrich Ranke, *Sprache und stil im Wälschen gast des Thomasin von Circlaria*, vol. 68 (Mayer & Müller, 1908); Although Eva Willms claims that they are not missing at all ("Sie fehlen keineswegs") and that depending on the anacrusis, many more can be counted than Ranke and Heusler consider. Thomasin and Eva Willms, *Der Welsche Gast*, De Gruyter Texte (Berlin ; New York: W. De Gruyter, 2004), 10

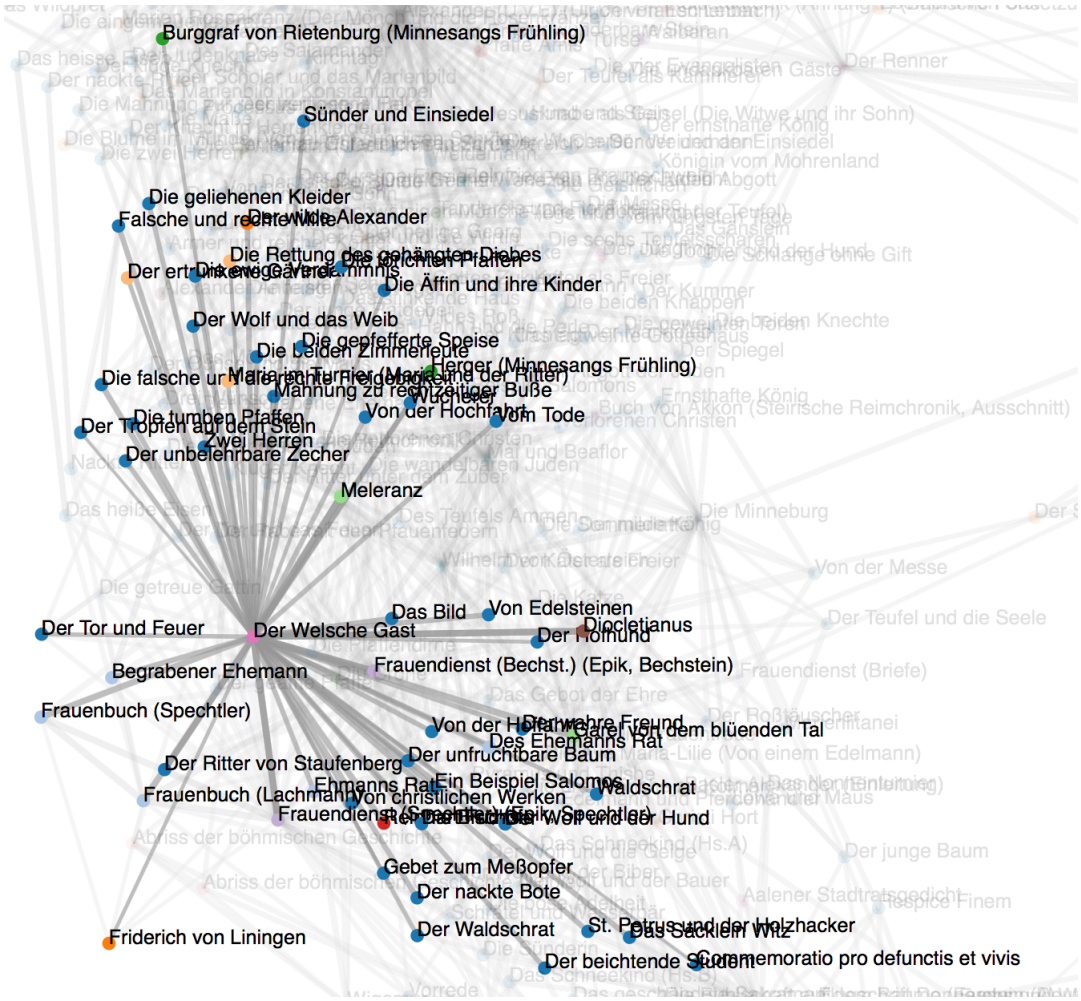

Figure 4.12: Connections for Thomasîn von Ziclaere's *Der Welsche Gast*

"Sprechmaschine" (speaking machine), and that a speaking partner became superfluous.[73] This "Sprechmaschine" is manifested in these simplifications, which Braun argues occurred much earlier, as even Konrad von Würzburg, Der Kanzler, and Ulrich von Liechtenstein do not have as complex of a structural form as the early *Minnesang* of Reinmar and Friedrich von Hausen.[74] This is illustrated in the force directed graph, as the later *Minnesang* fuses with general *Sangspruchdichtung*, as Braun also identifies, and begins to push toward the epic rhyme-pair tradition. As Heusler notes on several occasions, MHG epic verse experienced severe simplification in cadence and rhyme, which is shown by the resolution of the two

[73]Cramer, *Waz hilfet âne sinne kunst?: Lyrik im 13. Jahrhundert Studien zu ihrer Ästhetik*, 180.

[74]Braun identifies this in less hypotaxis and more parataxis.Braun, "Aufmerksamkeitsverschiebung. Zum Minnesang des 13. Jahrhunderts als Form- und Klangkunst," 215-216

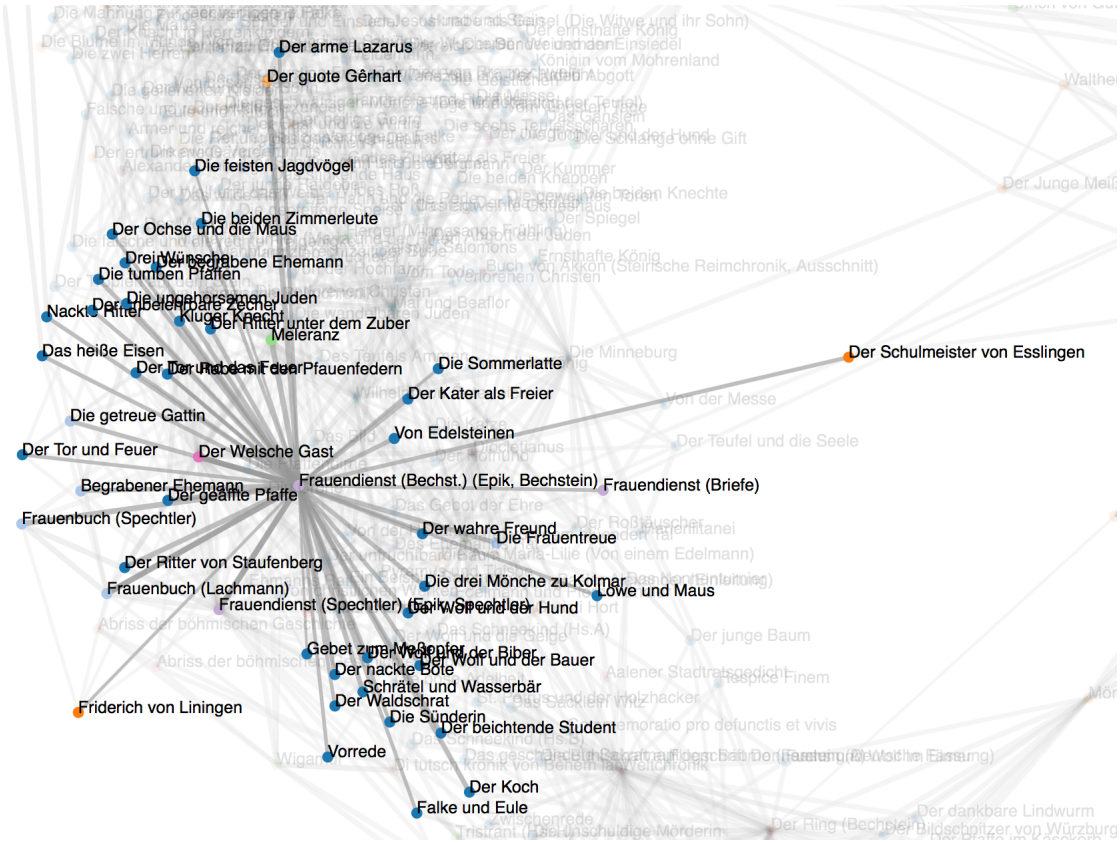

Figure 4.13: Connections for Ulrich von Liechtenstein's *Frauendienst*

forms in the bridge of the late 14th century Mönch von Salzburg.[75] Hugo von Montfort, active in the late 14th century, still remains in the lyrical grouping for his use of ABAB rhyme, but there is a significant increase in the closed, monosyllabic sequences, intensified by Mönch von Salzburg. Mönch von Salzburg's *tfidf* values resemble very much those of Heinrich Wittenwiler's early 15th century *Der Ring* in Group 9. This also coincides with more dense and rigid feet in the epic meter, as noted by Heusler.[76]

Interpretation

This method of clustering formal features measures several aspects of a text:

- **Rhythm and meter**: Because the syllable serves as the foundation of MHG meter,

[75]Much to Heusler's chagrin: "A persistent up and down is in its origin an un-germanic principle" ("Durchgeführtes Auf und Ab ist seinem Ursprung nach ein ungermanischer Grundsatz...") (Heusler, *Deutsche Versgeschichte: mit Einschluss des altenglischen und altnordischen Stabreimverses*, 9

[76]Heusler, *Deutsche Versgeschichte: mit Einschluss des altenglischen und altnordischen Stabreimverses*, 120.

sequence	*tfidf*
C-C-C-C-C-	0.309054
-C-C-C-C-C	0.203436
C-2O-XC-C-	0.153950
C-2C-C-C-C	0.135940
-XC-C-C-C-	0.126977

Table 4.14: Top five *tfidf* values for Hugo von Montfort

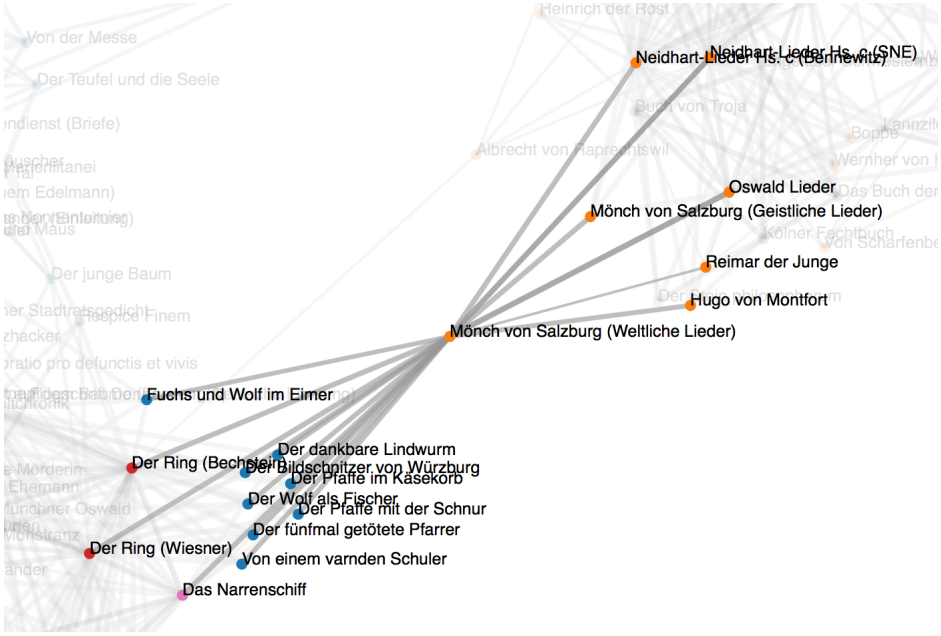

Figure 4.14: Connections for Mönch von Salzburg

sequence	*tfidf*
C-C-C-C-C-	0.409733
-C-C-C-C-C	0.283525
C-XC-C-C-C	0.154582
-C-XC-C-C-	0.137166
C-C-C-OC-C	0.135579

Table 4.15: Top five *tfidf* values for Mönch von Salzburg

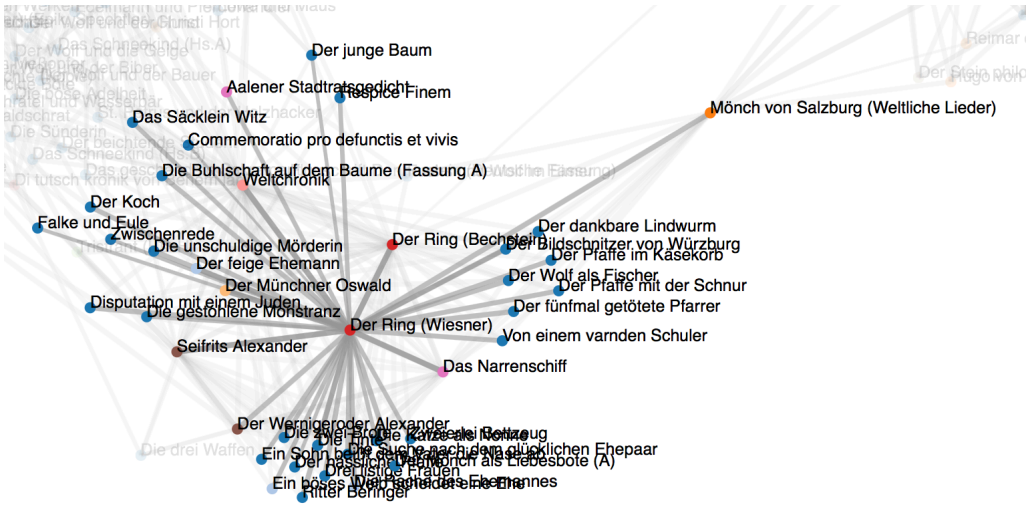

Figure 4.15: Connections for Heinrich Wittenwiler's *Der Ring*

sequence	*tfidf*
C-C-C-C-C-	0.366309
-C-C-C-C-C	0.242072
C-XC-C-C-C	0.141312
-C-XC-C-C-	0.138407
C-C-XC-C-C	0.136630

Table 4.16: Top five *tfidf* values for Heinrich Wittenwiler's *Der Ring*

this method is sure to discern metrical similarity between texts, which is evidenced in the grouping of the *Nibelungenstrophe*, lyrical poetry, the *Berner Ton*, and more general rhyme-pair epic poetry. The precise metrical scheme is slightly more difficult. All of MHG poetic meter is dependent upon on syllables, syllable properties, word boundaries[77], and word stress. The method presented captures nearly all of this except for stress. The next chapter will make a more sophisticated approach to capture this quality as well, though only for the epic poetry.

- **Sonority**: An aspect of syllable quality, sonority will be accounted for both indirectly in a quantitative manner, and directly in a sequential manner. Accordingly, texts sharing unique strings of open syllables will be more similar. Texts with more open syllables will inevitably have a higher ratio of vowels to consonants. Whether this can accurately discern the extent to which a song is sung, is a task still left to be undertaken.

[77]MHG is a stress initial language, and cadence is dependent on word boundries

- **Linguistic Complexity**: With word boundaries as features in the sequence, the size of words will also be drawn into the analysis. Due to both a simplification of form, as well as schwa reduction through the period, different stages of the German language had different syllable per word ratios, which manifests in the sequences captured above. Nevertheless, this analysis has shown that this does not detract significantly from a formal comparison.

This method accounts for most orthographical variants of MHG. While naturally not all variations can be considered, if we accept the assumption that authors composing verse were aware of metrical and rhythmical restraints, then it is *indifferent* whether one author was writing in Alemannic or Frankish. In order to keep the rhythm, the scribe or author would have adapted their own dialect to conform to the basic syllable properties measured, and this would result in a text similar to the model they desired (as supported by multiple versions of the same text for the *Nibelungenlied*, or multiple texts by the same author with high similarity as with Konran von Würzburg). One important consequence of this assumption is the requisite assumption that scribes were aware of meter and rhythm. I argue, based on my discussion of memory in Chapter 1, that this may be one of the few assumptions of scribes we can make confidently. If a poet or scribe forgot verses or even words, or was composing new lines as in the oral tradition, he or she likely fit whatever context they had into the rhythm of the text (recalling the discussion of formulas), and thus changing the text itself, but not the rhythm of the tradition. In this sense, this method may be more accurate in extracting a more authentic essence of the original precisely by abstracting from it.[78]

4.2 Form and Content

While there is much more to be worked out from this analysis, I do not wish to overstep the aims of this project in presenting new methods. To this end, I propose one extension of the above analysis to better understand the relationship between form and content in MHG literature. I start by posing a single question: What if the method presented does not abstract *enough*? What if these formal sequences texts share are actually exact lexical, or formulaic, matches? For example, the *Nibelungenlied* manuscripts and editions have a high similarity to one another, but what if this is not the rhythm and sound being captured as prototypically '*Nibelungenisch*', but rather the actual word phrases incrementing the *tf* of the *tfidf* matrix? In fact, for the *Nibelungen* cluster, this *should* be the case. But what if

[78]The obvious problem with these assumptions is scribal or digitization errors. This is a pitfall of most computational text analysis techniques and the only remedy is a larger corpus in hopes to drown out the noise. A further problem is text size. If a text is too small, a mere ten lines for example, there will not be enough sequential data when the data are normalized to generate useful measures. This is evident in the results, where certain short texts cluster together. On the opposite end, very long texts also have more opportunity to generate sequences that relate to other texts, especially if the sequence is relatively scarce in the corpus. While the L2 normalization of the *tfidf* values will account for these problems to a degree, it is still something to keep in mind with these sorts of analyses.

	seq. (10-grams)	lemmata	seq. (10-grams)	lemmata
Unigrams	.909 (9)	.916 (4)	.916 (4)	.910 (8)
Bigrams	-	.881 (3)	-	.878 (4)
Trigrams	-	.839 (3)	-	.844 (2)

Table 4.17: Cosine similarity for *Iwein* vs. *Erec* and *Parzival* across all methods

the group identified as the heroic epics by the formal measure is actually matching common phrases between the texts, hence tightly binding together the form and content? Is there a way to measure the separate contributions of each form and content to genre construction?

Removing Fuzzy Lexical Matches from Feature Sequence Matches

I first take three example texts, remove nearly all lexical matches that increment feature sequence frequencies via string matching, and recalculate the cosine similarities. For this task, I select Hartmann von Aue's *Iwein* and *Erec*, and Wolfram von Eschenbach's *Parzival*. These texts provide one of the best opportunities to test this hypothesis: two were written by the same author, so it can be hypothesized that the language, style, and vocabulary may be similar, and all three are Arthurian romances following the epic meter of the time and genre, many sharing elements. *Iwein* will be the lead test to modify due to its abundant transmission.[79] Table 4.17 shows *Erec* and *Parzival* with high similarity to *Iwein* in regards to both the formal method and the lexical method (presented below). Recalling that cosine similarity is calculated by indexing a *tfidf* matrix, similarity scores thus depend on frequencies of certain sequences (terms), weighted by their corpus wide frequency. My method to test for lexical matches involves identifying possible matches and excluding them from incrementing that sequence's frequency (*tf*) for *Iwein*, by removing the corresponding text to that sequence from the entire text altogether (thus a lower *tf* for the sequence in that text, but likely not affecting the *idf* of the sequence). Computationally, this is carried out by identifying all feature sequence matches between *Iwein* and the other two texts, and for each feature sequence match extracting all the corresponding character string matches for each text that were mapped to that sequence. Subsequently, each string of text characters for *Iwein* is compared to each string of text characters for *Erec* and *Parzival*, and if the Levenshtein ratio[80] is greater than .85, all the appearances of that exact text string in *Iwein* are removed. Some of what is removed included exact matches, orthographic variations, as well as very close matches:

[79] Of course, *Erec*'s sole transmission is in the 16th century *Ambraser Heldenbuch*.

[80] The Levenshtein ratio is defined by the Levenshtein distance (i.e., edit distance) divided by the alignment length. The Levenshtein distance is the least number of edits to one string of characters to match the other string of characters. Levenshtein, "Binary Codes Capable of Correcting Deletions, Insertions and Reversals"

Sequence: *CO-C-C-C-O*
'*herre ich hân von iu*'[81] ≈ '*herre ich hân von iu*'

Sequence: *-C-CO-C-XO*
'*wîp unde man â*'[82] ≈ '*wîp unde man ze*"

Sequence: *1C-C-C-OC-*
'*ob er noch rîters*'[83] ≈ '*ob der noch strîtes*'[84]

Every match with a Levenshtein ratio < .85 was retained:

Sequence: *-XOC-C-C-O*
'*iuwer zuht und iu*'[85] ≠ '*biutet gein dem ei*'[86] ≠ '*sînen zoum nam si*'[87] (etc.)

This theoretically retains formal matches that are not related lexically. Removing close lexical matches in *Erec* removed 2313 potential lexical matches of 112129 possible sequence matches (2.06%), which resulted in a loss of 10.23% of all sequence feature 10-grams. Removing close lexical matches in *Parzival* removed 1854 of potential lexical matches of 112129 possible sequence matches (1.7%), which resulted in a loss of 9.34% of all sequence feature 10-grams. The resulting new similarities are given in Table 4.19 and Table 4.20. Therefore, even with every remaining match in the *tfidf* inventories for *Iwein* and *Erec* and *Parzival* not sharing formulaic phrases, the formal similarity remains very strong. Naturally, similarity scores for *Iwein* to each text decreases as that thematic- or genre-specific formula (likely present in all most related texts) is removed from *Iwein*.

As a last test to understand the "power of the voice" and to analyze it "independent of content", I repeat the above process for the top 10 most similar texts to *Iwein* according to the formal sequencing measure, reducing the sequences in *Iwein* cumulatively.[88] Removing close lexical matches in *Iwein* from the top ten texts removed 18147 potential lexical matches of 112129 possible sequence matches (16.18%), which resulted in a loss of 40.65% of all feature sequences, implying that near 40% of *Iwein*'s formal features consist of lexical matches to at least one text in the top 10, truly identifying a group of thematically and formally related texts, or what many would label a 'genre'. The result is a list of the top 10 most similar texts

[81] "...sir, I have from you..."
[82] "...woman and man..."
[83] "...whether he still of a knight's..."
[84] "...whether he's still fighting..."
[85] "...your upbringing and..."
[86] "...exchange went to the..."
[87] "...took his bridle..."
[88] Zumthor, *Die Stimme und die Poesie in der mittelalterlichen Gesellschaft*, 63.

text	sim.
Rennewart	0.928854
Alexander (U.v.E) (Ulrich von Eschenbach)	0.923586
Wigalois, der Ritter mit dem Rade	0.923414
Parzival	0.916120
Daniel von dem blühenden Tal	0.915249
Barlaam und Josaphat	0.914278
Lanzelet	0.913639
Karl der Grosse	0.913152
Erec	0.908859
Willehalm (Wolfram)	0.905695

Table 4.18: Top ten cosine similarities before removal for *Erec*

text	sim.
Alexander (U.v.E) (Ulrich von Eschenbach)	0.907986
Wigalois, der Ritter mit dem Rade	0.906297
Rennewart	0.906072
Parzival	0.905858
Barlaam und Josaphat	0.899009
Daniel von dem blühenden Tal	0.896546
Karl der Grosse	0.896123
Willehalm (Wolfram)	0.894577
Lanzelet	0.893751
Erec	0.892356

Table 4.19: Top ten cosine similarities for *Iwein* after removal of *Erec* matches

to the reduced *Iwein*, one containing no sequences of lexical matching to its previous top 10. Table 4.21 shows that while 6 of the top 10 in similarity remained in the top 10 after removal, two particular texts, Wolfram's *Willehalm* and the anonymous *Biterolf und Dietleib*, surged in similarity rank. Both these texts clearly belong to the heroic epic genre in substance (though slightly less so for *Willehalm*), but were intentionally written in the Arthurian romance traditional rhyme pair form, and so, after lexical matches removed similarities within the thematic relations, fewer matches were made to the heroic epic, which decreased in similarity far less than the other Arthurian romances, such as Wirnt von Gravenberc's *Wigalois*. Two epic poems known for imitating the great poets before them, Ulrich von Türheim's *Rennewart* and Ulrich von Eschenbach's *Rennewart*, also fall sharply in ranking after matches are removed. Table 4.25 summarizes the correlation and rank changes with

text	sim.
Alexander (U.v.E) (Ulrich von Eschenbach)	0.907044
Rennewart	0.906273
Wigalois, der Ritter mit dem Rade	0.904758
Parzival	0.902403
Barlaam und Josaphat	0.897970
Karl der Grosse	0.896748
Daniel von dem blühenden Tal	0.895643
Lanzelet	0.895100
Erec	0.895042
Willehalm (Wolfram)	0.893535

Table 4.20: Top ten cosine similarities for *Iwein* after removal of *Parzival* matches

sequence	sim.
Parzival	0.771731
Willehalm (Wolfram)	0.766083
Tristan und Isold	0.763064
Alexander (U.v.E) (Ulrich von Eschenbach)	0.758367
Erec	0.756541
Der Renner	0.753834
Biterolf und Dietleib	0.753810
Barlaam und Josaphat	0.753227
Wigalois, der Ritter mit dem Rade	0.752769
Wilhelm von Wenden	0.751707

Table 4.21: Top ten cosine similarities after removal of top ten

each removal.[89] While correlation, rank change, and overlap decreases with removal of lexical matches, it is also clear that a genre can be identified through form itself, abstracting from its content. In other words, poets like Wolfram can imitate a genre through an intense engagement with form, not having as many suitable lexical phrases at hand. Table 4.21 clearly shows that even without intertextual, formulaic phrases, the Arthurian romances still group close together.

The specific matches removed and their frequency in *Iwein* and *Iwein*'s top 10 most similar texts are given in Table 4.23. This list alone comprises the essence of a MHG Arthurian romance, revealing in a ranked order not only the most common intertextual phrases and

[89]While Pearson's considers the magnitude of each similarity in calculating the correlation coefficient, Spearman's rank correlation was also calculated, which yielded similar results. This is relevant if one considers such a comparison from a practical, applied standpoint as being one of rank.

	r	overlap 10	overlap 20
Iwein	.999 (.999)	100%	85%
Parzival	.999 (.999)	100%	85%
Top 10	.991 (.991)	60%	70%

Table 4.22: Pearson's (and Spearman's) r for cosine similarity, and top 20 overlap before and after for *Erec*

formulas, but also ranked by importance of formal characteristics. Table 4.24 additionally shows which sequences were removed by frequency. The result is not necessarily a list of intertextualities, but if we consider this in light of Parry's work on the Homeric epic, i.e., "bards found and kept expressions which could be used in a variety of sentences, either as they stood or with slight modifications, and which occupied fixed places in the hexameter line," these genre-specific phrases were part of the repertoire for poets creating Arthurian romances.[90] Parry noted exactly which parts of the hexameter these formulas were most commonly employed in, which is exactly what the sequences show us. Apart from the closed, monosyllabic words beginning the epic meter, we see a very common 'OOC' cadence, but more importantly, this open cadence comes in the second verse of a rhyme pair (represented by the 1), matching 'swaz ir gebietet daz'. Parry happens to use the example "and X replied" (τὸν δ᾽ ἠμείβετ᾽ἔπειτα), of which we see several variations in the top matches.[91] As shown here, Parry argues that formulas are due "to the theme, their rhythm is fixed by the verse-form..."[92]

sequence	*# removed*
der künec artûs (the King Arthur)	76
dô sprach der künec (then spoke the king)	28
sî sprach lieber herre (she said dear sir)	24
als ich iu sage (as I tell you)	24
ich ie wart geborn (I ever was born)	22
lîbe und mit guote (dear and with good)	20
dô sprach der herre (then spoke the sir)	20
ich weiz wol daz (I know well that)	19
nâch âventiure reit (rode out for adventure)	18
biderbe unde guot (diligent and good)	18
unz an mînen tôt (us to my death)	16
dô sprach diu künegîn (then spoke the queen)	16
bete und mîn gebot (prayer and my order)	16

[90]Parry and Milman, *The making of Homeric verse: the collected papers of Milman Parry*, 9.
[91]Ibid., 10.
[92]Ibid., 272.

sequence	*# removed*
-1C-C-C-C-	354
-C-1C-C-C-	196
-XC-C-C-C-	195
-C-XC-C-C-	172
-C-OOC-1C-	164
C-C-C-C-C-	150
-OOC-1C-C-	119
-C-C-C-CC-	111
-C-C-C-OC-	96
-C-C-OOC-1	92

Table 4.24: Most frequenty removed sequences from top 10 similarity

hân sô wil ich (has, so I wish to)	15
den künec artûs (the King Arthur)	15
sprach diu künegîn (said the queen)	14
mit worten und mit (with words and with)	14
swaz ir gebietet daz (whatever you ordered that)	14
swaere daz er im (burdensome, that he)	14
in willekomen sîn (he be welcomed)	14
wil ich iu sagen (I wish to tell you)	14
im saelde und êre (in goodness and honor)	13
in dem lande nie (never in the land)	12
nâch rîterlîchen siten (according to knightly customs)	12
der künec artûs ze (the King Arthur to)	12
an dem lîbe und (holding dear and)	12
dô sprach der rîter (then spoke the knight)	12
einen slac daz er (a hit that he)	12
waere geschehen ob (would have happened if)	12
ich iu wil sagen (I wish to tell you)	12

Table 4.23: Most frequenty removed matches from top 10 similarity

Correlation with Lexical Methods

As alluded to above in the explanation of a *tfidf* value, the primary use of a *tfidf* matrix and cosine similarity analysis focuses on using words as terms. It is often conducted as a first step to extract features before constructing a machine learning model. In an attempt to

	unigrams	bigrams	trigrams
formal vs. lexical	.624 (.640)	.799 (.801)	.834 (.839)
top 20 overlap	21.8%	32.6%	36.2%

Table 4.25: Pearson's (and Spearman's) r for formal vs. lexical methods, and top 20 overlap

quantify the degree to which form and context is united on the level of an individual text, I calculate the correlation between the text cosine similarities generated by the formal method presented here and those generated by a lexical method; i.e., deriving cosine similarity from words (or more precisely, lemmata) theoretically yields the texts most similar in *substance*, while the method of prosodic sound patterning yields texts most similar in *form*. To this end, the cosine similarity for each text to every other text was calculated in the manner above, however, now with lemmata as the terms for the *tfidf* matrix. Similar to the formal patterning 10-grams, the cosine similarities were calculated separately for unigram, bigram, and trigram *tfidf* matrices because a formal feature sequence of length 10 likely stretches over multiple words. The results of the lemmata similarity is then compared to the results of the formal sequencing by calculating the correlation coefficient (both Spearman's and Pearson's r). The average correlation coefficient for each comparison is given in Table 4.25. It must also be remembered that there are distinct genres and subsets of text, especially on the formal level as shown above. Thus I also calculate the share of overlap for only the top 20 most similar texts for each text, as this indicates how closely the most similar texts relate in form and content. All coefficients are positive, indicating mild to strong correlation between the methods, i.e., calculating similarity between texts using prosodic sequences yields similar results to a method rooted in the lexical word, which should come as no surprise because MHG poets often composed in a form reflective of their substance. Yet as will be shown below, this is adhered to closer by some poets than others. This correlation increases when more information about content (bigram and trigram) is provided.

More interesting than the expected correlation is the correlation on the level of an individual text. Which texts exhibit the most similar and most different rank similarities between the two methods? Table 4.26 shows correlation overlap for each method, with the lexical method utilizing trigrams. In concrete terms, this table demonstrates that the most similar texts to *Eneide*, *Herzmaere*, and *Tristan und Isold*, whether measured by form or content, are very close if not the same as measured by both methods. In contrast, the most similar texts to *Der Ritterspiegel*, *Die Klage der Kunst*, and *Lohengrin* differ drastically between the two methods (none of the top 20 most similar texts measured by form are the same as the top 20 texts measured by content). In other words, this measures those texts embodying the essence of a genre (or the marriage of form and content) and those mismatching form and content. It is very enlightenting that the texts with the highest correlation between these two methods also happen to be the earliest, or what the scholarship considers, foundational texts: *Eneide*, *Tristan und Isold*, and *Gregorius*. This suggests that in fact these

early texts did formally establish a genre, which was then followed by successors intending to write into the same genre. On the opposite end, we see the expected texts identified in the grouping analysis above. The Berner Ton poems, while sharing formal characteristics with *Minnesang*, do not share thematic similarities. Wolfram's *Titurel*, as well as the *Jüngerer Titurel*, intentionally mix form and content, as does the lyrical version of *Herzog Ernst*.

text	overlap
Eneide	0.80
Herzmaere	0.80
Tristan und Isold	0.70
Gregorius	0.70
Willehalm (Wolfram)	0.70
Erec	0.65
Karl der Grosse	0.65
Diu urstende	0.65
Das Katzenauge	0.65
Tochter Syon	0.65
Der Trojanische Krieg	0.65
Aristoteles und Phyllis	0.60
Reinfried von Braunschweig	0.60
Willehalm (U.v.T.)	0.60
Walther von Klingen	0.60
Der arme Heinrich	0.60
Flore und Blanscheflur	0.60
Tristrant	0.60
Engelhard	0.60
...	...
Boppe	0.10
Eckenlied (Fassung L.)	0.10
Hugo von Mülndorf	0.05
Koninc Ermenrîkes Dôt	0.05
Burggraf von Regensburg (Minnesangs Frühling)	0.05
Herzog Ernst (Hs.D, strophig)	0.05
Bußgebet	0.05
Der Jüngere Titurel	0.05
Der Wunderer	0.05
Dietmar der Setzer	0.05
Titurel	0.05
Engeltaler Schwesternbuch	0.05
Lannzilet	0.05
Meinloh von Sevelingen (Minnesangs Frühling)	0.05

Virginal	0.05
Der von Wengen	0.05
Lohengrin	0.00
Die Klage der Kunst	0.00
Der Ritterspiegel	0.00

Table 4.26: Top 20 overlap of formal and lexical methods

The conclusion to draw from these results is that on a large scale form and content are highly correlated in MHG verse, though as seen in Table 4.26, poets varied (intentionally and not) the degree to which they conformed in uniting both. While the formal method yields a significant number of individual insights and a useful visualization, its most significant contribution is the questions it raises and insights it provides to the relationship of form and content in MHG literature. The age old question comes to mind: Does form follow content, or content follow form? And can form and content truly be separated? While there is still no definitive answer, the life cycle of MHG form through its Germanic and Romance influences can clearly be charted. Disambiguating form and content, as attempted in the previous section, reveals that for MHG form is rooted deeper in texts than specific words or phrases. This method also raises the unsettling proposition that by reducing poetic texts to a very basic sequence of syllable and rhyme properties, an algorithm plugging in words to match certain sequences could accurately generate a text within a specific genre grouping. After all, part of my argument assumes that MHG poets were trying to, consciously or subconsciously, fit themselves into a specific style of MHG composition. While poets most likely did not keep track of syllable sequences at the level analyzed in this chapter (a level Moretti notes is "hardly notice[d]"), it is highly probable, and displayed by many poets such as Konrand von Würzburg and Wolfram von Eschenbach, that they were contemplating the sound and rhythm of the poetry they had previously heard, or even read, to intentionally generate different aesthetics.[93] The courtly rhyme pairs distinguish themselves strongly from the strophic heroic epics, yet Wolfram presents us with a problem. In his *Titurel*, and to a lesser extent *Willehalm*, Wolfram intentionally reverses form and content, generating a "Verfremdungseffekt" ("defamiliarization effect") according to März in the sense of Shklovsky: "A love story about a couple belonging to the fairytale Arthurian world, which leads them to death, and does not receive the sound coloring of a pair-rhyme happy ending; a heroic myth, which denies its hero the gloriole, as Wolfram decides, is also denied the solemn heroic cadence."[94] Müller also highlights how the ending of each story does not fulfill the model. Heusler discusses how poets crossed the division of lyric and epic poetry, experimenting with form and content, citing many of the exact texts appearing at the bottom of Table 4.26:

[93]Moretti, *Distant reading*, 206–207.

[94]"Eine Liebesgeschichte eines der märchenhaften Artussphäre angehörigen Paares, die bloß in den Tod führt, erhält von Wolfram nicht die Klangfarbe des paarreimenden happy-ends; einem Heldenmythos, der seinem Helden die Gloriole derart verweigert, wie das Willehalm beschieden ist, wird auch der heroisch-pathetische Tonfall versagt." März, "Metrik, eine Wissenschaft zwischen Zählen und Schwärmen?," 327-328

> As far as unsung strophic literature goes, these followed more the free use of rhyme pairs in feet: heroic epics, Albrecht's *Titurel*, the *Lohengrin*. Others followed more the lyric style such as Ulrich's *Frauendienst* and Konrad's *Klage der Kunst*, *Tirol und Fridebrant*, *der Winsbecke*, *der Wartburgkrieg*, along with a couple smaller pieces. Wolfram's *Titurel* stanzas uniquely employ lyrical dactyls.[95]

Similarly, recalling the discussion in Chapter 1 of truth and sound, Wolfram utilizes sound through form to deceive, or defamiliarize, the listener to the genre of the text. This "Verfremdungseffekt" is measured above by the degree to which form and content are mismatched, i.e., the degree to which similarity changed after extracting genre-specfic sequential intertexualities and formulas. Florian Kragl suggests the idea of a different schemata for MHG literature, if a text only becomes a 'text' in its content being processed through a form at the same moment, form and content are tightly interwoven and cannot exist without one another (*geformte Materie*), suggesting that 'form' is better labeled 'semantic structure' (*Sinnstrukturen*).[96] Or is a text simply a combination of distinct form and content? While the preceding analysis supports these arguments, it is important to remember that these statistics are relative, and the goal of this study is not to produce raw statistics, but rather, as Zumthor writes, to reveal the 'consequences' of a literature dominated by a form that is driven by voice.[97] This analysis has certainly revealed the power of this formal characteristic.

[95] "Soweit unsangliche Buchdichtung in Strophen ging, folgte sie im Taktfüllen mehr dem freieren Brauch der Reimpaare: Heldenepos, Albrecht's *Titurel*, der *Lohengrin*. Mehr nach dem lyrischen Gleichtritt hinüber liegen Ulrich's *Frauendienst* und Konrad's *Klage der Kunst*, *Tirol und Fridebrant*, *der Winsbecke*, *der Wartburgkrieg*, nebst ein paar kleineren Stücken. Eigenartig von liedhaften Daktylen bestimmt sind Wolframs *Titurel*-strophen." Heusler, *Deutsche Versgeschichte: mit Einschluss des altenglischen und altnordischen Stabreimverses*, 102

[96] Kragl, "wort unde wîse. Formen des sangbaren Verses in der deutschen Literatur des Mittelalters," 41.

[97] Zumthor, *Die Stimme und die Poesie in der mittelalterlichen Gesellschaft*, 83.

Chapter 5

Meter and Emphasis

Poetic meter in the MHG tradition has always been a contentious and complex subject, as it requires a nuanced knowledge of not only MHG literature, but also a strong understanding of MHG linguistics, particularly phonology, and of the musical practices of the period.[1] Most work thus far has not been able to master all of these areas.[2] While this project does not attempt to fully unite these diverse fields, it does seek to take careful consideration of each in developing a computational model to better understand how medieval German poets crafted their words into meter. The increased popularity of machine learning algorithms and their application to textual data presents a particularly fruitful opportunity in a domain that has plagued MHG scholarship for years. Instead of a deductive approach, i.e., beginning with the assumption of trochaic alternation as the guiding principle, supervised learning allows for a large-scale inductive approach, supplying the algorithm with a wealth of specific examples from which general principles can be discerned.

While late 20th century scholarship neglected meter primarily due to theoretical disagreements and a lack of manuscript evidence, Christoph März recently re-framed MHG scholarship on meter in his article 'Metrik, eine Wissenschaft zwischen Zählen und Schwärmen', in which he attempts to revive a meter-based formal perspective.[3] According to Müller, form has two important functions and opportunities: it *reminds* us, recalling the discussion of

[1]This chapter has adapted much of the author's previous work in "Supervised Machine Learning for Hybrid Meter." Alex Estes and Christopher Hench, "Supervised Machine Learning for Hybrid Meter," *Proceedings of the Fifth Workshop on Computational Linguistics for Literature, NAACL-HLT 2016*, 2016, 1 Shortly after, Agirrezabal et al. undertook a similar project using the benchmark results set by Hench and Estes.Manex Agirrezabal et al., "Machine Learning for Metrical Analysis of English Poetry," *Proceedings of COLING 2016, the 26th International Conference on Computational Linguistics: Technical Papers*, 2016, 772–781 It has also been cited by work adapting the model to Spanish and Portuguese.Borja Navarro-Colorado, "A metrical scansion system for fixed-metre Spanish poetry," *Digital Scholarship in the Humanities*, 2017, fqx009; Adiel Mittmann et al., "Escansão automática de versos em português," 2016,

[2]Christoph März criticizes recent scholarship as being too linguistic in nature, and forgetting the aesthetic effect meter produces.März, "Metrik, eine Wissenschaft zwischen Zählen und Schwärmen?," 323

[3]Ibid.

memory in the first chapter, and it allows for *comparison*.[4] Both of these observations provide motivation for the following analyses. Poetic meter acts not only to support the memory of a performer or composer, but also affects the audience, prompting this comparative reception. März writes:

> I recall the experience that when you try to remember a poem, you often only remember the pattern—a few words may come along with that pattern or not. Also, if you forget parts of the text, the threads can be found again in certain passages by humming the rhythm of the verse.[5]

This act of remembering serves as an opportunity to identify connections between songs and texts (both formally and semantically), and compare texts, as März would have it. This comparison, when recognized by a performer or audience, can generate and add meaning to a song. Especially in the MHG tradition, a connection between form and content has always been presumed (and demonstrated in the previous chapter), particularly between form and genre. Yet März is also interested in lower level connections and references within genres. März asks whether these abstracted metrical schemata "transport" specific ideas, and if so, how they are created.[6] Klaus Kohrs asked a similar question decades earlier. Kohrs explains in Saussurian terms how meter itself can add signification to language, which it does not inherently carry: "With the metrical, that is even 'quasi-musical' formation of language as a symbolic and sonoric phenomenon the side of the Signifié is quasi sublimated, i.e., sensical and semantic references become virulent, which the 'natural' language does not have and does not need to have."[7] Hugo Kuhn presents the idea similarly in relation to music and melody, but emphasizes its "Gebrauchsfunktion" (use function), i.e., the use-cases for these artworks, as folksongs, religious uses, for the court, knights, etc.[8] This point is taken up by Cramer, questioning what the actual *Gebrauchsfunktion* for these artworks was, and whether our ideas of them are correct according to the sources.[9] But März crucially reshapes this question, instead of asking what meaning or function poetic meter may contain, he notes that meter is always determined relatively.[10] As Zumthor and Saussure have claimed about words and sound, there is no meaning in the base element itself, only in context and pattern.

[4] März, "Metrik, eine Wissenschaft zwischen Zählen und Schwärmen?," 325.

[5] "Ich erinnere an die Erfahrung, daß man, wenn man sich ein Gedicht wachzurufen versucht, oftmals nur dessen 'Schema' im Kopf hat - es fallen einem dann die Wörter dazu ein oder nicht. Auch daß man, wenn man Teile des Textes vergessen hat, den Faden an irgendeiner Stelle wiederfindet, indem und weil man den Versgang vor sich hinbrummt." ibid.

[6] Ibid.

[7] "Mit der metrischen, also 'quasi-musikalischen' Durchformung der Sprache als eines Zeichen- und Lautphänomens wird dann auch die Seite des Signifié quasi sublimiert, d.h. es werden Sinn- oder Bedeutungsbezüge virulent, die die 'natürliche' Sprache nicht hat und nicht zu haben braucht." Kohrs, "Zum Verhältnis von Sprache und Musik in den Liedern Neidharts von Reuental," 605

[8] Kuhn, *Text und Theorie*, 38.

[9] Ibid., 39.

[10] "Metrumbehaftetes [ist] immer auch relational bestimmt." März, "Metrik, eine Wissenschaft zwischen Zählen und Schwärmen?," 325

But for both words and poetic meter, this context must be extended beyond the contained object of a line of poetry to the body of referential objects. The aim of this chapter is to disambiguate these relative relationships. This project does not intend to argue that Andreas Heusler's metrical theory, or any other theory for that matter, is without fault, nor that specific metrical values even exist as such, but rather that implementing *any* framework inevitably teases out relative differences within a corpus.[11] Heusler decries the 19th century philologists for altering the text and making a statistical analysis of MHG metrics impossible, and for this reason he gives no statistics in his MHG study of meter.[12] Yet the focus here is not concrete, in terms of absolute numbers or statistics, but rather in establishing relative relationships between texts, which aggregated over a whole text or corpus, as shown in the second chapter, will not drown out clear characteristics.[13] März follows this argument to its end in calling these relationships 'echoes', while I title this project 'resonances' to emphasize the importance of sound.[14]

5.1 From Quantity to Quality

The distribution of Latin into distinct regional dialects had profound linguistic and literary implications for all of Europe. One notable consequence was on the quantitative poetry. Even before the Middle Ages, the syllable length of classical Latin had been nearly forgotten in the vernacular. Latin poetry had used quantitative meter, whereby syllable length was the organizing principle, and syllables could still be counted based on this length. However, the emerging dialects differed from Latin in that stress became a phonologically important feature, and so-called qualitative meter ('rhythmic poetry') predominated in the Romance languages. Reconciling these linguistic differences, MHG meter relied on both stress and syllable length. This hybrid metrical form poses unique challenges to scanning poetry and allowed for a diverse development in genre and style.[15] Yet this freedom is the cause for one of the main questions and theoretical problems in MHG research on meter: not necessarily Heusler's question of "*How* am I to measure it?"[16], but rather *what*: in a system of "measured syllable verse with free syllable counts"[17]—what is it that we can count, or

[11] This is similar to the approach taken in the previous chapter, in which intentionally *no* metrical theory was adopted, as the data was sufficient to group texts relatively simply with phonological features.

[12] Heusler, *Deutsche Versgeschichte: mit Einschluss des altenglischen und altnordischen Stabreimverses*, 4.

[13] Admittedly, idiosyncrasies in a specific text will, for this reason, be very difficult, if not impossible, to discern if edited texts are used for analysis.

[14] März, "Metrik, eine Wissenschaft zwischen Zählen und Schwärmen?," 326-327.

[15] Heusler, *Deutsche Versgeschichte: mit Einschluss des altenglischen und altnordischen Stabreimverses*, 74-75.

[16] "*Wie* hab ich zu messen?"

[17] "silbenzählende Verse mit freier Silbenzahl"

should count?[18] "What is countable in the verse?"[19] We could count syllables, but it is not clear if the poets did this naturally themselves, despite what the *Meistersänger* would like us to believe. We could also count syllable features, as in the previous chapter, and while a relative understanding of prosodic sound and rhythm may be garnered, the rhythmical and musical patterning (if a clear one existed) remains elusive. Herbert Bögl describes MHG verse in his *Abriss der mittelhochdeutschen Metrik: mit einem Übungsteil*: MHG "presents in an abstract language of symbols the sequence of syllables in a verse and weighs them taking into account their length and stress."[20] It is this "taking into account" that presents a difficult computational problem for analysis MHG meter, in that strict rules for length and stress are and cannot always be employed.

To illustrate this shift from a quantitative classical meter to a qualitative post-classical vernacular meter, I first consider the quantitative 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 two consonants or '×'.[21] All other syllables are short. The first line of Virgil's *Aeneid* serves as example:[22]

ārma vi|rumque ca|nō, Tro|jae quī|prīmus ab|ōrīs[23]
— ◡◡ | — ◡ ◡ |— — |— — | — ◡ ◡|— —

A widely cited poem displaying the shift from quantitative to qualitative rhythmic poetry in the Latin tradition is Bishop Auspicius of Toul's late 5th century letter to Arbogast, the Count of Trier, imitating the iambic dimeter already made famous by Ambrose.[24] The letter begins:

Praecelso exspectabili his Arbogasti comiti
Auspicius qui diligo salutem dico plurimam.[25]

The first hemistich — ◡ ◡ — ◡ — ◡ ◡ shows that a quantitative scansion would be ill-fitted to the rest of the verse, and that a strictly iambic scansion is preferred, with

[18]Heusler, *Deutsche Versgeschichte: mit Einschluss des altenglischen und altnordischen Stabreimverses*, 9, 13.

[19]"[W]as an Zählbarem enhält der Vers?" März, "Metrik, eine Wissenschaft zwischen Zählen und Schwärmen?," 323-324

[20]"stellt in abstrakter Zeichensprache die Abfolge der Silben in einem Vers dar und gewichtet sie *hinsichtlich* ihrer Länge und ihrer Betonung." Bögl, *Abriss der mittelhochdeutschen Metrik: mit einem Übungsteil*, 9

[21]Bruce Hayes, "Compensatory Lengthening in Moraic Phonology," *Linguistic Inquiry* 20 (1989): 253–306.

[22]— represents a long syllable and ◡ a short syllable.

[23]"I sing of arms and of a man,"

[24]Norberg and Ziolkowski, *An introduction to the study of medieval Latin versification*, 101; Ralph W. Mathisen, *People, Personal Expression, and Social Relations in Late Antiquity* (University of Michigan Press, 2003), 34

[25]"To the highest Count Abrogast waiting, whom I, Auspicius, love and say many greetings." Norberg and Ziolkowski, *An introduction to the study of medieval Latin versification*, 101

a paroxytone in the cadence. Much Latin poetry followed suit, and the medieval *Codex buranus* famously witnesses the intermingling of Latin and MHG rhythmic verse, clearly demonstrating that they were drawing from the same rhythmical schemata. Yet Germanic meter, on the other hand, did not originally follow the quantitative meter of antiquity, preferring organization through alliterative techniques. In fact, Heusler calls alliteration the "Hausmarke" (house brand) of the Germanic language family, dependent on the dynamic root syllable inherent in Germanic languages.[26] In addition to alliteration, a further marker of Germanic verse is the *Langzeile* (long line), traditionally consisting of two *Kurzzeilen* (short lines), an *Anvers* (first half of the line) and *Abvers* (second half of the line).[27] While this tradition began earlier, a classic example of Germanic alliterative verse is the 9th century Bavarian *Muspilli*:

> …sin tac piqueme, daz er touuan scal.
> uuanta sar so sih diu sela in den sind arheuit,
> enti si den lihhamun likkan lazzit,
> so quimit ein heri fona himilzungalon,
> daz andar fona pehhe: dar pagant siu umpi.
> sorgen mac diu sela, unzi diu suona arget,
> za uuederemo herie si gihalot uuerde.[28]

This alliterative verse dominated throughout most of OHG, and continued strong in the Nordic traditions. Around the same time that the *Muspilli* was written in the southeast, to the west Otfrid von Weißenburg in Alsace was beginning to incorporate characteristics of Old French poetry into his 9th century OHG verse, leaving the first evidence of Germanic language verse to break with the alliterative verse and incorporate end rhyme (referred to as a strictly Romance language influence by Heusler). Otfrid is generally considered the starting point[29] for a study of modern German verse.[30] Otfrid's *Evangelienbuch* became the model for this new Germanic verse, though he retained the *Langzeile* from the older Germanic tradition. Otfrid established many of the new metrical possibilities in cadence (monosyllabic full, bisyllabic ringing, and trisyllabic ringing) witnessed in the MHG period.[31] Much of the influence on Otfrid's style came from various writings on religion, heroic stories, and charms

[26]Heusler, *Deutsche Versgeschichte: mit Einschluss des altenglischen und altnordischen Stabreimverses*, 92-93.

[27]Ibid., 100.

[28]Wilhelm Braune and Ernst A Ebbinghaus, *Althochdeutsches Lesebuch* (Tübingen: Niemeyer, 1994) "his day comes, on which he will die. If the Soul then quickly makes it way and leaves the body lying there, then one army comes from the stars and another from hell: they fight over it (the soul). The soul may be worried until judgement is made as to which of the armies it will be brought."

[29]Though Heusler is sure to note that Otfrid was not the only OHG source in the 9th and 10th centuries.

[30]The break with alliteration was much stronger on the continent than in England, as demonstrated by Old English and Old Norse verse. Heusler, *Deutsche Versgeschichte: mit Einschluss des altenglischen und altnordischen Stabreimverses*, 8

[31]Ibid., 13.

recorded at the time.[32] Heusler argues that this freedom in verse came primarily from the church, specifically church songs. Heusler, echoing and intensifying the comments on sung lyric in Chapter 3, writes: "song more easily takes advantage of the prosodic freedom."[33] Concerning rhyme, for nearly 300 years there was only pair rhyme in the AABB form, occasionally AAA, until around 1150.[34] Otfrid's rhyme began as pure monosyllabic rhyme, and later developed into multi-syllable assonance and other types.[35] As the importance of rhyme grew, it became necessary for the rhyming syllable to also carry accent.[36] This new rhyme and accent provided an alternative means to tie verses together, but also ushered in new freedoms of measuring verse, as rhyme required syllables to relate to one another, something emphasized by the contemporary musicologists.[37] The form of the Ambrosian hymn is the closest metrically to Otfrid. The greatest difference lies within the construction of the line, where the syllable count is not certain, and divided lifts[38] are abundant.[39]

Otfrid's founding of the German rhythmic verse was what Heusler calls a "Germanicizing" of the Romance iambic verse: free filling of verses with syllables, anacrusis, and more varied cadences.[40] Despite the occasionally nationalistic statements made by Heusler, he does accurately chart out the development of Germanic verse and its influence from the Romance tradition, particularly in that the mixing of alliterative and pair-rhyme verse led to the early Germanic free filling of feet. Yet metrical conventions did exist in Otfrid's verse. The last foot was still strictly monosyllabic and verses could range from four to ten syllables, but were more often somewhere in between.[41] OHG verse often had feet with more syllables than MHG because OHG words simply had more medial syllables.[42] In contrast to MHG verse, OHG verse was more consistent with syllable length and duration.[43] In this sense, OHG verse was a "mediator" between Latin and alliterative verse.[44] With Otfrid, the German pair-rhyme *Vierheber* (four stresses) began to take shape:

Vuas	líuto	filu	in	flí	ze,[45]	
×	× ×	◡́ ◡ ×		—	◡̀	◡̂

[32]Incantations saw the greatest innovation in a move from older forms of verse to a distinctly Germanic verse.Heusler, *Deutsche Versgeschichte: mit Einschluss des altenglischen und altnordischen Stabreimverses*, 6

[33]"Gesang trägt leichter über die prosodische Freiheit hinweg." ibid., 32

[34]Ibid., 12.

[35]Ibid., 20.

[36]Ibid., 24,3 1.

[37]See Chapter 1; ibid., 9

[38]Akin to eighth notes in music

[39]Heusler, *Deutsche Versgeschichte: mit Einschluss des altenglischen und altnordischen Stabreimverses*, 35.

[40]Ibid., 36.

[41]Ibid., 43.

[42]ibid., 48; ibid., 126

[43]Ibid., 56.

[44]Ibid., 63.

[45]There were many hardworking peoples,

in | managemo | ága | lei | ze,[46]
× | ◡́ ◡ ◡̀ ◡ | x́ × | — | x̀ ◡̂

5.2 The Middle High German *Vierheber*

The most comprehensive and still referenced study of (Middle High) German meter is Andreas Heusler's three volume *Deutsche Versgeschichte*. Heusler's theory has been criticized incessantly over the years, but persists as the accepted theory for MHG meter today. März claims that as reluctant as we are to not use Heusler's theories, we use them because there is simply no better alternative.[47] While attempts have been made to supplement or critique Heusler's work, especially the existence of the fundamental 'Takt' (measure, as in music), it has proven difficult for alternative theories to escape temporal restraints. If there is no 'Takt', is there no foot, or stress alternation?[48] As März observes, many of the alternative theories do not differ significantly from Heusler's, only Franz Saran's '*Schallanalyse*' (acoustic analysis) is suggested by März as a plausible alternative to better incorporate the actual voice of the verse.[49]

What follows is a description of MHG epic meter in the Heusler tradition, with supplement from other, mostly pedagogical, resources. The following framework is then employed to construct a supervised machine learning model of scansion.[50]

The predominating pattern in all MHG verse is an alternation between stressed and unstressed syllables.[51] MHG epic verse employs trochaic tetrameter: each line has four feet, and each foot is a trochee; this is known as the *Vierheber* for the four lifts in a verse. 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?[52]
| x́ × | x́ × | x́ × | x́ × |

[46]with such great zeal,

[47]März, "Metrik, eine Wissenschaft zwischen Zählen und Schwärmen?," 318.

[48]Ibid., 319.

[49]Ibid., 321-322.

[50]While Heusler's theory is certainly debatable, it must be reiterated that the intent of this work is not necessarily the absolute, *real* meter of MHG, but once again *relative* differences throughout the corpus, which could be revealed by an array of different theories. Furthermore, Heusler does not note any serious differences in meter between regions of MHG, from Veldeke in the north, to the Austrian southeast, to the Frankish northwest, there was little variation in the general *Vierheber* (four stress) patterning.Heusler, *Deutsche Versgeschichte: mit Einschluss des altenglischen und altnordischen Stabreimverses*, 77

[51]Helmut Tervooren, *Minimalmetrik zur Arbeit mit mittelhochdeutschen Texten* (Göppingen: Kümmerle Verlag, 1997).

[52]Henry Wadsworth Longfellow, *The poems of Longfellow, including Evangeline, The song of Hiawatha, The courtship of Miles Standish, Tales of a wayside inn.* (New York: B. A. Cerf, D. S. Klopfer, 1932).

Similarly, the prototypical 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.[53] If a foot is filled by one syllable, the syllable must be phonologically heavy (containing a long vowel or ending in a consonant). If the foot is filled by three syllables, either the first two or the last two syllables are often phonologically light.[54]

It is in these atypical feet that the influence of quantitative meter, where syllable length is a key factor, becomes evident. The foot must be slightly redefined to account for this. Phonologically, syllable length is measured in morae, a unit of time such that a short syllable has one mora and a long syllable has two morae.[55,56] A foot in MHG meter is more precisely defined as having two morae, not necessarily two syllables.[57] Indeed the mora, not the syllable, has been called the fundamental unit of MHG verse,[58] although the mora functions differently in this poetic tradition than in its phonological definition. If a foot has only one syllable, the syllable must be heavy because a heavy syllable is two morae and the MHG foot requires two morae. A light syllable cannot be the only syllable in a foot, since it cannot be two morae. If a foot has three syllables, two are often light because half morae are most often light syllables (the first half mora of a pair must always be light), together forming one mora.[59] 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 light, and a double mora must be phonologically heavy. Phonological length is otherwise irrelevant and any syllable can be one mora.[60]

In addition to length, as a function of morae, syllables are also assigned stress. There are three stress values: primary, secondary, and 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.[61]

[53]Anna Domanowski et al., *Mittelhochdeutsche Metrik Online*, https://www.uni-muenster.de/MhdMetrikOnline/, Accessed on 04-02-2016, 2009.

[54]Excepted are several end syllables in divided falls such as '-er', '-el', and 'ez' (more below).ibid.

[55]Anthony Fox, *Prosodic Features and Prosodic Structure: The Phonology of Suprasegmentals* (New York: Oxford University Press, 2000).

[56]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.

[57]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, a double mora is a half note, and a half mora is an eighth note.Bögl, *Abriss der mittelhochdeutschen Metrik: mit einem Übungsteil*

[58]Tervooren, *Minimalmetrik zur Arbeit mit mittelhochdeutschen Texten*, 1.

[59]Occasionally very weakly stressed long syllables can also count as a half mora.

[60]Heusler, *Deutsche Versgeschichte: mit Einschluss des altenglischen und altnordischen Stabreimverses*, 111.

[61]The metrical distinction between different degrees of stress is rooted in phonological reality: Heinz J. Giegerich, *Metrical Phonology and Phonological Structure: German and English* (Cambridge: Cambridge University Press, 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

The final mora of the final foot of a line is omitted by convention. This is construed as a pause, or musically a rest, and receives its own symbol in the scansion ⏜, even though there is no corresponding word or syllable.[62] A short, word final syllable may also be elided before a word beginning with a vowel. Finally, MHG epic verse permits up to three syllables in anacrusis (or pickup notes, 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:

- **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[63]
- **half mora - secondary stress** (◡̀): a short syllable with secondary stress
- **half mora - unstressed** (◡): an unstressed syllable
- **double mora** (—): a stressed long syllable; double morae always carry primary stress
- **elision** (e̩): 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:[64]

Ein |	ritter |	sô ge|lêret |	was[65]	
× |	×́ × |	×́ ×|	×́ × |	×́ ⏜

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[66]	
× |	×́ × |	◡́ ◡ × |	— |	×̀	⏜

second and fourth syllables.

[62]The lyric tradition did allow for a feminine full cadence, which filled the entire last foot, this generally does not appear in the epic tradition.

[63]Tervooren, *Minimalmetrik zur Arbeit mit mittelhochdeutschen Texten*, 5.

[64]Note that this notation differs slightly from that which is used for classical verse.

[65]"There was a knight so learned"

[66]"he looked extensively,"

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

|die ein |ritter in |sîner |jugent[67]
| ×́ × |×́ ◡ ◡ |×́ × | ◡́ ◡ ◡̂

Line 8 shows an elided syllable in the second foot:

dar |an be|gunde̩ er |suo|chen[68]
× |×́ × | ×́ × | — | ×̀ ◡̂

5.3 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.[69] The second approach assumes nothing of the meter, and seeks to determine it by marking syllables and identifying patterns.[70] This project draws more on the latter. Previous scholarship has also 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, and thus supervised learning presents itself as an attractive method. After initial results of this project were published, similar studies were undertaken for English, Spanish, and Portuguese poetry with the results here as benchmark.[71]

[67]"which a knight [should have] in his youth."

[68]'in [these books] he began to search,"

[69]Charles O. Hartman, *Virtual Muse: Experiments in Computer Poetry* (Hanover, N.H.: Wesleyan University Press, 1996).

[70]M. R. Plamondon, "Virtual Verse Analysis: Analysing Patterns in Poetry," *Literary and Linguistic Computing* 21 (Supplement 1 2006): 127–141; Gareth McAleese, "Improving Scansion with Syntax: an Investigation into the Effectiveness of a Syntactic Analysis of Poetry by Computer using Phonological Scansion Theory," -, 2007, Erica Greene, Tugba Bodrumlu, and Kevin Knight, "Automatic analysis of rhythmic poetry with applications to generation and translation," in *Proceedings of the 2010 conference on empirical methods in natural language processing* (Association for Computational Linguistics, 2010), 524–533; Manex Agirrezabal et al., "ZeuScansion: a tool for scansion of English poetry," *Finite State Methods and Natural Language Processing*, 2013, 18–24; Borja Navarro, "A computational linguistic approach to Spanish Golden Age Sonnets: metrical and semantic aspects," in *Proceedings of the Fourth Workshop on Computational Linguistics for Literature* (Denver, Colorado, USA: Association for Computational Linguistics, 2015), 105–113.

[71]Estes and Hench, "Supervised Machine Learning for Hybrid Meter"; Agirrezabal et al., "Machine Learning for Metrical Analysis of English Poetry"; Navarro-Colorado, "A metrical scansion system for fixed-metre Spanish poetry"; Mittmann et al., "Escansão automática de versos em português"

Computational Metrics and Middle High German

A strictly rule-based approach to scanning MHG epic meter was undertaken by Friedrich Dimpel in 2004.[72] As Dimpel's work is the only of its kind in this field, it deserves special consideration here. As part of his dissertation and continuing work at the University of Erlangen, Dimpel developed a set of tools named *ErMaStat* (*Erlanger-Mittelalter-Statistik*), crafted specifically for MHG epic poetry.[73] Although sure to admit the shortcomings of such an approach, the opening pages of his introduction to *ErMaStat* reveal his stylometric intentions in making such a suite of tools:

> Whenever one attempts to approach literary, scholarly questions with quantitative processes, then one must assume that texts from different authors (or different periods of an author's work) demonstrate certain distinct characteristics on a phonological, morphological, lexical, and syntactical level, which allow themselves to be captured quantitatively.[74]

Dimpel's impressive list of variables include: 1) syllable, word, and line count, 2) vowel and consonant counts, 3) function words (specific parts of speech), 4) alliteration, assonance, and enjambment, 5) suffixes, 6) word frequencies, 7) prefixes, 8) common words (a finer measurement than word frequency), 9) word combinations (naïve bigrams), and 10) a metrical analysis. His intention, as revealed above, is to model style, or characteristics of style, in order to compare texts and estimate probabilities of works being written by the same author.

Dimpel continues with three examples. In the first example, he takes four of the better known MHG epics: *Parzival*, *Tristan*, *Wigalois*, and *Willehalm*. Using the variables above, he calculates and averages significance values, showing that *Parzival* and *Willehalm*, both written by Wolfram von Eschenbach, do in fact have a lower degree of difference than the other works to one another from a quantitative, stylistic point of view. Dimpel is also able to determine the contributions from individual variables. Dimpel's second analysis concerns the grouping of Wolfram's *Parzival* into chapters and the thesis proposed by Elisabeth Karg-Gasterstädt of four different sound types, following the work of Eduard Sievers.[75] Dimpel's *ErMaStat* supports Karg-Gasterstädt's hypothesis as a possibility. His last example considers the date of authorship of Hartmann von Aue's *Iwein* with respect to Hartmann's *Erec*.

[72] Friedrich Dimpel, *Computergestützte textstatistische Untersuchungen an mittelhochdeutschen Texten - Bd. II. Daten und Programme* (Tübingen: A. Francke, 2004).

[73] Friedrich Dimpel, "Textstatistische Analysen an mittelhochdeutschen Texten," *Jahrbuch für Computerphilologie* 6 (2004).

[74] "Wann immer man versucht, sich literaturwissenschaftlichen Fragestellungen mit quantifizierenden Verfahren zu nähern, so muss man dabei voraussetzen, dass Texte von unterschiedlichen Autoren (oder unterschiedlicher Schaffensperioden eines Autors) bestimmte unterscheidende Merkmale auf phonologischer, morphologischer, lexikalischer und syntaktischer Ebene aufweisen, die sich quantitativ erfassen lassen." ibid.

[75] Elisabeth Karg-Gasterstädt, *Zur entstehungsgeschichte des Parzival*, Sächsische forschungsinstitute in Leipzig. Forschungsinstitut für neuere philologie. I. Altgermanistische abteilung ... [Veröffentlichungen] 2 (Halle (Saale): M. Niemeyer, 1925).

Dimpel's approach to MHG meter specifically is also admirable, first programming for alternation and then hierarchically creating rules to correct alternation for stress.[76] Though his work must be commended for its accuracy and linguistic engagement, it is a laborious task, inflexible, and extremely language specific. My intention here is not to duplicate his work, nor dismiss it. Supervised learning is a new approach to an old problem for MHG. It also provides an opportunity for the "drei-stufige" (three-level) scansion Dimpel has not yet attempted, but notes is a great challenge to modeling MHG meter. There are also advantages of particular interest to humanists. A supervised method will learn to scan more as a human than a strictly rule-based approach, perhaps remaining truer to the poetic tradition, and giving insight into what poses difficulties for human scanners. It also allows for greater versatility, and a chance to analyze the prosody beyond the epic meter, and perhaps even prose.[77]

A second main difference to Dimpel's work is the intent with such a system. Dimpel has continued to pursue very compelling investigative work in author identification, stylometry, and manuscript traditions. While the proposed system below can certainly carry out similar work, this project intends here to highlight the varying use of language chronologically and their contribution to genre.[78] Thus this project aims to foreground particular story traditions, and how poetic meter plays a central role in declaring their uniqueness, once more investigating the relationship between MHG form and content.

Supervised Learning

The machine learning approach taken in this project is supervised, i.e., the computer is provided with annotated data in the form:

ein/MORA WBY/WBY rit/MORA_HAUPT ter/MORA WBY/WBY sô/MORA_HAUPT WBY/WBY ge/MORA lê/MORA_HAUPT ret/MORA WBY/WBY was/MORA_HAUPT

The algorithm then learns which of the annotated features (described below) are important, and subsequently how to classify any given syllable. In contrast to other automated scansion systems, a supervised approach learns how the human annotators scanned based on a set of provided features and annotated data, as the algorithm identifies which features were deemed important by the humans who annotate them. When working with human productions, such as poetry, this is an attractive advantage. Yet there are both advantages and disadvantages to this method. On the one hand, the resulting model will take contextual and situational factors into account, factors that a strictly rule-based approach may not, due to the multiple layers of rules and probabilities constructed. If a poet attempts a certain stylistic move during a section of the narrative it may be captured (via a combination of feature weights) by the model (if the human annotators noticed it in a similar passage). On the other hand,

[76] As a deductive approach, it bears similarities to the algorithmic syllabification in Chapter 2.

[77] **dimpel_textstatistische_2004; dimpel_automatische_2015**.

[78] As medievalists, we are well aware of the problems of authorship: See Bumke, "Der unfeste Text"

	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 5.1: Summary statistics for annotated dataset

this leads the model to scan poetry more like the annotators, which naturally narrows the limits of interpretation and follows the practices of a specific theoretical school. Nevertheless, as mentioned in Chapter 1, this final chapter seeks to push the limits of modeling MHG texts, aiming to reveal new aspects of these texts at the cost of making such assumptions.

5.4 Data

Because supervised learning is a novel approach to poetic meter, annotated metrical data do not exist for MHG or most other languages. Following the Heusler scansion tradition system outlined above, syllables of MHG epic poetry were annotated 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*.[79] 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 reproduced in Table 5.1.

Syllabification was performed prior to annotation according to the system detailed in Chapter 2.[80] Annotation was carried out by the author and colleague Alex Estes, who are both trained in MHG scansion.[81] In the case that a line exhibits multiple permissible scansions, priority is given to the scansion that 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 coefficient for the inter-annotator agreement is .962 (confusion matrix given in Table 5.2). The greatest disagreement for the

[79]Incorporating different poems from different poets accommodates varying styles of writing, but it also introduces more variability, which will become the foundation of the model.

[80]Hench, "Phonological Soundscapes in Medieval Poetry."

[81]Although neither author is a native speaker of NHG, 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 1 \ Annotator 2	×	×́	×̀	—	◡	◡́	ẹ	◡̀
×	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
ẹ	0	0	0	0	0	0	9	0
◡̀	0	0	0	0	0	0	0	1

Table 5.2: Inter-annotator agreement confusion matrix

human annotators was among unstressed and stressed morae, and between unstressed morae and unstressed half morae, implying both some stress and some value disagreement.

5.5 Workflow

I present a new workflow for the automated scansion of poetic meter (MHG and other). The process begins with the syllabification of texts, as described in Chapter 2. Because the syllable is the base unit for many poetic traditions, it is also what needs to be annotated. After syllabification is the metrical annotation, requiring experts in scansion and the texts themselves.[82] After annotation, features must be developed and extracted to help the model understand what part of a verse is important in assigning rhythmical values. Many of the most important features are phonological. After these features are identified, an algorithm must be developed for obtaining and annotating these features before they, along with the syllables themselves, are sent to the model. After feature identification and extraction, there is a process of development and validating the model, in order to choose the most suitable parameters for the task. The model then makes predictions for each syllable based on the features and parameters supplied. The model also yields marginal probabilities for each metrical value for any given syllable considering the other syllables in the line, i.e., each syllable comes with a list of probabilities for each of the possible metrical values. Simply taking the most probable sequence for any given line can be quite accurate, and considering that there is not yet a constraint to the four stresses in MHG epic meter, would be very helpful in eventually constructing a model for MHG lyrical poetry, which does not adhere strictly to the *Vierheber* qualities. In a sense, the bare model is a model best suited to predicting stress.

[82]For the presented model, the accuracy will be highest if the text is standardized and includes markers of long vowels because the annotated texts were such, and the extracted features depend upon this. However, the model can still scan any sort of text input with a reduced accuracy.

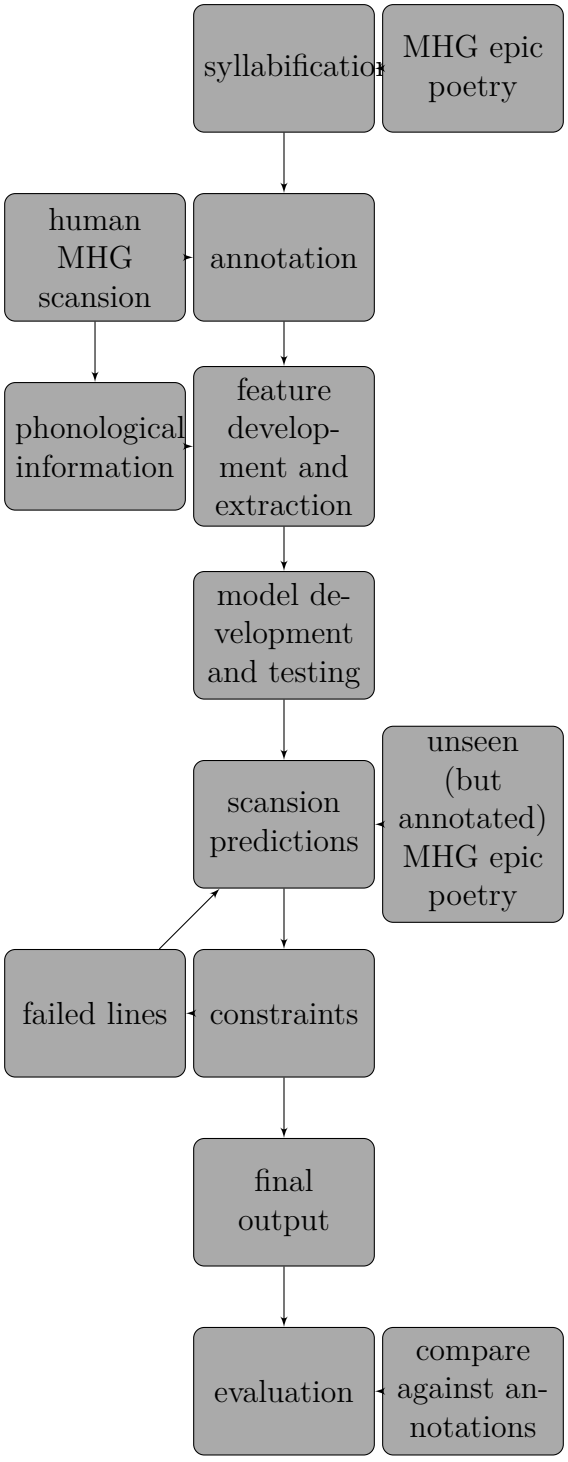

Figure 5.1: Flow chart for MHG scansion.

For the epic poetry, and the model described here, the predictions must be further processed through specific constraints, as the model will not impose strict rules unless instructed to do so. If the predicted sequence passes each constraint, it is considered the final scansion. If not, the line is sent for further processing. The following tests check for impossible scansions in MHG epic meter:

- **Four stresses**. The heart of the *Vierheber* is exactly four stresses per line.[83] Any more or less fails the line.
- **Double morae must be heavy**. Phonologically, to carry the weight of two morae, the syllable *must be heavy*. If a light double mora is identified, the line fails.
- **First syllables of divided lifts must be light**. While divided falls are allowed to be either light or heavy depending on the end syllable (though usually light), the first syllable of a divided lift *must be light*, if not, the line fails.[84]
- **Elided syllables must be light**. An elided syllable must end in a short vowel.
- **Alternation**. Two lifts cannot follow one another unless a lift follows a double mora, otherwise stress alternation is not upheld and the line fails.

If the line fails any of the above tests, it is sent for further processing, where, depending on the phonology of the syllables and the rules above, all phonologically possible metrical values as well as their marginal probabilities for each syllable are combined into every possible sequence for all the syllables in the line. Illegal combinations according to the tests above are ruled out of the set, and the set is then ranked by the sum of the probabilities of the metrical values for each syllable. The most probable, legal scansion is then selected. Thus errors in the model stem primarily from incorrect identification of the language's natural stress, as all certain metrical patterning errors are sorted out. The model itself is intended to account for this natural stress by taking advantage of the information provided to it through the annotated data. The constraints help the model further cut out possibilities that conform to the natural stress, but not the metrical environment (which is relatively weakly learned by the model), yielding an ultimately highly accurate model. After the model is constructed, it is possible to predict the rhythmical patterns of unseen poetry (although to evaluate the accuracy, the values must be known and annotated).

[83]Generally, this rules out the *stumpf* (blunt) cadence, which carries only three stresses. The *stumpf* cadence is rarely the only possible scansion (often a double mora can be assigned to fill the feet), though as Bögl points out, some lines in *Erec*, for example, leave the *stumpf* cadence as the only possible scansion.Bögl, *Abriss der mittelhochdeutschen Metrik: mit einem Übungsteil*, 26

[84]Domanowski et al., *Mittelhochdeutsche Metrik Online*.

5.6 Models and Features

Two baseline models were developed[85]: 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, though not explicitly accepting features beyond the syllables themselves. The n-gram model consists of cascading trigram, bigram, unigram, and regular expressions models, i.e., first a value is predicted based on the previous two values, if possible; otherwise it is predicted based on the previous one value, and if the first two models fail, it is predicted solely based on the value 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.[86] The n-gram model was implemented with default settings.

The Brill model[87], 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. This approach is very similar to Dimpel's enumeration of hierarchical rules for MHG scansion.

To compare to the baseline, efforts were focused on constructing a Conditional Random Fields (CRF) model.[88] The decision to implement a CRF model was predicated on the interpretability of CRF modeling and understanding the primary features for MHG scansion. A CRF model fits the problem of scansion better than a traditional Hidden Markov Model (HMM) because HMM models only consider relationships between each state and the observation (i.e., the previous syllable and metrical value), and the HMM utilizes the joint probability distribution, while the CRF model utilizes the conditional distribution and can incorporate information from not only each state, but the entire observation, or in this case line (i.e., line length, cadence features, etc.).[89] In the model, each syllable contains the

[85]The results for all models were internally 10-fold cross-validated.

[86]This proved important to recognize stress alternation.

[87]Eric Brill, "Transformation-based error-driven learning and natural language processing: A case study in part-of-speech tagging," *Computational linguistics* 21, no. 4 (1995): 543–565; implemented with the help of NLTKSteven Bird, Edward Loper, and Ewan Klein, *Natural Language Processing with Python* (O'Reilly Media Inc., 2009)

[88]John D. Lafferty, Andrew McCallum, and Fernando C. N. Pereira, "Conditional Random Fields: Probabilistic Models for Segmenting and Labeling Sequence Data," in *Proceedings of the Eighteenth International Conference on Machine Learning*, ICML '01 (Morgan Kaufmann Publishers Inc., 2001), 282–289; The implementation of the CRF model was expedited with the help of *crfsuite*.Naoaki Okazaki, *CRFsuite: a fast implementation of Conditional Random Fields (CRFs). http://www.chokkan.org/software/crfsuite/*, 2007, `http://www.chokkan.org/software/crfsuite/`

[89]Future work might consider an alternative in sacrificing interpretability for accuracy utilizing a Bidirectional Long Short Term Memory (BLSTM) neural network, though this is not attempted in this project because the CRF model proves very accurate when considering Cohen's Kappa.

features for the syllable itself, but also those of every syllable in the line, marked by index. The features and their motivations are:

- **Position within line**: the last mora of a line is always stressed (except in masculine bisyllabic cadences), and double morae occur most often in the third foot. If there is anacrusis, these syllables will be unstressed morae.
- **Length of syllable in characters**: longer syllables (in characters and phonemes) 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 two characters, last character, and last two characters.
- **Elision**: the last two characters of the previous syllable and the first two characters of the current syllable are identified as one feature to detect conditions for elision.
- **Syllable weight and length**: syllables ending in a vowel or consonant are open or 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**: MHG is a stress initial language.

The model was tuned only on the cross-validated 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 to the model itself were made after the model features and parameters were selected. However, the additional rules described above were added in order to increase accuracy for the epic meter specifically.

5.7 Results

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 5.2 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. The final results of the CRF model are given in Table 5.3 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 preferred CRF model achieves an F-score of .924 on the cross-validated development data and .909 on the held-out testing**

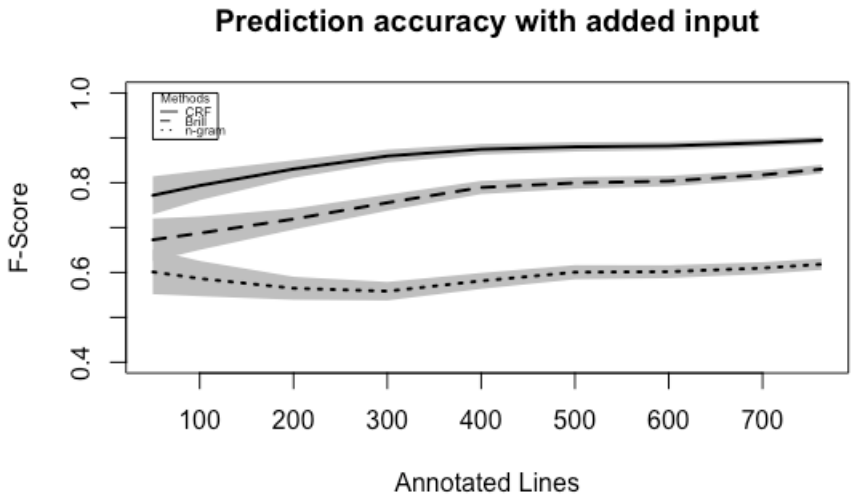

Figure 5.2: Tagging accuracy with added input.

data. Supervised learning thus proves to also be an economical option for languages with complex meter.

The top ten highest scoring features of the CRF model and rules of the Brill model are given in Table 5.4. It is evident that the CRF model takes advantage of the phonological features provided, an advantage over the baseline models. Top CRF features (1) and (5) suggest any heavy syllable is likely stressed, and often a stressed mora. The CRF model also discerned cadence from the line patterning, exhibited in top CRF features (2) and (4), noting that except for the rather uncommon occurrence of a cadence with a divided lift in the last foot (masculine bisyllabic cadence), the last syllable is always a stressed mora. Elision appears frequently in the top CRF features (3) and (6). Anacrusis is recognized in top CRF features (7) following the prototypical patterning:

Ein | ritter | sô ge|lêret | was[90]
× | ×́ × | ×́ ×|×́ × | ×́ ⏜

Where there are eight syllables, if seven syllables down the line is the last syllable of the line (EOL), and alternation is regular, that focal syllable will be unstressed in the anacrusis. Top CRF features (8), (9), and (10) each consider words and word boundaries, specifically that double morae often occur at the beginning of a multi-syllabic word, and often that word is bisyllabic, with the second syllable ending in 'en', such as 'mae-ren' or 'rî-ten'. Top CRF feature (10) notes that unstressed half morae often occur after the first syllable of multi-syllabic words (half morae are only stressed when beginning a word).

[90] "There was a knight so learned" Hartmann and Mertens, *Der arme Heinrich*, l. 1

			held-out	
metrical value	**F**	**obs.**	**F**	**obs.**
mora - unstr.	.938	2405	.937	253
mora - prim.*	.949	2463	.951	253
double mora	.881	424	.928	34
half mora - unstr.	.672	231	.541	41
half mora - prim.	.822	107	.667	11
elision	.773	65	.667	2
(macro) average	.924		.909	

*Morae and half morae with secondary stress were not predicted, rather determined based on wordboundaries after prediction.

Table 5.3: CRF model F-score for individual metrical values and (macro) average in development and on held-out data

CRF	**Brill**
(1) — if heavy syll.	(1) $\acute{\times}$ → × if at word boundary and following syll. is $\acute{\times}$
(2) not $\acute{\times}$ if next syll. is EOL	(2) $\acute{\times}$ → — if followed by $\grave{\times}$ and word boundary
(3) ẹ if first char. of next syll. is 'e'	(3) × → $\grave{\times}$ if end of line
(4) $\acute{\times}$ if EOL	(4) $\grave{\times}$ → $\acute{\times}$ if monosyllabic
(5) $\acute{\times}$ if heavy syll.	(5) × → $\acute{\times}$ if following syll. is 'ge'
(6) ẹ if first char. of next syll. is 'i'	(6) $\acute{\times}$ → × if 'ist' +2 syll.
(7) × if +7 syll. not EOL	(7) ẹ → ⌣ if 'der', 'den', 'diu' +2 syll.
(8) — if next syll. same word	(8) ⌣ → ẹ if 'ein', 'ich', 'er' +2 syll.
(9) — if next syll. ends in 'en'	(9) ẹ → $\grave{\times}$ if 'ge' +2 syll.
(10) not ⌣ if beg. of word	(10) — → × if 'al' -1 syll.

Table 5.4: Top ten CRF features and Brill rules

		Predicted					
		×	×́	⏑	—	⏑́	ę
True	×	239	4	8	2	0	0
	×́	6	243	3	0	1	0
	⏑	11	7	20	1	0	2
	—	0	1	1	32	0	0
	⏑́	1	3	1	0	6	0
	ę	0	0	0	0	0	2

Table 5.5: CRF confusion matrix

The Brill model adopts a more general rule for alternation in top Brill feature (1). Notably, the Brill model takes greater advantage of word boundaries in (1) and (2), while these features rank lower in the CRF model. The Brill model inevitably also notes the influence of specific words or prefixes. The unstressed past participle prefix 'ge' ranks as a top five rule for the Brill model. If 'ist' is two syllables down the line, the current syllable's assignment is changed from stressed to unstressed, or if 'ein', 'ich', or 'er' is two syllables away, the original half mora assignment is changed to an elision.

The scores from both models confirm extant MHG metrical theory (as it was employed for the annotation), but suggest new methods of approach for students of MHG meter. Instead of first marking stress, as suggested by *Minimalmetrik* Tervooren, *Minimalmetrik zur Arbeit mit mittelhochdeutschen Texten* and the pedagogically oriented website *Mittelhochdeutsche Metrik Online*,[91] it may be useful to first determine the cadence and anacrusis by counting the number of syllables in the line, and looking for heavy syllables at the end of the line. Stress can then be marked in the remaining syllables, and metrical values can be assigned based on phonological features. This method may be particularly useful for non-native German speakers, who may have less feeling for the natural stress of German. These results and insights support our feature decisions and our implementation of a CRF model.

5.8 Errors and Challenges

Investigating the errors and challenges of a supervised model presents the opportunity for the greatest new insights into the field, and the advantage over an unsupervised approach. The confusion matrix for the CRF model in Table 5.5 shows the errors made in the prediction of the held-out data. The confusion matrix highlights the greatest trouble of the model in predicting both stressed and unstressed half morae, particularly the latter. This situation is mirrored in the human inter-annotator agreement matrix, demonstrating, as may be expected, that the machine learning model makes errors as the human annotators might. The unstressed mora and half mora confusion, common in both human and machine annotation, is understandable, as these two are the most phonologically ambiguous metrical values

[91] Domanowski et al., *Mittelhochdeutsche Metrik Online*.

in MHG meter. Double morae, stressed half morae, and elisions all have the phonological restrictions listed above, and stressed morae are evidently less confused with double morae, likely due to clear alternation in the surrounding environment. Unstressed morae and half morae have hardly any restrictions (only that they are likely not heavy syllables). This then generates further stress confusion between unstressed morae and stressed morae seen in both the computer model and human annotation.

If an algorithm can be trained to scan MHG meter similar to how a human might, it may be interesting to see what is considered difficult for the algorithm. This can be computed by taking the average of each syllable's maximum marginal probability over the syllables in a line. In this sense, each syllable has a marginal probability for each possible metrical value.[92]. The lower the average of the maximum marginal probability for a line, the less confident the model is about its provided scansion, and vice versa.[93] With the model, this can be computed for any text (annotated or not), but let us first look at the model text used extensively for annotation and instruction, Hartmann's *Der arme Heinrich*. Unsurprisingly, the easiest lines for the model to scan are lines that hold true to the trochaic tetrameter patterning, inclusive of the common one syllable anacrusis:

(1) sus | trouc ouch | mich mîn | tumber | wân[94]
× | ×́ × | ×́ × | ×́ × | ×́ ⌢

(2) ge | frumten | sô ge| sundez | hin[95]
× | ×́ × | ×́ ×|×́ × | ×́ ⌢

(3) Ein | ritter | sô ge|lêret | was[96]
× | ×́ × | ×́ ×|×́ × | ×́ ⌢

[92] $marginal(t, s)$ is the marginal probability of tag t for syllable s, and thus $\sum_{t'} marginal(t', s) = 1$. For a line with N syllables, the average of each syllable's maximum marginal probability over the syllables in a line is thus $\frac{1}{N} \sum_{i=1}^{N} \arg\max_t marginal(t, w_i)$

[93] While this chapter will consider the confidence of the model as the *difficulty* of scanning a line of MHG poetry, this measure could also be used to distinguish between a metrical line of poetry and a line of prose. This approach is very similar to Anttila and Heuser's study of English and Finnish.Anttila and Heuser, "Phonological and metrical variation across genres" Such a study using the model presented here for MHG would undoubtedly be a fruitful application.

[94] "Thus I was also deceived by my foolish belief," l. 400. Average of the maximum marginal probability over all the syllables in the line: .9994

[95] "beneficially so healthy (their child) in (to death)"l. 1034. Average of the maximum marginal probability over all the syllables in the line: .9991

[96] "There was a knight so learned," l. 1. Average of the maximum marginal probability over all the syllables in the line: .9991

(4) ich | weiz wol | daz er | selbe | giht[97]
× | ×́ × | ×́ ×| ×́ × | ×́ ⌢

(5) er | ist ein | vil ver|schaffen | gouch[98]
× | ×́ × | ×́ ×| ×́ × | ×́ ⌢

In these examples we also see very distinct prosodic stress patterning. 'sus', 'ouch', 'mîn', 'ber', 'ge', 'ten', 'ge', 'dez', 'ein', 'ter', 'ge', 'ret' are all common unstressed MHG syllabes both in poetry and prose, while 'trouc', 'mich', 'tum', 'wân', 'frum', 'sô', 'sun', 'hin', 'rit', 'sô', 'lêr', 'was' are all either semantically significant, or heavy, stressed syllables. Moreover, the stress of all multi-syllabic is clear: 'gefrumten', 'gesundez', 'gelêret', and 'verschaffen' have unstressed prefixes, while the rest follow the typical word-initial stress. Any MHG scholar would notice the clear trochaic quality of these lines, and few would disagree with the typical trochaic scansion. One must also wonder how the poets understood and wrote these lines. Were they particularly easy to craft? Do they carry less importance in the story? Or, as noted in the discussion of truth and lies in the first chapter, are these lines particularly *true* (and just not informative?), because the poet did not need extra effort to fit the truth into a legal line of MHG poetry?

The computer model has difficulties with foreign words, outlier line lengths, and uncommon prosodic and metrical patternings:

(6) cordis | specu | la |tor[99] (preferred)
×́ × | ×́ ×| — | ×̀ ⌢

cor|dis spe|cu la|tor (computer model)
— | ×̀ ×| ×́ ×| ×̀ ⌢

(7) dâ | hiez sî | ûf | gân[100]
— | ×́ ×| — | ×́ ⌢

[97] "I know well that he himself confirms,"l. 1162. Average of the maximum marginal probability over all the syllables in the line: .9988

[98] "He is a very much nonsensical fool," l. 725. Average of the maximum marginal probability over all the syllables in the line: .9987

[99] "the one who seeks the heart," l. 1357. Average of the maximum marginal probability over all the syllables in the line: .247

[100] "She was ordered to go on top (of the table)." l. 1206. Average m arginal probability: .274

syllable	value	max. marginal prob.
cor	—	.124
dis	×̀	.132
spe	×	.134
cu	×́	.204
la	×	.342
tor	×̀	.549

Table 5.6: Example (6), average probability .247

(8) nû | râtet | mir | alle durch | got[101] (preferred)
× | ×́ × | — | ×́ ⏑ ⏑ | ×́ ⏜
nû | râtet | mir al | le durch | got (model)
× | ×́ × | ×́ ×| ×́ × | ×́ ⏜

(9) ich en| kun| de ze | saler | ne[102]
× × | — | ×̀ × | ×́ × | ×̀ ⏜

(10) dem ist | ouch | niht ze | wol[103]
×́ × | — | ×́ × | ×́ ⏜

(11) einen | fremeden | tôt | niht ver| tragen[104] (preferred)
× × | ×́ ⏑ ⏑ | — | ×́ × | ⏑́ ⏑ ⏜
einen | freme| den tôt | niht ver| tragen (computer model)
× × | ×́ ×| ×́ × | ×́ × | ⏑́ ⏑ ⏜

While the Latin in example 6 does not pose any problem for the medievalist scanner, the model cannot discern the long syllable in specu*lā*tor, and having learned MHG, it would never

[101] "Now give me advice you all through God," l. 1482. Average of the maximum marginal probability over all the syllables in the line: .496

[102] "I could not to Salerne." l. 1018. Average of the maximum marginal probability over all the syllables in the line: .504

[103] "He is not doing so well." l. 600. Average of the maximum marginal probability over all the syllables in the line: .505

[104] "to bear another death?" l. 1329. Average of the maximum marginal probability over all the syllables in the line: .516

syllable	value	max. marginal prob.
dâ	—	.028
hiez	×́	.114
sî	×	.114
ûf	—	.114
gân	×́	.998

Table 5.7: Example (7), average probability .274

syllable	value	max. marginal prob.
nû	×	.921
râ	×́	.644
tet	×	.381
mir	×́	.151
al	×	.137
le	×́	.310
durch	×	.427
got	×́	.999

Table 5.8: Example (8), average probability .496

syllable	value	max. marginal prob.
ich	×	.768
en	×	.631
kun	—	.379
de	×̀	.348
ze	×	.263
sa	×́	.473
ler	×	.505
ne	×́	.668

Table 5.9: Example (9), average probability .504

syllable	value	max. marginal prob.
dem	$\acute{\times}$	.014
ist	$\times$	.014
ouch	—	.013
niht	$\acute{\times}$	.997
ze	$\times$	.997
wol	$\acute{\times}$	.997

Table 5.10: Example (10), average probability .505

syllable	value	max. marginal prob.
ei	$\times$	.791
nen	$\times$	.956
fre	$\acute{\times}$	.309
me	$\times$	.166
den	$\grave{\times}$	.167
tôt	$\times$	.233
niht	$\acute{\times}$	.411
ver	$\times$	.420
tra	$\acute{\smile}$	.712
gen	$\smile$	.996

Table 5.11: Example (11), average probability .516

presume a stressed syllable on the third syllable of a multi-syllabic word.[105] The maximum marginal probabilities in Table 5.6 are striking compared even to the other difficult lines, as the model is unsure about nearly every syllable.[106] The other troubling cases are more relatable. In example 7, we are confronted with the minimum number of syllables that a MHG epic verse is permitted to contain, one which the model nearly guesses correctly. Adding to the difficulty is that each word is monosyllabic, and as documented above, the model prefers bisyllabic words as double morae. The double morae are also only two characters (or phonemes) in length, an adverb and a preposition. Table 5.7 shows that the model is not confident about any of the assignments except for the very last syllable, which, as a monosyllabic ultima, is likely a stressed mora.

Example 8 appears as classic trochaic tetrameter following the form of an easy prediction for the model, yet the natural stress of the bisyllabic 'alle' prevents this scansion, forcing 'mir' to be in a stressed position in the preferred scansion. The model is least confident about the three syllables it in fact scans incorrectly, ('mir al-le') but evidently believes the

[105]However, were the long $\bar{a}$ marked it may have scanned the line correctly.

[106]In fact, using these probabilities, it is likely a promising task to be able to determine whether a word is MHG or not.

alternation to be the stronger choice than a midline, monosyllabic double mora followed by a divided lift, which itself is a rare occurence.

While the model correctly predicts example 9, it has little confidence in the middle of the line. A two syllable anacrusis is not uncommon, though not frequent, and the two light, open syllables 'de' as a stressed syllable and 'ne' as a final stressed syllable further add to confusion, although to give the model the benefit of the doubt, 'salerne' is a proper noun.

The model also correctly predicts example 10, though again MHG prosody would not suggest 'dem' or 'ouch' as double morae. The duration of the line is quite typical, and the maximum marginal probabilities confirm this. It is particularly striking in this example that the model is very uncertain about the beginning of the line. While 'dem' is clearly a rare double mora, in this position, it would even be considered a rare stress, and would rather be scanned as part of the anacrusis. In contrast to the front of the line, the model has great confidence in the end of the line, where 'ze' would commonly be unstressed (often even elided or in a divided lift), and the heavy syllables 'niht' and 'wol' are assigned the remaining stresses.

Example 11 shows another outlier in line length (10 syllables). Whether or not we accept 'fremeden' as three syllables or two ('fremden' is attributed in manuscript A) the line contains a two syllable anacrusis (notably only one word), ending in a three syllable word, the first syllable of which is unstressed. Once again, the syllables, about which the model is least certain, are also those incorrectly scanned. Yet the model evidently believes a double mora for 'tôt' and a divided lift are less likely than retaining alternation. The masculine bisyllabic cadence, particularly difficult to scan on first read, is correctly identified.

The marginal probabilities of each metrical value for each syllable in a line additionally allow us to calculate the difficulty of scanning any given MHG *Vierheber* text as a whole (and even specific sections of any text). While this is strictly the difficulty for the model, as shown above, the model is a close approximate for a human annotator, experiencing similar difficulties. Moreover, the examples illustrate that the typical trochaic tetrameter causes no problem for the model nor the human scanner, while unexpected double morae, and longer anacrusis are cause to stop and think, particularly when syllables with relatively few phonemes, or monosyllabic words, are properly assigned double mora value. To sort the *Vierheber* texts in the MHDBDB corpus by difficulty of the meter, the median and mean of the maximum marginal probabilities over all the syllables in a line for all lines in each text are calculated, yielding Table 5.12. We see the top of the chart populated by the top and bottom of the *Vierheber* cluster presented in the previous chapter, as these texts have very consistent and restricted cadences, as noted by Heusler. While Konrad von Würzburg and the anonymous author of Reinfried von Braunschweig utilize the ringing cadence frequently, Ulrich von Liechtenstein does not, yet both do so consistently and in a predictable manner. The major use of simple trochaic tetrameter (measured more finely below), or double morae only in the penultimate foot, likely further pushed texts toward the top of the list (evidenced by the model's ease in predicting strictly trochaic verse). An odd mix of texts appear at the bottom of Table 5.12. Both Ulrich von Türlin's *Willehalm*, or *Arabel*, and Wolfram's *Willehalm* evidently employ more difficult metrical schemata, as does Der Stricker's *Daniel*

von dem blühenden Tal, perhaps a consequence of his unique *Strickerkadenz*. Particularly interesting is the difficulty of scanning *Der Welsche Gast*, which, as will be demonstrated below, follows the style of Ulrich von Liechtenstein quite closely in cadence selection, which would lead one to assume an easier scansion task. Because Thomasîn was not writing in his native language, his native language being Italian, perhaps he more often confuses stress in MHG, and thus creates difficulties for the model despite a very simple trochaic patterning.

text	median	mean
Reinfried von Braunschweig	0.984907	0.952727
Der Schwanritter	0.984840	0.954346
Das Turnier von Nantes	0.984484	0.953066
Alexius	0.983046	0.953657
Herzmaere	0.982992	0.956888
Heinrich von Kempten	0.982073	0.948779
Pantaleon	0.981337	0.944830
Der Trojanische Krieg	0.981231	0.945008
Engelhard	0.980778	0.947011
Silvester	0.980702	0.948039
Der guote Gêrhart	0.978524	0.939438
Frauendienst (Bechst.) (Epik, Bechstein)	0.977328	0.934517
Herzog Ernst (Hs.D, strophig)	0.976312	0.926170
Alexander (R. v. E.) (Rudolf von Ems)	0.974599	0.930398
Barlaam und Josaphat	0.969378	0.926035
Frauendienst (Büech.) (Büechlîn, Bechstein)	0.966405	0.923810
Meleranz	0.965263	0.915054
Wigalois, der Ritter mit dem Rade	0.963911	0.919376
Gauriel von Muntabel	0.962624	0.911146
Biterolf und Dietleib	0.962019	0.917653
Helmbrecht	0.961616	0.917587
Tristan (H.v.F.)	0.960467	0.913434
Tandareis und Flordibel	0.958841	0.912529
Der arme Heinrich	0.956336	0.911926
Parzival	0.950174	0.911735
Dietrich und Wenezlan	0.948923	0.907889
Gregorius	0.948497	0.904120
Iwein	0.947784	0.902092
Lanzelet	0.947157	0.899856
Herzog Ernst (Hs. B)	0.946568	0.903171
Alexander (U.v.E) (Ulrich von Eschenbach)	0.945071	0.903506
Tristan (Ulrich v. Türheim)	0.944975	0.899818
Walberan	0.943629	0.897919

Laurin	0.943434	0.900584
Dietrichs Flucht	0.942704	0.898055
Willehalm (Wolfram)	0.942339	0.904198
Flore und Blanscheflur	0.942235	0.899201
Eneide	0.936819	0.889650
Karl der Grosse	0.935828	0.898457
Der Welsche Gast	0.934193	0.898327
Der Schlegel	0.933589	0.899505
Alexander Anhang	0.932088	0.887358
Erec	0.931306	0.890927
Daniel von dem blühenden Tal	0.922973	0.884164
Willehalm (U.v.T.)	0.918577	0.893880
Lambrechts Alexander (Strassburger Hs.)	0.894552	0.864145

Table 5.12: Model ease of scanning *Vierheber* texts based on median and mean maximum marginal probabilities over all syllables in a line for all line in the text

5.9 Characterizing the Middle High German *Vierheber*

Previous scholarship and forewords to edited editions enjoy characterizing the rhythm of epic poetry, particularly as it pertains to specific texts or poets. For example, in his introduction to MHG meter (a pedagogical text), Herbert Bögl writes:

> Hartmann and Gottfried prefer a leaner verse with fewer syllables, Wolfram, however, is more familiar with extensive anacrusis and more syllables in a foot. The following generation of poets avoided if possible feet with many syllables as well as monosyllabic feet, and prefer a monosyllabic anacrusis in contrast to the multisyllabic anacrusis, or none at all. Thus the verses of for example Konrad von Würzburg contain genearlly 6 to 9 syllables. Fluid stress is occassionally found in Wolfram's poetry, and more so in Gottfried's, especially when his verse employs the word *Minne*.[107]

[107]Hartmann und Gottfried bevorzugen den schlankeren, silbenärmeren Vers, Wolfram kennt eher umfangreiche Auftakte und silbenreichere Taktfüllungen. Die nachfolgende Dichtergeneration meidet silbenreiche ebenso wie einsilbig gefüllte Takte tunlichst und bevorzugt den einsilbigen Auftakt gegenüber dem mehrsilbigen oder dem auftaktlosen Vers, so dass die Verse - z. B. bei Konrad von Würzburg - in aller Regel 6 bis 9 Silben enthalten. Mit schwebender Betonung muss man gelegentlich bei Wolfram rechnen, mehr noch bei Gottfried von Straßburg, vor allem wenn seine Verse mit dem Wort Minne einsetzen. Bögl, *Abriss der mittelhochdeutschen Metrik: mit einem Übungsteil*, 27

With the model above, we now have the tools to investigate some of these claims, and to better characterize the MHG epic, in both genre and chronology. We can also use these measures to begin to investigate the cause of this variation between texts.

Cadence

Cadence is one of the most notable opportunities for a poet or text to distinguish itself. Generally, MHG allows for five types of cadence:[108]

1. monosyllabic masculine: $|\acute{\times}\ \ \widehat{\ }$
2. bisyllabic masculine: $|\acute{\smile}\ \smile\ \ \widehat{\ }$
3. feminine full: $|\acute{\times}\ \ \times$
4. bisyllabic feminine: — $|\grave{\times}\ \ \widehat{\ }$
5. trisyllabic feminine: $|\acute{\times}\ \ \times\ \ |\grave{\times}\ \ \widehat{\ }$
6. Stricker cadence: The last syllable of a monosyllabic masculine cadence (primary stress) rhyming with the last syllable (secondary stress) of a bisyllabic feminine cadence[109]

MHG epic poetry is considered "rhyme pair" poetry; generally, rhyming lines will have the same cadence (the Stricker cadence naturally uniquely excepted).[110] These cadences may be better collapsed into 'full' (*voll*) or 'ringing' (*klingend*), where 'full' ends with a primary stressed syllable in the last foot, and 'ringing' ends with a secondary stressed syllable in the last foot (the word finishes resounding, falling from the primary stress in the double mora of the previous foot). An automated process of determining cadence allows for aggregate statistics on relationships, and even content extraction. But first it is useful to determine the extent to which poets took advantage of these different cadences. Table 5.13 gives the proportion of ringing cadences for each text. There is notable variation in the use of cadence. As Heusler and Ranke have pointed out, Ulrich von Liechtenstein and Thomasîn von Zirclaria (writing in vastly different areas of Germany), both use the ringing cadence sparingly.[111] Heusler goes as for as to say that it was "banned" from Ulrich's *Frauendienst*, although Eva Willms claims that "they aren't missing at all!"[112] and that depending on the Auftakt, many more can be counted than Ranke and Heusler consider.[113] In fact, this is

[108]Domanowski et al., *Mittelhochdeutsche Metrik Online*.

[109]Ibid.

[110]Bögl, *Abriss der mittelhochdeutschen Metrik: mit einem Übungsteil*, 26.

[111]Thomasin and Willms, *Der Welsche Gast*, 10; Heusler, *Deutsche Versgeschichte: mit Einschluss des altenglischen und altnordischen Stabreimverses*, 132-133; Ranke, *Sprache und stil im Wälschen gast des Thomasin von Circlaria*

[112]"Sie fehlen keineswegs." Thomasin and Willms, *Der Welsche Gast*, 10

[113]Heusler, *Deutsche Versgeschichte: mit Einschluss des altenglischen und altnordischen Stabreimverses*, 133.

what the model is detecting, as will be shown in the next section, that there is still debate as to whether a cadence should be scanned with a longer anacrusis and a ringing cadence, or a shorter anacrusis and a full cadence (a masculine bisyllabic cadence). The advantage of the supervised learning model is that it does not need to follow strict rules, but can learn from the human annotation provided. Thus, if one familiar with the works used as training data (*Der arme Heinrich*, *Parzival* and *Wigalois*), one would likely scan *Frauendienst* and *Der Welsche Gast* as having at least some ringing cadences, likely due to heavy syllables, commonly read as double morae, in the penultimate foot.

text	*klingend*
Das Turnier von Nantes	0.649654
Flore und Blanscheflur	0.630902
Eneide	0.627848
Lambrechts Alexander (Strassburger Hs.)	0.625217
Reinfried von Braunschweig	0.605649
Herzog Ernst (Hs. B)	0.595416
Lanzelet	0.592436
Helmbrecht	0.577479
Der Schwanritter	0.576602
Karl der Grosse	0.574975
Erec	0.571358
Der Trojanische Krieg	0.565297
Daniel von dem blühenden Tal	0.565264
Gregorius	0.564903
Willehalm (Wolfram)	0.557317
Dietrichs Flucht	0.555369
Engelhard	0.553506
Pantaleon	0.548193
Der arme Heinrich	0.540222
Gauriel von Muntabel	0.539256
Der Schlegel	0.537114
Dietrich und Wenezlan	0.537074
Tristan (Ulrich v. Türheim)	0.528813
Biterolf und Dietleib	0.521926
Alexander (U.v.E) (Ulrich von Eschenbach)	0.521857
Laurin	0.519088
Silvester	0.514171
Iwein	0.509309
Parzival	0.500443
Tristan (H.v.F.)	0.498766
Alexius	0.494334

Herzmaere	0.489796
Herzog Ernst (Hs.D, strophig)	0.479547
Tandareis und Flordibel	0.472036
Heinrich von Kempten	0.471429
Alexander (R. v. E.) (Rudolf von Ems)	0.465160
Meleranz	0.462276
Willehalm (U.v.T.)	0.460470
Barlaam und Josaphat	0.456013
Wigalois, der Ritter mit dem Rade	0.453288
Frauendienst (Büech.) (Büechlîn, Bechstein)	0.451781
Alexander Anhang	0.450000
Der guote Gêrhart	0.437283
Walberan	0.423539
Der Welsche Gast	0.360482
Frauendienst (Bechst.) (Epik, Bechstein)	0.293866

Table 5.13: Ratio of *klingend* cadences

Similar to the method presented in Chapter 3, it might prove fruitful to look at the specific language associated with these cadences. As the use of these cadences vary, Table 5.14 shows only the proportions of normalized lemma frequencies appearing in full and ringing cadences. Thus 'tragen' ('to carry/bear'), 'man' ('one'), 'jehen' ('affirm'), etc., appear over twice as frequently in full cadences, while 'gawein' (a name), 'heizzen' ('to call'), 'nie' ('never'), etc. appear nearly twice as likely in ringing cadences. It is no coincidence that common rhyme pairs emerge in the full cadences: 'jehen', 'sehen' ('to see'), 'komen' ('to come') are commonly employed in bisyllabic masculine cadences, 'lîp' ('body') and 'wîp' ('woman') in monosyllabic masculine cadences. Both Parzival's and Gawein's name more commonly appear in ringing cadences.

lemma	full/ringing
tragen	2.444500
man	2.346501
jehen	2.344162
stân	2.321150
wîp	2.300589
hant	2.157679
komen	2.065876
lîp	2.054963
bî	2.050834
ligen	2.035317
sehen	1.814803
herre	1.782718

ritter	1.780655
knabe	1.764819
lâzen	1.723517
...	...
niht	0.774092
wesen	0.770259
mêr	0.768657
parzival	0.761294
rîten	0.744589
minne	0.740137
si	0.735134
wol	0.707283
prîs	0.689569
haben	0.630292
werden	0.602728
von	0.588760
nie	0.561995
heizzen	0.560767
gawein	0.542168

Table 5.14: Lemma and cadence ratios for *Parzival*

But how is this manifested on the level of the text? Does the poet intentionally incorporate cadence changes? Where are groups of the most verses with a single cadence? In *Parzival*, where there are nearly an exact equal amount of full and ringing cadences, the greatest group of verses with a ringing cadence is when Feirefîz cites all the his battles, beginning with "rois papirîs von trogodjente."[114] Yet the greatest group of verses with a full cadence provides more insights. Gawein returns to Schastel marveile badly wounded and Arnive begins to heal him. Gawein is reinvigorated:

|vróuwe, |mîne |sín |nè, ‿
die mir |wā́ |rèn ent|rún |nèn, ‿
die |hábet |ír ge|wún|nèn ‿
|wíder |ín mîn |hér|zè: ‿
ouch |sénftet |sích mîn |smér|zè. ‿
swaz |ích |kréfte (o)der |sínne |hā́n, ‿
|díe hât |íuwer |díenest|màn ‿
|gár von |íuwern |schúl|dèn. ‿[115]

[114]Verse 770:1

[115]"Lady, my senses, which have escaped me, you've managed to bring them back into my heart: My pains are also subsiding. Whatever strength of senses that I may possess, then I have only you to thank." Eschenbach et al., *Parzival*, 580:8 - 580:15

We hear a dramatic string of double morae as Gawein declares his loyalty to her, emphasizing the fusion of physical and emotional excitement—'sinne', 'herze', 'smerze', with uncharacteristically sparse lines. Arnive responds:

> si sprach: |"hérre, |íuwern |húl|dèn ⌢
> |súl wir uns |álle |nâ̊|hèn ⌢
> und |dés mit |tríuwen |gâ̊|hèn. ⌢
> |nû volct |mír und en|rédet niht |víl. ⌢
> |éine |wúrz ich iu |gében |wil, ⌢
> dâ|vón ir |slâfet: |dást iu |gúot. ⌢
> |ézzens, |trínkens |kéinen |múot ⌢
> |súlt ir |háben |vór der |náht. ⌢
> |sô̊ kumt iu |wíder |íuwer |máht, ⌢
> sô |trit(e) ich |íu mit |spîse |zúo, ⌢
> daz |ír wol |bî|tet unz |vrúo." ⌢
> eine |wúrz si |légete (i)n |sînen |múnt: ⌢
> dô |slíef er |án der |sélben |stúnt. ⌢
> |wól si |sîn mit |décke |phlác. ⌢
> |álsus |ûber|slíef den |tác ⌢
> der |êren |rîche (u)nd |lásters |árm, ⌢
> |lác al |sánfte (u)nd |ím was |wárm. ⌢
> et|swénne (i)n |dóch in |slâ̊fe |vrô̊s, ⌢
> |dáz er |hésc|hète (u)nd |nô̊s, ⌢
> |állez |vón der |sálben |kráft. ⌢[116]

Arnive initially responds softly, in a similar manner with double morae in 'hulden', 'nâhen', and 'gâhen', showing Gawein kindness and respect. The following 17 lines, all in masculine cadence, convey a significant amount of information in a rushed tone, notably bearing more syllables per line, contrasting the previous drawn-out conversation. This commences precisely with the line "nû volct mir und enredet niht vil", as Arnive instructs Gawein to stop speaking, or rather stop with the drawn-out speech with double morae, I need to convey some important information; time also condenses accordingly. In the first exchange, the real time of exchange is depicted, while in the second exchange, a sequence of events is listed. These cadence types segue into a broader discussion of what Heusler calls '*Versfüllung*', viz. 'the fill of a verse'.

[116] "She spoke: 'Sir, we all aim to achieve your grace, and to go about this with loyalty. Now listen to me and do not speak much. I will give you a medicinal herb, which will make you sleep: that is good. You won't have any desire to eat or drink until the beginning of the night. Then you'll regain your strength. I will bring you food later, so that you'll greet us in the morning.' She placed the medicinal herb in his mouth and he went to sleep in the same moment. She carefully covered him with sheets. He slept for the entire day, rich with honor, poor in shame, he lay sotly and was warm. Whenever he froze in his sleep, then he coughed and sneezed, all due to the power of the ointment." Eschenbach et al., *Parzival*, 580:16 ff.

Versfüllung

As shown in the annotated data for the model, MHG verses have a minimum of five syllables and can range to having over 10 (although rare). This range of verse length is due to varying use of double morae and divided lifts/falls, as well as the use of anacrusis. While there can be no more than three syllables in anacrusis, even a one syllable ancrusis can significantly change the rhythm, as it takes over the observed rest in the last foot of the previous line. Depending on the anacrusis, a verse may be either synaptic or asynaptic. If the following verse begins with an anacrusis continuing the alternation (the anacrusis replaces the rest in the previous verse), then this verse is said to be synaptic.[117] If there is no anacrusis, the alternation is not preserved (the rest is observed), and the next verse begins asynaptic. As Heusler notes, the length of the anacrusis changed significantly through high *Minnesang* and MHG verse development, including the epic verse.[118] Table 5.15 shows the mean for various measures of line length. These results support Bögl's generalization of Hartmann's relatively short verses, and Wolfram's relatively long verses, as well as Heusler's claim that "Veldeke alienates with his many slim verses. Likewise in *Erec*,"[119] and that Ulrich von Eschenbach, an imitator of Wolfram, wrote sparse verses.[120]".[121] This has a significant impact on how a work would be presented to an audience. A clear example of this different rhythmical environment is provided by Hartmann in his *Gregorious*. Hartmann begins to describe the board sent out to sea accompanying baby Gregorius, on which his mother wrote of his noble origins and the sin that led to his birth:

720 Ein |tável |wárt ge|trágen |dár
721 der |vróuwen |díu daz |kínt ge|bár,
722 diu |víl guot |hél|fènbein |wás,
723 ge |zíeret |wól als |ích ez |lás
724 von |gólde und |vón ge |stéi|nè,
725 |dáz ich |níe de|héi|nè ⌢
726 |álsô |gúo|tè ge|wán. ⌢
727 |dâ |schréip diu |múoter |án ⌢
728 |sô si |méiste |máh|tè ⌢
729 |vón des |kíndes |áh|tè: ⌢
730 wan si |hâte |dés ge|dín|gèn ⌢
731 daz ez |gót |sólde |bríng|èn ⌢
732 den |líu|tèn ze hán|dèn

[117]Bögl, *Abriss der mittelhochdeutschen Metrik: mit einem Übungsteil*, 22.

[118]Heusler, *Deutsche Versgeschichte: mit Einschluss des altenglischen und altnordischen Stabreimverses*, 106-107.

[119]"Veldeke befremdet durch seine vielen magern Verse. Desgleichen der Erec"

[120]Heusler, *Deutsche Versgeschichte: mit Einschluss des altenglischen und altnordischen Stabreimverses*, 101.

[121]Ibid., 101, 112.

733 die |gót an |ím er|kán|dèn.[122]

The first five verses begin with a one syllable anacrusis, forcing a more metered breathing, and anticipation (not to mention recall for the performer) of the next verse. The next five lines are asynaptic, all without anacrusis, realizing the pause in the last foot and allowing for consecutive stressed syllables. Additionally, the middle lines all make use of double morae, beginning at the end of the verse, and working its way to the front in verse 727. The last verses of the group all end with double morae in the penultimate feet. The traveling double mora and shift in use of anacrusis, especially in consecutive verses, serves to highlight the poet's perspective on this board and redirect focus to its unique quality of containing text. Not only is anacrusis used as a stylistic, but it also provides us with a relative notion of the temporal space between verses. As Table 5.15 points out, Wolfram's verses frequently flow into the next, while Heinrich von Veldeke's more often observes the pause.

text	anacrusis	line length	foot length
Willehalm (U.v.T.)	1.121375	8.152451	1.977771
Willehalm (Wolfram)	1.030784	7.987001	1.943790
Frauendienst (Bechst.) (Epik, Bechstein)	1.023924	8.285598	2.049723
Parzival	0.997824	8.008826	1.965527
Karl der Grosse	0.968125	7.718863	1.876516
Pantaleon	0.957831	7.857275	1.925548
Das Turnier von Nantes	0.946367	7.625433	1.873414
Der Schwanritter	0.945682	7.778552	1.910399
Der Welsche Gast	0.939152	7.910625	1.966617
Frauendienst (Büech.) (Büechlîn, Bechstein)	0.936577	7.963510	1.970460
Herzog Ernst (Hs.D, strophig)	0.913838	7.839861	1.949811
Der Trojanische Krieg	0.913553	7.784082	1.921001
Der Schlegel	0.889908	7.790659	1.921323
Alexander (U.v.E) (Ulrich von Eschenbach)	0.881250	7.766500	1.920988
Herzog Ernst (Hs. B)	0.855007	7.573659	1.872170
Laurin	0.844115	7.735419	1.931071
Lanzelet	0.841297	7.530671	1.851255
Dietrich und Wenezlan	0.839679	7.498998	1.843019
Alexius	0.838527	7.803116	1.941454
Daniel von dem blühenden Tal	0.828912	7.573399	1.864914
Heinrich von Kempten	0.828571	7.842857	1.959307
Tristan (H.v.F.)	0.827356	7.761725	1.944678

[122]"A board was brought there to the lady who bore the child. The board was made of ivory and adorned beautifully, as I have read, with gold and with jewels: I have never had one so nice. The mother wrote on it to the best of her abilities about the child's origin, because she had the thought that God would bring the child to the hands of people, who would recognize God's workings in him." Hartmann and Mertens, *Der arme Heinrich*, ll. 729-732

Engelhard	0.826107	7.736470	1.928301
Dietrichs Flucht	0.823941	7.565123	1.877209
Der arme Heinrich	0.822760	7.569653	1.872902
Lambrechts Alexander (Strassburger Hs.)	0.812609	7.420290	1.820870
Gregorius	0.811533	7.523714	1.871443
Iwein	0.808795	7.558060	1.876490
Silvester	0.798928	7.698391	1.924614
Tandareis und Flordibel	0.782163	7.712322	1.936201
Biterolf und Dietleib	0.782148	7.638370	1.910000
Tristan (Ulrich v. Türheim)	0.781292	7.662825	1.920754
Helmbrecht	0.778926	7.481405	1.864325
Erec	0.774000	7.462842	1.865234
Walberan	0.771017	7.621297	1.915666
Wigalois, der Ritter mit dem Rade	0.759180	7.657728	1.931882
Barlaam und Josaphat	0.754516	7.700198	1.946321
Der guote Gêrhart	0.745809	7.623699	1.932081
Herzmaere	0.744898	7.668367	1.933107
Meleranz	0.738971	7.547779	1.904027
Flore und Blanscheflur	0.727080	7.313765	1.835540
Gauriel von Muntabel	0.717167	7.479952	1.886595
Alexander (R. v. E.) (Rudolf von Ems)	0.663663	7.411422	1.908049
Eneide	0.655593	7.241861	1.822112
Reinfried von Braunschweig	0.652001	7.251965	1.846291
Alexander Anhang	0.638571	7.286190	1.860159

Table 5.15: Anacrusis and filling of feet

Perhaps the most interesting aspect of MHG verse is the *beschwerte Hebung*, or double mora. As already seen, double morae allow for a syllable to carry twice the weight (and likely thus twice the duration) of a normal syllable. These instances are often considered important stylistic choices in indicating semantically, or contextually, important words in a verse: "only then did the monosyllabic foot become a 'declamatory machinery'; it exhilaratingly disrupted the up and down pacing and increasingly yielded the natural stress of the language."[123] Heusler notes that double morae were quite common in early MHG epics such as Heinrich von Veldeke's *Eneide*, and even had an "altertümlich", viz. antiquated, effect in later works containing many double morae, such as Hartmann's *Erec* and *Gregorius*.[124] The trend then began to move away from this "Germanic"[125] characteristic toward adopting a more strict

[123]"dann erst wurde der einsilbige Takt zum 'feinen deklamatorischen Kunstmittel'; dann unterbrach er belebend das vorherrschende Auf-ab und brachte den natürlichen Satzfall steigernd heraus." Heusler, *Deutsche Versgeschichte: mit Einschluss des altenglischen und altnordischen Stabreimverses*, 118

[124]Ibid., 101.

[125]"The monosyllabic foot was of course a German heirloom; not introduced by the French." ("Der einsilbige Takt war ja deutsches Erbstück; nicht welsche Einfuhr.") ibid., 118

style of stress alternation from the French tradition, seen in Rudolf von Ems' *Alexander* and nearly eliminated by Ulrich von Liechtenstein and Konrad von Würzburg, who aim for almost exclusively strict stress alternation.[126] Heusler understands Konrad as the endpoint, after whom German verse began to be written more freely once again. Table 5.16 confirms most of these observations. Of these texts, the top five most common double morae are: 'rîche' ('rich', 'powerful'), 'waere' ('would be'), 'maere' ('story'), 'ere' ('honor'), 'mêre' ('more') (closely followed by 'sêre' ('very')). All double morae must be heavy phonologically, but it is interesting that except for rîche, the top five are all *open* heavy syllables with *front* vowels. Furthermore, each of the top five double morae begins with a *sonorant* consonant (/m/, /n/, /w/, /j/, /l/, /r/), and except 'rîche', ends in 're'.[127] Sonorant consonants are voiced consonants with continuous airflow. Clearly, MHG poets aimed for the greatest resounding word to place in the double morae position, perhaps pointing to a greater vocal tradition in MHG epic verse than previously acknowledged. This observation supports the discussion of singing and syllabic phonology in Chapter 3, whereby open syllables allow for an unobstructed flow of air from the mouth, and may be preferred for notes of extended length (or melismatic syllables), especially considering the relative dearth of open, heavy syllable words in MHG. Table 5.17 shows the most common double mora for each text. It is quite clear that double morae characterize the content of a text quite well, which can come as no surprise given that double morae would present themselves most prominently in memory. Double morae are also often proper noun names.[128]

text	% double morae
Lambrechts Alexander (Strassburger Hs.)	0.097617
Eneide	0.095105
Flore und Blanscheflur	0.088926
Erec	0.080958
Daniel von dem blühenden Tal	0.079729
Herzog Ernst (Hs. B)	0.078354
Gregorius	0.077372
Lanzelet	0.077010
Helmbrecht	0.076015
Karl der Grosse	0.074808
Der arme Heinrich	0.074564
Alexander Anhang	0.072936
Dietrichs Flucht	0.071523
Iwein	0.070903

[126]Heusler, *Deutsche Versgeschichte: mit Einschluss des altenglischen und altnordischen Stabreimverses*, 101.

[127]Although the pronunciation of MHG 'w' is not certain.

[128]In respect to Wolfram, Heusler writes: "Monosyllabic feet served to denote propoer names." ("Einsilbiger Takt dient auch der Auszeichnung von Eigennamen.") Heusler, *Deutsche Versgeschichte: mit Einschluss des altenglischen und altnordischen Stabreimverses*, 119

Reinfried von Braunschweig	0.069159
Gauriel von Muntabel	0.068819
Dietrich und Wenezlan	0.063602
Das Turnier von Nantes	0.063528
Der Schlegel	0.062627
Biterolf und Dietleib	0.060387
Tristan (Ulrich v. Türheim)	0.060161
Laurin	0.058949
Alexander (U.v.E) (Ulrich von Eschenbach)	0.057748
Walberan	0.056834
Willehalm (Wolfram)	0.052707
Alexander (R. v. E.) (Rudolf von Ems)	0.051154
Wigalois, der Ritter mit dem Rade	0.050662
Der Schwanritter	0.049955
Meleranz	0.049557
Der Trojanische Krieg	0.048831
Engelhard	0.048730
Pantaleon	0.047476
Parzival	0.046408
Tandareis und Flordibel	0.045926
Tristan (H.v.F.)	0.045665
Silvester	0.045198
Barlaam und Josaphat	0.043626
Herzmaere	0.042582
Herzog Ernst (Hs.D, strophig)	0.041741
Willehalm (U.v.T.)	0.040853
Frauendienst (Büech.) (Büechlîn, Bechstein)	0.040585
Alexius	0.040388
Der guote Gêrhart	0.040166
Heinrich von Kempten	0.038583
Der Welsche Gast	0.037997
Frauendienst (Bechst.) (Epik, Bechstein)	0.006331

Table 5.16: Double morae ratio

text	double morae
Alexander (R. v. E.) (Rudolf von Ems)	rîche, persî, maere, waere, wîgant
Alexander (U.v.E) (Ulrich von Eschenbach)	rîche, waere, maere, swaere, daz
Alexander Anhang	alsô, daz, mit, stat, rîche
Alexius	eufêmiân, ougen, reine, sêre, lougen
Barlaam und Josaphat	lêre, rîche, sêre, mêre, arbeit

Biterolf und Dietleib	maere, waere, wîgant, rîche, rîchen
Daniel von dem blühenden Tal	waere, niht, daz, sô, alle
Das Turnier von Nantes	engellanden, reine, handen, guldîn, gesteine
Der Schlegel	alten, behalten, wære, besunder, ende
Der Schwanritter	brâbant, hiute, swaere, beiden, maere
Der Trojanische Krieg	waere, rîche, stunden, werden, reine
Der Welsche Gast	mêre, solde, wârheit, niht, sinne
Der arme Heinrich	heinrich, herre, wære, arbeit, güete
Der guote Gêrhart	rîche, güete, guote, waere, muote
Dietrich und Wenezlan	bôlân, strîten, rîten, berne, zîten
Dietrichs Flucht	maere, dietrîch, bernaere, berne, gerne
Eneide	rîche, wâre, troiân, mâre, turnûs
Engelhard	werden, maere, erden, waere, engelhart
Erec	waere, êrec, herre, mêre, êre
Flore und Blanscheflur	waere, sô, daz, mêre, swaere
Frauendienst (Bechst.) (Epik, Bechstein)	helm, heinrîch, mîn, alsô, sîn
Frauendienst (Büech.) (Büechlîn, Bechstein)	êre, mêre, sêre, swaere, maere
Gauriel von Muntabel	waere, maere, walbân, rîche, gâwân
Gregorius	waere, swaere, guote, herre, maere
Heinrich von Kempten	heinrich, truhsaeze, kempten, hende, zîten
Helmbrecht	helmbreht, gotelinde, maere, waere, muoter
Herzmaere	cleine, frouwen, sinne, beide, schouwen
Herzog Ernst (Hs. B)	rîche, lande, wîgant, maere, waere
Herzog Ernst (Hs.D, strophig)	daz, strîten, maere, liute, swaere
Iwein	waere, êre, arbeit, îwein, mêre
Karl der Grosse	ruolant, heiden, êre, sêre, rîche
Lambrechts Alexander (Strassburger Hs.)	alexander, rîche, wâren, dô, daz
Lanzelet	waere, maere, êre, wîgant, artûs
Laurin	laurîn, dietrîch, wîgant, dietleip, berne
Meleranz	rîche, zîten, wîgant, waere, maere
Pantaleon	werden, erden, ougen, heiden, worden
Parzival	gâwân, rîche, maere, waere, minne
Reinfried von Braunschweig	sinne, minne, wære, rîche, mære
Silvester	werden, erden, reine, waere, haete
Tandareis und Flordibel	rîche, wîgant, maere, waere, êre
Tristan (H.v.F.)	minne, herzen, marke, sinne, küneginne
Tristan (Ulrich v. Türheim)	maere, waere, ýsôt, minne, sôte
Walberan	laurîn, gerne, waere, berne, dietrîch
Wigalois, der Ritter mit dem Rade	manheit, waere, korntîn, rîche, maere
Willehalm (U.v.T.)	niht, markîs, waere, daz, minne
Willehalm (Wolfram)	rîche, waere, markîs, minne, heimrîch

Table 5.17: Double morae word frequency

To measure Heusler's diversity of 'Versfüllung' in its entirety, I adopt a popular measure of diversity, originally developed in the context of information theory and widely used in environmental science, that of Shannon entropy.[129] Shannon entropy fits this application quite well, considering Shannon's original problem dealt with string prediction given a set of characters, the entropy measure essentially quantifies the degree of certainty in predicting a random character from a string dataset, given a finite number of characters. Similarly, a useful measure of 'Versfüllung' entails the degree of certainty in predicting a specific foot or line. To this end, I calculate the Shannon entropy for a complete foot (excluding anacrusis and the last foot) and for the entire line, sampling 1,000 feet with replacement from each text. The results are shown in Table 5.18. Once again, the earlier, more "Germanic" texts display a wider diversity in metrical patterning, while Ulrich von Liechtenstein and Konrad von Würzburg aim toward monotonous trochees. We see all the works attributed to Hartmann von Aue in the top ten for diversity. Save *Parzival* and *Der Welsche Gast*, the top ten also contains the works most often discussed by scholars.

text	foot	line
Eneide	1.252917	3.763736
Lambrechts Alexander (Strassburger Hs.)	1.194705	3.980579
Iwein	1.167059	3.405163
Daniel von dem blühenden Tal	1.140783	3.498383
Helmbrecht	1.099641	3.039134
Der arme Heinrich	1.094388	3.217970
Laurin	1.030382	3.407795
Erec	1.005811	3.294478
Gregorius	0.980259	3.133976
Willehalm (Wolfram)	0.979950	3.159355
Tristan (Ulrich v. Türheim)	0.971779	3.185109
Alexander Anhang	0.969258	3.248891
Karl der Grosse	0.968051	2.983325
Der Schlegel	0.934412	3.277041
Meleranz	0.922810	2.891814
Barlaam und Josaphat	0.917566	2.968009
Flore und Blanscheflur	0.917468	3.115846
Herzog Ernst (Hs. B)	0.914392	3.191947
Gauriel von Muntabel	0.906261	3.041430
Tristan (H.v.F.)	0.903151	3.103300
Dietrichs Flucht	0.894119	3.041568

[129]C. E. Shannon, "A mathematical theory of communication," *The Bell System Technical Journal* 27, no. 3 (July 1948): 379–423.

Walberan	0.884103	3.246322
Willehalm (U.v.T.)	0.881089	3.078347
Heinrich von Kempten	0.875890	2.665526
Parzival	0.868807	3.217873
Alexander (U.v.E) (Ulrich von Eschenbach)	0.867947	2.963053
Tandareis und Flordibel	0.867046	2.642869
Biterolf und Dietleib	0.854046	2.824735
Dietrich und Wenezlan	0.852161	2.974738
Pantaleon	0.850183	2.725092
Lanzelet	0.848197	3.022754
Alexius	0.845937	2.837503
Das Turnier von Nantes	0.809099	2.304515
Der Trojanische Krieg	0.803099	1.792531
Engelhard	0.800056	1.944227
Herzmaere	0.770970	2.488530
Alexander (R. v. E.) (Rudolf von Ems)	0.768860	2.413441
Der Schwanritter	0.761848	2.485584
Der Welsche Gast	0.710632	2.645962
Silvester	0.703183	2.412775
Frauendienst (Büech.) (Büechlîn, Bechstein)	0.690972	2.655543
Herzog Ernst (Hs.D, strophig)	0.666231	2.460030
Der guote Gêrhart	0.624613	2.388511
Wigalois, der Ritter mit dem Rade	0.622915	2.714158
Reinfried von Braunschweig	0.621006	1.908564
Frauendienst (Bechst.) (Epik, Bechstein)	0.303572	1.881531

Table 5.18: Shannon entropy for 1000 random line samples

While each of these measures should be further investigated for individual texts, this project has illuminated the possibilities of an applied model for MHG scansion, despite disagreement in the scholarship over the correct theory.

5.10 Comparison to sound patterning method

Using this machine learning model developed for MHG epic poetry scansion, we can compare the model's output to the features encoded in the previous chapter using prosodic sequence patterning. While not providing any practical use, such a comparison can shed light on just how connected phonology and meter in MHG are, though we already know from the model and pedagogical texts that meter is highly dependent on phonology. All of the texts used in the metrical analysis of the previous chapter were thus both scanned by line and encoded with the sequence method described in Chapter 4. Each line sequence was assigned

all the observed scansions for that particular sequence; e.g. sequence 'C-C-C-C-OC-O-C' is observed ten times in the corpus, but has only two observed scansions:

$$|\acute{\times} \times |\acute{\times} \times |\acute{\times} \smile \smile |\acute{\times} \ \widehat{\ }$$

and

$$\times |\acute{\times} \times |\acute{\times} \times |\acute{\times} \times |\acute{\times} \ \widehat{\ }$$

66% of all 55,253 sequences have only one unique observed scansion, though many of these only have one observation. 22.3% of sequences with at least five observations have only one unique observation. The most certain sequence for scansion purposes is 'C-C-C-CC-CC-C', which is observed 502 times and carries only the scansion:

$$\times |\acute{\times} \times |\acute{\times} \times |\acute{\times} \times |\acute{\times} \ \widehat{\ }$$

This implies that if the verse ends in two closed bisyllabic words followed by a closed monosyllabic word, the scansion must be the above typical trochaic tetrameter. On the other end of the spectrum we see the sequence 'C-CC-C-OC-OC' observed 51 times with 11 different scansions[130]:

(1) $\times |\text{—} |\grave{\times} \smile \smile |\acute{\times} \times |\acute{\times} \ \widehat{\ }$

(2) $\times |\text{—} |\grave{\times} \times |\acute{\times} \smile \smile |\acute{\times} \ \widehat{\ }$

(3) $\times |\text{—} |\grave{\times} \times |\acute{\times} \times |\breve{\smile} \smile \ \widehat{\ }$

[130]The Shannon entropy was also calculated here for each sequence to rank diversity of scansions within an observed phonetic sequence

(4) × × |x́ × |— |x̀ × |x́ ⌢

(5) × × |x́ × |x́ × |— |x̀ ⌢

(6) × |x́ ⏑ ⏑ |— |x̀ × |x́ ⌢

(7) × |x́ ⏑ ⏑ |x́ × |— |x̀ ⌢

(8) × |x́ × |— |x́ × |⏑ ⏑ ⌢

(9) × |x́ × |x́ ⏑ ⏑ |— |x̀ ⌢

(10) × |x́ × |x́ × |x́ × |x́ ⌢

(11) x́ × |x́ × |x́ × |⏑ ⏑ ⌢

Notably, both sequences are eight syllables (the average length of a MHG line), yet it is suprising just how much scansion can vary (or not) based solely on the phonology of the syllables.

5.11 MHG Scansion GUI Software for General Use

In hopes of supporting students of Middle High German, this project has also developed a graphical interface to scan MHG epic texts titled *Automatische Mittelhochdeutsche Skandierungssystem* (AMS). The software is available online.[131]

[131] `https://github.com/henchc/MHG_Scansion`

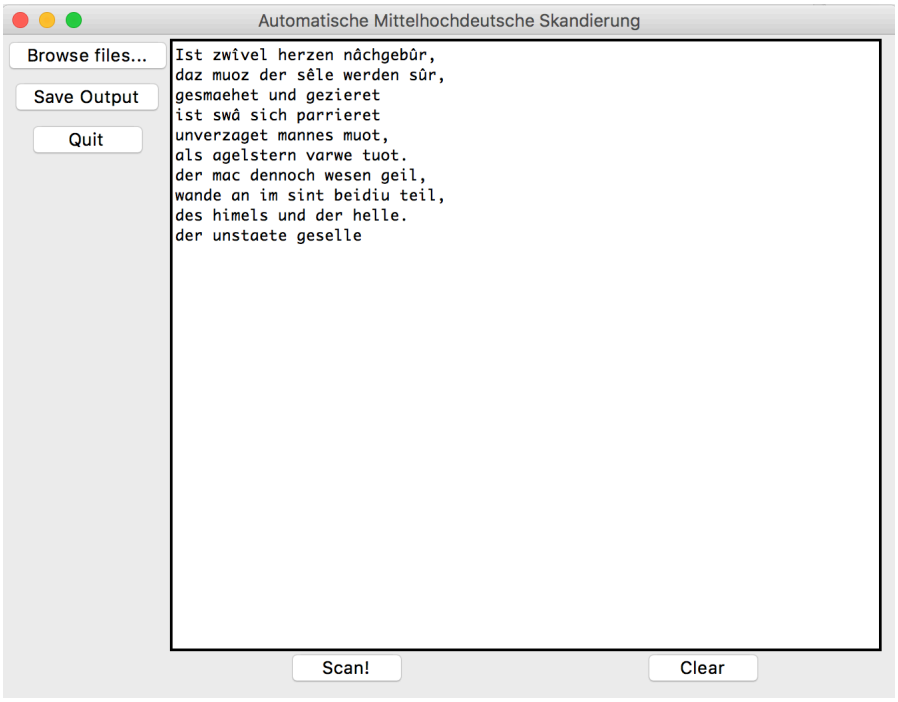

Figure 5.3: *Automatische Mittelhochdeutsche Skandierung* software

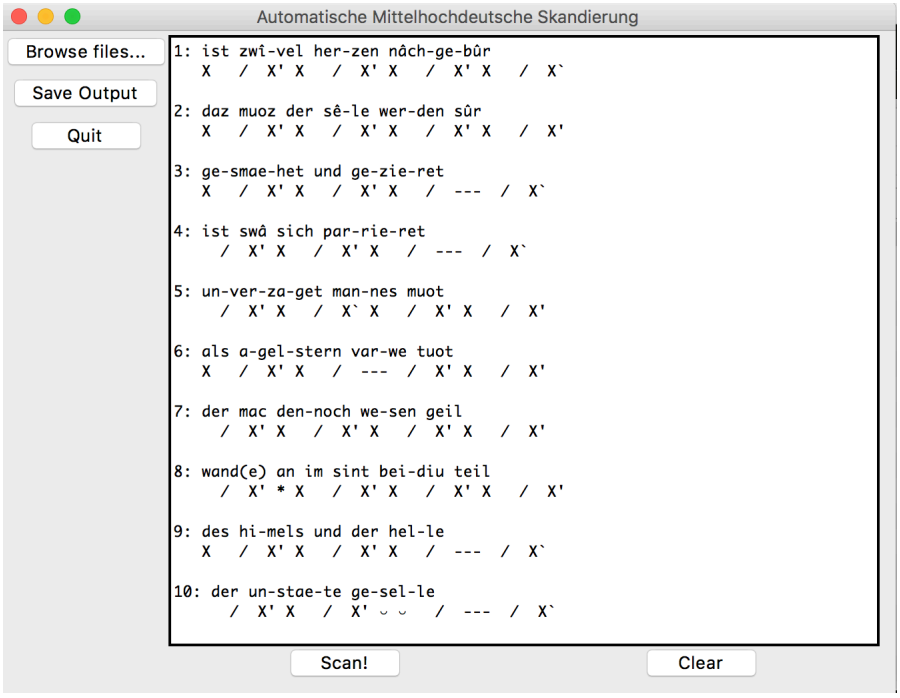

Figure 5.4: *Automatische Mittelhochdeutsche Skandierung* output

Chapter 6

Conclusion and Further Work

6.1 Sensory Perception

The first chapter of this project suggested that we may be able to improve our knowledge of medieval sensory perception through a quantitative analysis without having a source of data equivalent to what can be gathered today from cutting edge technology. By aggregating representations of sound over the extant medieval German lyric and epic corpus, we have indeed seen that the poets, and likely the audience, were intimately aware of rhythmical and auditory patternings. In fact, the discussion in the first chapter suggests that these senses may have played an even more central role in the culture, particuarly to its memory, than today. While I have not intended to argue that poets were actually composing poetry by counting syllable properties, I have argued that distinct, imitable styles existed and are chronologically traceable. Moreover, these styles group thematically and cluster among poets and their successors. These patterns necessarily affect cognition through associative patterning, of which poets took advantage to produce different effects depending on the content. We see intentional juxtapositions, or what März calls a "Verfremdungseffekt", when these expectations are challenged.

This project has devoted much attention to the syllable. As demonstrated in the first chapter, the syllable was a major focus of medieval composers of poetry and music. The syllable serves as a fundamental unit of sound for us to analyze this sensory perception. Voice, rhythm, and memory all draw upon the rhythmic and auditory products rooted in the syllable. Medievalists are aware of the importance of the voice to the performance of these texts, but this aspect cannot be emphasized enough. As Stock correctly reminds us, if we hope to better understand these texts, their composition, and audience, we must attempt to characterize the role of the voice in this performance.[1] Without any recordings, we can only look at the phonological evidence that remains. Unfortunately, this evidence is noisy due to linguistic and orthographic variation. This project has demonstrated that abstracting from this noise to the syllabic level can help unite an otherwise diverse corpus.

[1] Stock, "Das volle Wort - Sprachklang im späteren Minnesang."

6.2 Digital Humanities and Medieval Studies

This project has presented several new methods and techniques for the digital humanities to navigate the medieval landscape. Medievalists are certainly presented with unique challenges when compared to scholars of other periods. Our texts are arguably less "stable", we are not certain about the origin of our texts, there is heavy linguistic variation, editing practices often mediate what we consume as literature, our corpus size is much smaller, etc. I have tried here to offer several different paths for medievalists to explore. Most importantly, I present a method to unite a corpus that is linguistically diverse by abstracting from orthographic noise to the syllable.

I hope to have also overcome the question digital humanists fear most: "Have we learned anything *new*?" Often we use digital methods to confirm what we already know, or quantify it. Moretti believes this makes our knowledge more "exact".[2] By working on a formal level of the text, one that is often not as noticeable to readers today, this project has revived observations made by previous generations by confirming many of Heusler's claims, but also discovered new observations beyond making our knowledge more "exact". It is easier for us to internalize words and concepts of a corpus, but more difficult to aggregate, associate, and generalize about rhythmical patterns (especially when these patterns are yet to be fully understood). Thus a rhythmical analysis has proved to be fruitful in producing new knowledge in the scholarship.

This project also emphasizes *relative* similarities and differences within a corpus. Especially when working with historical texts, establishing an absolute is nearly impossible. We cannot know for certain what medieval German sounded like. Nevertheless, these methods have allowed us to compare how different poets understood the relationship between sounds (and even symbols) in the language. This method also affords us some room for inaccuracy. Just as we will never know what medieval German sounded like, we also will likely never discover whether Heusler's theory of medieval German meter is "correct". However, Heusler's theory can be still be utilized. When we look at the differences between texts, we see how they differ *in reference to Heusler's theory*, but underlying Heusler's theory is simple phonological evidence (mediated through manuscripts). Therefore, while we may say that one text exhibits a greater usage of double morae, what we are truly saying is that this text differs from others in its use of fewer syllables in a line and likely more heavy syllables in that line (among other things). How that was realized in practice, we cannot be certain, but it is very likely that it sounded different than feet filled with half morae, or several light syllables. However we interpret it, it is a different auditory experience for the audience, which can be aggregated and associated with different content. I hope that future work can build on this direction of looking at *relative* relationships as opposed to focusing on *absolute* statistics.

[2]Franco Moretti, ""Operationalizing": or, the function of measurement in modern literary theory," *Pamphlets of the Stanford Literary Lab*, no. 6 (2013).

6.3 Interdisciplinary Work

Such digital work in medieval studies requires interdisciplinary skills, which are not foreign to medievalists. Our training often includes multiple languages, an insurmountable time period, and a strong philological foundation. This collection of skills suggests that the digital humanities may find more friends among medievalists than not, and I hope this is the case. Nevertheless, most work in the digital humanities requires us to reach across disciplinary divides, and even outside of the humanities. This project drew upon the knowledge of linguists, musicologists, and computer scientists.

Beyond medieval Germanists, the methods presented in this paper would also undoubtedly be of use in related disciplines. The borrowings amongst medieval traditions would only become more apparent through a sound and rhythmic interdisciplinary comparison of verse. Temporally, these methods also present new opportunities to connect poets, as we know that the *Meistersänger* revered the *Minnesänger*, and intentionally imitated their work.

6.4 Future Work

The methods demonstrated in this project have proven to be fruitful when applied to the MHG corpus, illuminating hitherto unnoted prosodic idiosyncracies, and aiding in visualizing the development of MHG verse. This project has offered both a novel universal syllabification algorithm, useful for all languages but intended for low-resource historical languages, and a customized syllabification algorithm for MHG. The unique, syllable-based formal structure of MHG allows for a fascinating investigation of sound, which allows for calculations of discrepencies and unities in form and content. The previous chapter has brought research on MHG meter one step further, presenting a novel supervised method for poetic scansion (a model already acknowledged and implemented by other traditions with the results here as benchmark[3]), permitting a glimpse into the most decisive features, as well as an idea of difficulties even human scanners encounter when reading MHG poetry. It is the hope that these new methodologies are scrutinized, improved upon, and implemented further in studies of MHG sound and rhythm, and specifically aid in the reading of a more narrow corpus, or even a single text, and that MHG literature will begin to once again be considered in light of its rhythmic and sonorous qualities.

While this project has hesitated to draw any conclusions about the role of music in this poetry, it has hinted at a stronger relationship than previously acknowledged. I hope that music's role in this corpus is further researched along the line taken in this project.

This project has also begun to approach one of the greatest problems for medievalists, that of linguistic normalization and digital work. I hope to have set a standard here for subsequent digital work, whereby authors do not hesitate to use the texts available to them,

[3]Agirrezabal et al., "Machine Learning for Metrical Analysis of English Poetry"; Navarro-Colorado, "A metrical scansion system for fixed-metre Spanish poetry"; Mittmann et al., "Escansão automática de versos em português."

yet are aware of the discrepancies vis-à-vis the manuscripts, and provide relevant statistics to support their use of the corpus.

Bibliography

Adsett, Connie R, and Yannick Marchand. "A comparison of data-driven automatic syllabification methods." In *International Symposium on String Processing and Information Retrieval*, 174–181. Springer, 2009.

Agirrezabal, Manex, Inaki Alegria, Mans Hulden, and IXANLP Group. "Machine Learning for Metrical Analysis of English Poetry." *Proceedings of COLING 2016, the 26th International Conference on Computational Linguistics: Technical Papers*, 2016, 772–781.

Agirrezabal, Manex, Bertol Arrieta, Aitzol Astigarraga, and Mans Hulden. "ZeuScansion: a tool for scansion of English poetry." *Finite State Methods and Natural Language Processing*, 2013, 18–24.

Anttila, Arto, and Ryan Heuser. "Phonological and metrical variation across genres." In *Proceedings of the Annual Meetings on Phonology*, vol. 3. 2016.

Arvaniti, Amalia. "Rhythm, Timing and the Timing of Rhythm." *Phonetica* 66, no. 1 (April 2009): 46–63.

Assmann, Jan. "Die Katastrophe des Vergessens. Das Deuteronomium als Paradigma kultureller Mnemotechnik," 1994.

Aue, Hartmann von. *Der arme Heinrich*. Bibliothek des Mittelalters 6. Deutscher Klassiker Verlag, 2004.

Aurelian. *The discipline of music: Musica disciplina*. Colorado College Music Press Translations 3. Colorado Springs: Colorado College Music Press, 1968.

Aurelian and Lawrence Arthur Gushee. *Musica disciplina*. 21. American Institute of Musicology, 1975.

Bartlett, Susan, Grzegorz Kondrak, and Colin Cherry. "Automatic Syllabification with Structured SVMs for Letter-to-Phoneme Conversion." In *ACL 2008*, 568–576. 2008.

———. "On the syllabification of phonemes." In *Proceedings of Human Language Technologies: The 2009 Annual Conference of the North American Chapter of the Association for Computational Linguistics*, 308–316. Association for Computational Linguistics, 2009.

Bartsch, Karl. *Die Schweizer Minnesänger*. Frauenfeld: Huber, 1964.

Bartsch, Karl, and Helmut de Boor, eds. *Das Nibelungenlied*. 20th ed. Deutsche Klassiker des Mittelalters 3. Wiesbaden: Brockhaus, 1972.

Batts, Michael S., Saint Gall (Switzerland), Bayerische Staatsbibliothek, and Fürstlich Fürstenbergische Hofbibliothek (Donaueschingen, Germany), eds. *Das Nibelungenlied. Paralleldruck der Handschriften A, B und C nebst Lesarten der übrigen Handschriften*. Tübingen: Niemeyer, 1971.

Bäuml, Franz H. "The Oral Tradition and Middle High German Literature." *Oral Tradition* 1 (1986): 398–445.

Bäuml, Franz Heinrich, and Agnes M. Bruno. "Weiteres zur mündlichen Überlieferung des Nibelungenliedes." *Deutsche Vierteljahrsschrift für Literaturwissenschaft und Geistesgeschichte* 46 (1972): 479–493.

Bede, Bede, and Calvin B. Kendall. *Libri II De arte metrica ; et, De schematibus et tropis: The art of poetry ; and, Rhetoric*. Edited by Stiftsbibliothek Sankt Gallen. Bibliotheca Germanica 2. Saarbrücken [Germany]: AQ-Verlag, 1991.

Benjamin, Walter. *Das Kunstwerk im Zeitalter seiner technischen Reproduzierbarkeit: drei Studien zur Kunstsoziologie*. 2nd ed. Edition Suhrkamp 28. Frankfurt am Main: Surhkamp, 1968.

Bennewitz, Ingrid. "Von Nachtigallen, Krähen, Hühnern, und Sägern: Überlegungen zu Aufführung und Sängerrollen im Minnesang, speziell bei Neidhart." In *Edition und Interpretation: neue Forschungsparadigmen zur mittelhochdeutschen Lyrik: Festschrift für Helmut Tervooren*, edited by Helmut Tervooren and Johannes Spicker, 73–85. Stuttgart: Hirzel, 2000.

Bernhard, Delphine, and Anne-Laure Ligozat. "Es esch fàscht wie Ditsh, oder net? Étiquetage morpho-syntaxique de l'alsacien en passant par l'allemand." *Actes de TALARE 2013: Traitement Automatique des Langues Régionales de France et d'Europe*, 2013, 209–220.

———. "Hassle-free POS-tagging for the Alsatian dialects." *Non-Standard Data Sources in Corpus Based-Research*, 2013, 85–92.

Bird, Steven, Edward Loper, and Ewan Klein. *Natural Language Processing with Python*. O'Reilly Media Inc., 2009.

Bögl, Herbert. *Abriss der mittelhochdeutschen Metrik: mit einem Übungsteil*. Hildesheim; New York: G. Olms, 2006.

Bohrn, Isabel C., Ulrike Altmann, Oliver Lubrich, Winfried Menninghaus, and Arthur M. Jacobs. "When we like what we know – A parametric fMRI analysis of beauty and familiarity." *Brain and Language* 124, no. 1 (January 2013): 1–8.

Bolinger, D. L. "Pitch accent and sentence rhythm." In *Abe, Kanekiyo, Forms of English: accent, morpheme, order*, 139–180. Cambridge: Harvard University Press, 1965.

Boll, Katharina. *Also redete ein vrowe schoene: Untersuchungen zu Konstitution und Funktion der Frauenrede im Minnesang des 12. Jahrhunderts*. Würzburger Beiträge zur deutschen Philologie, Bd. 31. Würzburg: Königshausen & Neumann, 2007.

Braun, Manuel. "Aufmerksamkeitsverschiebung. Zum Minnesang des 13. Jahrhunderts als Form- und Klangkunst." *Wolfram-Studien*, Veröffentlichungen der Wolfram von Eschenbach-Gesellschaft, 21 (2013): 203–230.

Braune, Wilhelm, and Ernst A Ebbinghaus. *Althochdeutsches Lesebuch*. Tübingen: Niemeyer, 1994.

Brill, Eric. "Transformation-based error-driven learning and natural language processing: A case study in part-of-speech tagging." *Computational linguistics* 21, no. 4 (1995): 543–565.

Brunner, Horst, and Helmut Tervooren. "Einleitung: Zur Situation der Sangspruch- und Meistergesangsforschung." *Zeitschrift für deutsche Philologie* 119, no. 2000 (2000): 1–9.

Bumke, Joachim. "Der unfeste Text." *Germanistische Symposien Berichtsbanden* 17 (1996): 118–129.

Carruthers, Mary. "Sweetness." *Speculum* 81, no. 4 (2006): 999–1013.

Charles Kensington Salaman Esq., Hon. Mem. Acad. "On the English Language as a Language for Music." *Proceedings of the Musical Association* 3, no. 1 (January 1, 1876): 120–139.

Chomsky, Noam, and Morris Halle. *The sound pattern of English*. 1st MIT Press paperback. ed. Cambridge, Mass: MIT Press, 1991.

Clement, Tanya, and Stephen McLaughlin. "Measured Applause: Toward a Cultural Analysis of Audio Collections." *Journal of Cultural Analytics*, May 23, 2016. `http://culturalanalytics.org/2016/05/measured-applause-toward-a-cultural-analysis-of-audio-collections/`.

Cramer, Thomas. *Waz hilfet âne sinne kunst?: Lyrik im 13. Jahrhundert Studien zu ihrer Ästhetik*. Philologische Studien und Quellen, Heft 148. Berlin: E. Schmidt, 1998.

Curschmann, Michael. "The Concept of the Oral Formula as an Impediment to Our Understanding of Medieval Oral Poetry." *Medievalia et Humanistica* 8 (1977): 63–76.

Dasher, Richard, and Dwight Bolinger. "On pre-accentual lengthening." *Journal of the International Phonetic Association* 12, no. 2 (December 1982): 58–71.

Davis, Hugh H. "The 'De rithmis' of Alberic of Monte Cassino: A Critical Edition." *Mediaeval Studies* 28 (January 1, 1966): 198–227. ISSN: 0076-5872.

Dillon, Emma. *The sense of sound: musical meaning in France, 1260-1330*. New York: Oxford University Press, 2012.

Dimpel, Friedrich. *Computergestützte textstatistische Untersuchungen an mittelhochdeutschen Texten - Bd. II. Daten und Programme*. Tübingen: A. Francke, 2004.

———. "Textstatistische Analysen an mittelhochdeutschen Texten." *Jahrbuch für Computerphilologie* 6 (2004).

Dobozy, Maria. *Re-membering the present: the medieval German poet-minstrel in cultural context*. Disputatio 6. Turnhout, Belgium : Abingdon: Brepols; Marston [distributor], 2005.

Domanowski, Anna, Yochanan Rauert, Hanno Rüther, and Tomas Tomasek. *Mittelhochdeutsche Metrik Online*. https://www.uni-muenster.de/MhdMetrikOnline/. Accessed on 04-02-2016, 2009.

Egenes, Thomas. *Introduction to Sanskrit*. Motilal Banarsidass Publ., 1996.

Ellendt, Joh. Ernst. *Ueber den Einfluss des Metrums auf Wortbildung und Wortverbindung bei Homer*. Königsberg: Druck der Universitäts-Buch- und Steindruckerei von E.J. Dalkowski, 1861.

Erickson, Raymond, and Claude V. Palisca, eds. *Musica enchiriadis ; and, Scolica enchiriadis*. Music theory translation series. New Haven: Yale University Press, 1995.

Eschenbach, Wolfram von, Karl Lachmann, Eberhard Nellmann, and Dieter Kuhn. *Parzival*. 1 Aufl. Bibliothek deutscher Klassiker. Frankfurt am Main: Deutscher Klassiker Verlag, 1994.

Estes, Alex, and Christopher Hench. "Supervised Machine Learning for Hybrid Meter." *Proceedings of the Fifth Workshop on Computational Linguistics for Literature, NAACL-HLT 2016*, 2016, 1.

Fassler, Margot E. "Accent, Meter, and Rhythm in Medieval Treatises "De rithmis"." *The Journal of Musicology* 5, no. 2 (April 1987): 164–190.

Firth, John Rupert. "Sounds and prosodies." *Transactions of the Philological Society*, 1948, 127–152.

Fox, Anthony. *Prosodic Features and Prosodic Structure: The Phonology of Suprasegmentals*. New York: Oxford University Press, 2000.

Fulton, Rachel. ""Taste and see that the Lord is sweet" (Ps. 33:9): The Flavor of God in the Monastic West." *The Journal of Religion* 86, no. 2 (April 2006): 169–204.

Giegerich, Heinz J. *Metrical Phonology and Phonological Structure: German and English*. Cambridge: Cambridge University Press, 1985.

Goodwin, Jonathan, and John Holbo, eds. *Reading Graphs, Maps & Trees: Responses to Franco Moretti*. Anderson, SC: Parlor Press, 2011.

Gottfried, Friedrich Ranke, and Rüdiger Krohn. *Tristan*. 2nd ed. Universal-Bibliothek, 4471-4472. Stuttgart: Reclam, 1981.

Grabe, E., and E. Low. "Durational variability in speech and the rhythm class hypothesis." *Papers in Laboratory Phonology* Vol. 7 (January 1, 2002): 515–546.

Greene, Erica, Tugba Bodrumlu, and Kevin Knight. "Automatic analysis of rhythmic poetry with applications to generation and translation." In *Proceedings of the 2010 conference on empirical methods in natural language processing*, 524–533. Association for Computational Linguistics, 2010.

Gumbrecht, Hans Ulrich. "Rhythmus und Sinn." In *Materialität der Kommunikation*, 1st ed., edited by Hans Ulrich Gumbrecht, Karl Ludwig Pfeiffer, and Monika Elsner, 714–729. Suhrkamp-Taschenbuch Wissenschaft 750. Frankfurt am Main: Suhrkamp, 1988.

Gumbrecht, Hans Ulrich, Karl Ludwig Pfeiffer, and Monika Elsner, eds. *Materialität der Kommunikation*. 1st ed. Suhrkamp-Taschenbuch Wissenschaft 750. Frankfurt am Main: Suhrkamp, 1988.

Haferland, Harald. *Hohe Minne: zur Beschreibung der Minnekanzone*. Beihefte zur Zeitschrift für deutsche Philologie 10. Berlin: Erich Schmidt Verlag, 2000.

———. "Minnesang als Posenrhetorik." In *Text und Handeln: zum kommunikativen Ort von Minnesang und antiker Lyrik*, edited by Albrecht Hausmann, Cornelia Logemann, and Christian Rode, 65–105. Beihefte zum Euphorion 46. Heidelberg: Winter, 2004.

Hartman, Charles O. *Virtual Muse: Experiments in Computer Poetry*. Hanover, N.H.: Wesleyan University Press, 1996.

Hartmann and Volker Mertens. *Der arme Heinrich*. Bibliothek des Mittelalters. Cambridge, [England] : Ann Arbor, MI: Chadwyck-Healey ; ProQuest Information / Learning, 2005.

Haustein, Jens. "Grenzgänger Formexperimente in der Sangspruchdichtung des Marner, Konrads von Würzburg und Frauenlobs." In *Sangspruchdichtung um 1300: Akten der Tagung in Basel vom 7. bis 9. November 2013*, edited by Gert Hübner, Dorothea Klein, and Universität Basel, 249–262. Spolia Berolinensia 33. Hildesheim: Weidmann, 2015.

Hayes, Bruce. "Compensatory Lengthening in Moraic Phonology." *Linguistic Inquiry* 20 (1989): 253–306.

Haymes, Edward R. *Mündliches Epos in mittelhochdeutscher Zeit*. Vol. 164. Göppinger Arbeiten zur Germanistik. Göppingen, 1975.

Heinrich Düntzer. *Homerische Abhandlungen von Heinrich Düntzer*. Edited by National Central Library of Florence. Hahn, 1872.

Heinzle, Joachim. *Einführung in die mittelhochdeutsche Dietrichepik*. De Gruyter Studienbuch. Berlin ; New York: W. de Gruyter, 1999.

Hench, Christopher. "Alienated Bodies and the lîp/wîp Connection." *Interdisciplinary Journal for Germanic Linguistics and Semiotic Analysis* 21, no. 2 (2016): 165–189.

———. "Phonological Soundscapes in Medieval Poetry." *Proceedings of the Joint SIGHUM Workshop on Computational Linguistics for Cultural Heritage, Social Sciences, Humanities and Literature, ACL 2017*, 2017, 46.

Hennig, Ursula, ed. *Das Nibelungenlied: nach d. Hs. C*. 1. Aufl. Altdeutsche Textbibliothek 83. Tübingen: Niemeyer, 1977.

Heusler, Andreas. *Deutsche Versgeschichte: mit Einschluss des altenglischen und altnordischen Stabreimverses*. 2nd ed. Vol. 2. Grundriss der germanischen Philologie. Berlin: W. De Gruyter, 1956.

———. *Deutsche Versgeschichte: mit Einschluss des altenglischen und altnordischen Stabreimverses*. 2nd ed. Vol. 2. Grundriss der germanischen Philologie. Berlin: W. De Gruyter, 1956.

Hirschmann, Rudolf. "A survey of computer-aided research in early German." *Computers and the Humanities* 8, no. 5 (1974): 279–284.

Hirschmann, Rudolf, and Winfried Lenders. "Computer-assisted study of early German: The Mannheim symposium of 1973." *Computers and the Humanities* 8, no. 3 (1974): 179–181.

Hooper, Joan B. "The Syllable in Phonological Theory." *Language* 48, no. 3 (September 1972): 525.

Isidore and Stephen A Barney. *The etymologies of Isidore of Seville*. Cambridge, UK; New York: Cambridge University Press, 2006.

Jeserich, Philipp, and Michael J. Curley. *Musica naturalis: speculative music theory and poetics, from Saint Augustine to the late Middle Ages in France*. Rethinking theory. Baltimore, Maryland: Johns Hopkins University Press, 2013.

Jespersen, Otto. *Lehrbuch der Phonetik;* in collaboration with Robarts - University of Toronto. Leipzig, Teubner, 1904.

Kahn, Daniel. "Syllable-based generalizations in English phonology." PhD diss., MIT, 1976.

Karg-Gasterstädt, Elisabeth. *Zur entstehungsgeschichte des Parzival*. Sächsische forschungsinstitute in Leipzig. Forschungsinstitut für neuere philologie. I. Altgermanistische abteilung ... [Veröffentlichungen] 2. Halle (Saale): M. Niemeyer, 1925.

Kemenaten, Albrecht von, and Julius Zupitza. "Virginal." In *Dietrichs Abenteuer*. Berlin: Weidmann, 1870.

Kohrs, Klaus Heinrich. "Zum Verhältnis von Sprache und Musik in den Liedern Neidharts von Reuental." *Deutsche Vierteljahrsschrift für Literaturwissenschaft und Geistesgeschichte* 43, no. 4 (1969): 604.

Kondrak, Garrett Nicolai Lei Yao Grzegorz. "Morphological Segmentation Can Improve Syllabification." *ACL 2016*, 2016, 99.

Kragl, Florian. "wort unde wîse. Formen des sangbaren Verses in der deutschen Literatur des Mittelalters." *Literaturwissenschaftliches Jahrbuch* 52 (2011): 31–80.

Kraus, Carl von. *Gottfried von Neifen*. Vol. Deutsche Liederdichter. Tübingen: M. Niemeyer.

Kuhn, Hugo. *Minnesangs Wende*. Vol. 1. Hermaea n.F. Tübingen: Niemeyer, 1967.

———. *Text und Theorie*. Vol. 2. His Kleine Schriften. Stuttgart: Metzler, 1969.

Kuhn, Hugo, and Christoph Cormeau. *Minnelieder Walthers von der Vogelweide: ein Kommentar*. Untersuchungen zur deutschen Literaturgeschichte 33. Tübingen: Niemeyer, 1982.

Lafferty, John D., Andrew McCallum, and Fernando C. N. Pereira. "Conditional Random Fields: Probabilistic Models for Segmenting and Labeling Sequence Data." In *Proceedings of the Eighteenth International Conference on Machine Learning*, 282–289. ICML '01. Morgan Kaufmann Publishers Inc., 2001.

Leitzmann, Albert, ed. *Wolfram von Eschenbach*. Altdeutsche Textbibliothek, Nr. 12-16. Tübingen: M. Niemeyer, 1960.

Levenshtein, V. I. "Binary Codes Capable of Correcting Deletions, Insertions and Reversals." *Soviet Physics Doklady* 10 (February 1, 1966): 707.

Longfellow, Henry Wadsworth. *The poems of Longfellow, including Evangeline, The song of Hiawatha, The courtship of Miles Standish, Tales of a wayside inn.* New York: B. A. Cerf, D. S. Klopfer, 1932.

Lord, Albert Bates, Stephen A. Mitchell, and Gregory Nagy. *The singer of tales*. 2nd ed. Harvard studies in comparative literature 24. Cambridge, Mass: Harvard University Press, 2000.

Maetschke, S. R., K. S. Kassahn, J. A. Dunn, S.-P. Han, E. Z. Curley, K. J. Stacey, and M. A. Ragan. "A visual framework for sequence analysis using n-grams and spectral rearrangement." *Bioinformatics* 26, no. 6 (March 15, 2010): 737–744.

Mari, Giovanni. *I trattati medievali di ritmica latina*. Vol. 11. U. Hoepli, 1899.

März, Christoph. "Der Silben Zall, der Chunsten Grunt. Die gezälte Silbe in Sangspruch und Meistergang." *Zeitschrift für deutsche Philologie* 119, no. 2000 (2000): 73–84.

März, Christoph. "Metrik, eine Wissenschaft zwischen Zählen und Schwärmen?" In *Mittelalter: neue Wege durch einen alten Kontinent*, edited by Jan-Dirk Müller and Horst Wenzel. Stuttgart: Hirzel, 1999.

Mathisen, Ralph W. *People, Personal Expression, and Social Relations in Late Antiquity*. University of Michigan Press, 2003.

Maurer, Friedrich. *Die politischen Lieder Walthers von der Vogelweide*. 2nd ed. Tübingen: Niemeyer, 1964.

McAleese, Gareth. "Improving Scansion with Syntax: an Investigation into the Effectiveness of a Syntactic Analysis of Poetry by Computer using Phonological Scansion Theory." -, 2007.

Mertens, Volker. "Was ist Rhythmus in Minnesang?" In *Aus dem Takt: Rhythmus in Kunst, Kultur und Natur*, 175–198. Bielefeld: Transcript, 2005.

Mittmann, Adiel, et al. "Escansão automática de versos em português," 2016.

Moretti, Franco. *Distant reading*. London: Verso, 2013.

———. *Graphs, maps, trees: abstract models for a literary history*. Verso, 2005.

———. "Network Theory, Plot Analysis." *New Left Review*, II, no. 68 (April 2011): 80–102.

———. ""Operationalizing": or, the function of measurement in modern literary theory." *Pamphlets of the Stanford Literary Lab*, no. 6 (2013).

———. "The End of the Beginning." *New Left Review*, 2nd ser., no. 41 (October 2006): 71–86.

Müller, Jan-Dirk. "Die Fiktion höfischer Liebe und die Fiktionalität des Minnesangs." In *Text und Handeln. Zum kommunikativen Ort von Minnesang und antiker Lyrik*, 47–64. 2004.

———. ""Ir sult sprechen willekomen": Sänger, Sprecherrolle und die Anfänge volkssprachlicher Lyrik." In *Minnesang und Literaturtheorie*, edited by Ute von Bloh and Armin Schulz, 107–128. Tübingen: Niemeyer, 2001.

Navarro, Borja. "A computational linguistic approach to Spanish Golden Age Sonnets: metrical and semantic aspects." In *Proceedings of the Fourth Workshop on Computational Linguistics for Literature*, 105–113. Denver, Colorado, USA: Association for Computational Linguistics, 2015.

Navarro-Colorado, Borja. "A metrical scansion system for fixed-metre Spanish poetry." *Digital Scholarship in the Humanities*, 2017, fqx009.

Nolan, M. "Medieval Sensation and Modern Aesthetics: Aquinas, Adorno, Chaucer." *the minnesota review* 2013, no. 80 (January 1, 2013): 145–158.

Norberg, Dag Ludvig, and Jan M. Ziolkowski. *An introduction to the study of medieval Latin versification*. Washington, D. C: Catholic University of America Press, 2004.

Obermaier, Sabine. "Der Dichter als Handwerker - Der Handwerker als Dichter. Autorkonzepte zwischen Sangspruchdichtung und Meistersang." *Zeitschrift für deutsche Philologie* 119, no. 2000 (2000): 59–72.

Okazaki, Naoaki. *CRFsuite: a fast implementation of Conditional Random Fields (CRFs). http://www.chokkan.org/software/crfsuite/*, 2007. `http://www.chokkan.org/software/crfsuite/`.

Palmer, Carloline, and Michael Kelly. "Linguistic prosody and musical meter in song." *Journal of Memory and Language* 31, no. 4 (August 1992): 525–542.

Parry, Milman, and Adam Milman. *The making of Homeric verse: the collected papers of Milman Parry*. Oxford: Clarendon Press, 1971.

Patel, Aniruddh D. *Music, language, and the brain*. New York: Oxford University Press, 2008.

Paul, Hermann, Hugo Moser, Ingeborg Schröbler, and Siegfried Grosse. *Mittelhochdeutsche Grammatik*. Tübingen: Max Niemeyer, 1982.

Pfaff, Fridrich, and Baden (Germany). *Die grosse Heidelberger Liederhandschrift, in getreuem textabdruck*. Heidelberg: C. Winter, 1898.

Plamondon, M. R. "Virtual Verse Analysis: Analysing Patterns in Poetry." *Literary and Linguistic Computing* 21 (Supplement 1 2006): 127–141.

Prendergast, Christopher. "Evolution and Literary History." *New Left Review*, 2nd ser., no. 34 (August 2005): 40–62.

Prince, Alan. "Quantitative consequences of rhythmic organization." *Chicago Linguistic Society* 26, no. 2 (1990): 355–398.

Ramus, F., M. Nespor, and J. Mehler. "Correlates of linguistic rhythm in the speech signal." *Cognition* 73, no. 3 (December 17, 1999): 265–292.

Ranke, Friedrich. *Sprache und stil im Wälschen gast des Thomasin von Circlaria*. Vol. 68. Mayer & Müller, 1908.

Rathay, Johann. "Über den Unterschied zwischen Lied und Spruch bei den Lyrikern des 12. und 13. Jahrhunderts." PhD diss., Leopoldstädter real- und obergymnasiums, 1875.

Rettelbach, Johannes. "Minnelied und Sangspruch: Formale Differenzen und Interferenzen bei der Tonkonstitution im 13. Jahrhundert." In *Sangspruchdichtung. Gattungskonstitution und Gattungsinterferenzen im europäischen Kontext*, 153–168. Berlin: De Gruyter, 2007.

Robertson, Stephen. "Understanding inverse document frequency: on theoretical arguments for IDF." *Journal of Documentation* 60, no. 5 (October 1, 2004): 503–520.

Roethe, Gustav. *Die Gedichte Reinmars von Zweter*. Nachdr. der Ausg. Leipzig, 1887. Amsterdam: Rodopi, 1967.

Rogova, Kseniya, Kris Demuynck, and Dirk Van Compernolle. "Automatic syllabification using segmental conditional random fields." *Comput. Linguist. Neth. J* 3 (2013): 34–48.

Ross, Deborah, Jonathan Choi, and Dale Purves. "Musical intervals in speech." *Proceedings of the National Academy of Sciences* 104, no. 23 (2007): 9852–9857.

Rousseau, Jean-Jacques, John T. Scott, and Jean-Jacques Rousseau. *Essay on the origin of languages and writings related to music*. Vol. 7. The collected writings of Rousseau. Hanover, N.H.: University Press of New England, 1998.

Runow, Holger. "Hât ieman sîn sô snellen... Rezeptionsbedingungen des Sangspruchs um 1300 zwischen Mündlichkeit und Schriftlichkeit." In *Sangspruchdichtung um 1300: Akten der Tagung in Basel vom 7. bis 9. November 2013*, edited by Gert Hübner, Dorothea Klein, and Universität Basel, 89–108. Spolia Berolinensia 33. Hildesheim: Weidmann, 2015.

Saussure, Ferdinand de, Charles Bally, Albert Sechehaye, and Albert Riedlinger. *Course in general linguistics*. LaSalle, Ill: Open Court, 1986.

Saville, Jonathan. *The Medieval Erotic Alba: Structure As Meaning*. iUniverse, 1999.

Schlager, Karlheinz. "Die Neumenschrift im Licht der Melismentextierung." *Archiv für Musikwissenschaft* 38, no. 4 (1981): 296.

Schmid, Hans, ed. *Musica et Scolica enchiriadis: una cum aliquibus tractatulis adiunctis*. Recensio nova post Gerbertinam altera ad fidem omnium codicum manuscriptorum. Veröffentlichungen der Musikhistorischen Kommission 3. München: Verlag der Bayerischen Akademie der Wissenschaften : In Kommission bei der C.H. Beck'schen Verlagsbuchhandlung, 1981.

Schneider, Hermann. "Mittelhochdeutsche Spruchdichtung." In *Mittelhochdeutsche Spruchdichtung*, 134–145. Wege der Forschung 154. Darmstadt, 1972.

Schnell, Rüdiger. "Prosaauflösung und Geschichtsschreibung im deutschen Spätmittelalter." In *Literatur und Laienbildung im Spätmittelalter und in der Reformationszeit: Symposion Wolfenbüttel 1981*, edited by Ludger Grenzmann, Karl Stackmann, and Deutsche Forschungsgemeinschaft, 215–248. Germanistische Symposien, Berichtsbände 5. Stuttgart: J.B. Metzler, 1984.

Schwarz, Werner. "Notes on formulaic expressions in MHG poetry." In *Beiträge zur mittelalterlichen Literatur*, 14–24. 1984.

Schweikle, Günther. “Heinrich von Stretelingen.” In *Die deutsche Literatur des Mittelalters, Verfasserlexikon*, Zweite, völlig neu bearbeitete Auflage, vol. 3. Berlin ; New York: W. de Gruyter, 1978.

———. *Minnesang*. Vol. 244. Metzler, 1995.

Shannon, C. E. “A mathematical theory of communication.” *The Bell System Technical Journal* 27, no. 3 (July 1948): 379–423.

Sharma, Rama Nath, and Pāṇini. *The Aṣṭādhyāyī of Pāṇini*. New Delhi: Munshiram Manoharlal Publishers, 1987.

Simrock, Karl Joseph, and Wilhelm Wackernagel. *Gedichte Walthers von der Vogelweide*. Berlin: Vereinsbuchhandlung, 1833.

Singhal, Amit, Chris Buckley, and Mandar Mitra. “Pivoted Document Length Normalization.” In *Proceedings of the 19th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval*, 21–29. SIGIR ’96. New York, NY, USA: ACM, 1996.

Sparck Jones, Karen. “A statistical interpretation of term specificity and its application in retrieval.” *Journal of Documentation* 28, no. 1 (January 1, 1972): 11–21.

Springeth, Margerete, Nikolaus Morocutti, and Daniel Schlager. *Mittelhochdeutsche Begriffsdatenbank (MHDBDB). Universität Salzburg.* `http://www.mhdbdb.sbg.ac.at/`. Accessed: 2016-10-01.

Stock, Markus. “Das volle Wort - Sprachklang im späteren Minnesang.” In *Text und Handeln: zum kommunikativen Ort von Minnesang und antiker Lyrik*, edited by Albrecht Hausmann, Cornelia Logemann, and Christian Rode, 65–105. Beihefte zum Euphorion 46. Heidelberg: Winter, 2004.

Strohschneider, Peter. “”nu sehent, wie der singet!”: vom Hervortreten des Sängers im Minnesang.” In *”Aufführung” und ”Schrift” in Mittelalter und früher Neuzeit*, 7–30. 1996.

Tervooren, Helmut. *Minimalmetrik zur Arbeit mit mittelhochdeutschen Texten*. Göppingen: Kümmerle Verlag, 1997.

———. “”Spruch” und ”Lied”.” In *Schoeniu wort mit süezeme sange: philologische Schriften*, edited by Susanne Fritsch-Staar and Johannes Spicker. Philologische Studien und Quellen 159. Berlin: E. Schmidt, 2000.

Thomasin and F. W. von Kries. *Der Welsche Gast*. Göppinger Arbeiten zur Germanistik 425. Göppingen: Kümmerle-Verlag, 1984.

Thomasin and Eva Willms. *Der Welsche Gast*. De Gruyter Texte. Berlin ; New York: W. De Gruyter, 2004.

Tischler, Hans. "Rhythm, Meter, and Melodic Organization in Medieval Songs." *Revue belge de Musicologie / Belgisch Tijdschrift voor Muziekwetenschap* 28 (1974): 5.

Tomović, Andrija, Predrag Janičić, and Vlado Kešelj. "N-Gram-based classification and unsupervised hierarchical clustering of genome sequences." *Computer methods and programs in biomedicine* 81, no. 2 (2006): 137–153.

Treitler, Leo. "Homer and Gregory: the transmission of epic poetry and plainchant." *The Musical Quarterly* 60, no. 3 (1974): 333–372.

———. "The "Unwritten" and "Written Transmission" of Medieval Chant and the Start-Up of Musical Notation." *The Journal of Musicology* 10, no. 2 (1992): 131–191.

Vennemann, Theo. "On the Theory of Syllabic Phonology." *Linguistische Berichte* 18 (1972): 1–18.

Volkovich, Z., V. Kirzhner, A. Bolshoy, E. Nevo, and A. Korol. "The method of N-grams in large-scale clustering of DNA texts." *Pattern Recognition* 38, no. 11 (November 2005): 1902–1912.

Warning, Rainer. "Lyrisches Ich und Öffentlichkeit bei den Trobadors." In *FS Hugo Kuhn (1979)*, 120–159. 1979.

Wenzel, Horst. *Hören und Sehen, Schrift und Bild: Kultur und Gedächtnis im Mittelalter*. 1st ed. München: C.H.Beck, February 27, 1995.

Wiese, Richard. "The Prosodic Structure of German." In *The phonology of German*. The phonology of the world's languages. New York: Oxford University Press, 2000.

Wolfram, Joachim Heinzle, Dorothea Diemer, and Peter Diemer. *Willehalm*. Deutscher Klassiker-Verlag im Taschenbuch 39. Frankfurt am Main: Dt. Klassiker-Verl, 2009.

Ziegler, Vickie L. *The leitword in Minnesang: stylistic analysis and textual criticism*. The Penn State series in German literature. University Park: Pennsylvania State University Press, 1975.

Zumthor, Paul. *Die Stimme und die Poesie in der mittelalterlichen Gesellschaft*. München: Fink, 1994.

Zumthor, Paul, and Marilyn C. Engelhardt. "The text and the voice." *New literary history* 16, no. 1 (1984): 67–92.

Appendix A

LegaliPy

```
from __future__ import unicode_literals # for python2 compatibility
# -*- coding: utf-8 -*-
# created at UC Berkeley 2015
# Authors: Christopher Hench & Alex Estes © 2016
# ================================================================================

'''This program syllabifies any text in any language
solely on the Onset Maximization principle (Principle of Legality).
Input is text file'''

import codecs
import re
import csv
import sys
from datetime import datetime
import string
from collections import Counter


def cleantext(text):
    text = ''.join([x for x in text if not x.isdigit()])

    remove_char = string.punctuation + ' 〈〉•–»«¿·'
    for char in text:
        if char in remove_char:
            text = text.replace(char, '')

    return (text.split()) # return list of words, alt tokenize
```

```
def getonsets(text):

    vowels = 'āēėęīįœōūůaeiouyàáâäæãåèéêëîïíìôöòóøõûüùúÿ'

    tokens = cleantext(text)

    onsets = []
    for word in tokens:
        word = word.lower()
        onset = ""
        for letter in word:
            if letter not in vowels: # onset is everying up to first vowel
                onset += letter
            else:
                break
        onsets.append(onset)

    onsets = [x for x in onsets if x != ''] # get rid of empty onsets

    # now remove onsets caused by errors, i.e. less than .02% of onsets
    freq = Counter(onsets)

    total_onsets = 0
    for k, v in freq.items():
        total_onsets += v

    onsets = []
    for k, v in freq.items():
        if (v / total_onsets) > .0002:
            onsets.append(k)

    return (onsets, tokens)


def legalipy(word, onsets):

    longest_onset = len(max(onsets, key=len))
    vowels = 'āēėęīįœōūůaeiouyàáâäæãåèéêëîïíìôöòóøõûüùúÿ'
    vowelcount = 0
    revword = word[::-1] # reverse word to build onsets from back

    syllset = []
    for letter in revword:
```

```
        if letter.lower() in vowels:
            vowelcount += 1
        else:
            pass

    if vowelcount == 1: # monosyllabic
        syllset.append(revword)

    # begin main algorithm
    elif vowelcount > 1:
        syll = ""

        # following binaries trigger different routes
        onsetbinary = 0
        newsyllbinary = 1

        for letter in revword:

            if newsyllbinary == 1: # i.e. if we have a new syllable
                if letter.lower() not in vowels:
                    syll += letter

                else:
                    syll += letter
                    newsyllbinary = 0
                    continue

            elif newsyllbinary == 0: # i.e. no longer new syllable

                if syll == "": # fixes last syllable
                    syll += letter

                elif (letter.lower() in onsets and syll[-1].lower() in vowels):
                    syll += letter
                    onsetbinary = 1

                elif (letter.lower() + syll[-1].lower() in [ons[-2:] for ons in
                    onsets] and syll[-2].lower() in vowels):
                    syll += letter
                    onsetbinary = 1

                elif (letter.lower() + syll[-2:][::-1].lower() in [ons[-3:] for ons
                    in onsets] and syll[-3].lower() in vowels):
                    syll += letter
```

```
                onsetbinary = 1

            elif (letter + syll[-3:][::-1].lower() in [ons[-4:] for ons in
                onsets] and syll[-4].lower() in vowels):
                syll += letter
                onsetbinary = 1

            # order is important for following two due to onsetbinary
            # variable
            # i.e. syllable doesn't end in vowel (onset not yet found)
            elif letter.lower() in vowels and onsetbinary == 0:
                syll += letter

            # i.e. syllable ends in vowel, onset found, restart syllable
            elif letter.lower() in vowels and onsetbinary == 1:
                syllset.append(syll)
                syll = letter

            else:
                syllset.append(syll)
                syll = letter
                newsyllbinary = 1

        syllset.append(syll)

    # reverse syllset then reverse syllables
    syllset = [syll[::-1] for syll in syllset][::-1]

    return (syllset)

# MAIN PROGRAM HERE
if __name__ == '__main__':

    print("\n\nLegaliPy-ing...\n")

    sfile = sys.argv[1] # input text file to syllabify
    with open(sfile, 'r', encoding='utf-8') as f:
        text = f.read()

    onsets = getonsets(text)

    toprintl = []
    for token in onsets[1]:
        toprintl.append(legalipy(token, onsets[0]))
```

```
toprint = ""
for word in toprintl:
    for syll in word:
        if syll != word[-1]:
            toprint += syll
            toprint += "-"
        else:
            toprint += syll
    toprint += " "

onsetprint = (" - ".join([x for x in onsets[0]]) + '\n\n')

prologue = "Following onsets > .02 percent deemed 'legal':\n"

fmt = '%Y/%m/%d %H:%M:%S'
date = "LegaliPyed on " + str(datetime.now().strftime(fmt))

finalwrite = date + "\n\n" + prologue + onsetprint + toprint

with open('LegaliPyed.txt', 'w', encoding='utf-8') as f:
    f.write(finalwrite)

print("\nResults saved to LegaliPyed.txt\n\n")
```

Appendix B

SonoriPy

```
from __future__ import unicode_literals # for python2 compatibility
# -*- coding: utf-8 -*-
# created at UC Berkeley 2015
# Authors: Christopher Hench, Alex Estes

'''This program syllabifies words based on the Sonority Sequencing Principle
    (SSP)'''

import codecs


def sonoripy(word):

    def no_syll_no_vowel(ss):
        # no syllable if no vowel
        nss = []
        front = ""
        for i, syll in enumerate(ss):
            # if following syllable doesn't have vowel,
            # add it to the current one
            if not any(char in vowels for char in syll):
                if len(nss) == 0:
                    front += syll
                else:
                    nss = nss[:-1] + [nss[-1] + syll]
            else:
                if len(nss) == 0:
                    nss.append(front + syll)
                else:
```

```
                nss.append(syll)

    return nss

# SONORITY HIERARCHY, MODIFY FOR LANGUAGE BELOW
# categories can be collapsed into more general groups
vowels = 'aeiouy'
approximates = ''
nasals = 'lmnrw' # resonants and nasals
fricatives = 'zvsf'
affricates = ''
stops = 'bcdgtkpqxhj' # rest of consonants

vowelcount = 0 # if vowel count is 1, syllable is automatically 1
sylset = [] # to collect letters and corresponding values
for letter in word.strip(".:;?!)('" + '"'):
    if letter.lower() in vowels:
        sylset.append((letter, 5))
        vowelcount += 1 # to check for monosyllabic words
    elif letter.lower() in approximates:
        sylset.append((letter, 4))
    elif letter.lower() in nasals:
        sylset.append((letter, 3))
    elif letter.lower() in fricatives:
        sylset.append((letter, 2))
    elif letter.lower() in affricates:
        sylset.append((letter, 1))
    elif letter.lower() in stops:
        sylset.append((letter, 0))
    else:
        sylset.append((letter, 0))

# below actually divides the syllables
newsylset = []
if vowelcount == 1: # finalize word immediately if monosyllabic
    newsylset.append(word)
if vowelcount != 1:
    syllable = '' # prepare empty syllable to build upon
    for i, tup in enumerate(sylset):
        if i == 0: # if it's the first letter, append automatically
            syllable += tup[0]
        # lengths below are in order to not overshoot index
        # when it looks beyond
        else:
```

```
# add whatever is left at end of word, last letter
if i == len(sylset) - 1:
    syllable += tup[0]
    newsylset.append(syllable)

# MAIN ALGORITHM BELOW
# these cases DO NOT trigger syllable breaks
elif (i < len(sylset) - 1) and tup[1] < sylset[i + 1][1] and \
        tup[1] > sylset[i - 1][1]:
    syllable += tup[0]
elif (i < len(sylset) - 1) and tup[1] > sylset[i + 1][1] and \
        tup[1] < sylset[i - 1][1]:
    syllable += tup[0]
elif (i < len(sylset) - 1) and tup[1] > sylset[i + 1][1] and \
        tup[1] > sylset[i - 1][1]:
    syllable += tup[0]
elif (i < len(sylset) - 1) and tup[1] > sylset[i + 1][1] and \
        tup[1] == sylset[i - 1][1]:
    syllable += tup[0]
elif (i < len(sylset) - 1) and tup[1] == sylset[i + 1][1] and \
        tup[1] > sylset[i - 1][1]:
    syllable += tup[0]
elif (i < len(sylset) - 1) and tup[1] < sylset[i + 1][1] and \
        tup[1] == sylset[i - 1][1]:
    syllable += tup[0]

# these cases DO trigger syllable break
# if phoneme value is equal to value of preceding AND following
# phoneme
elif (i < len(sylset) - 1) and tup[1] == sylset[i + 1][1] and \
        tup[1] == sylset[i - 1][1]:
    syllable += tup[0]
    # append and break syllable BEFORE appending letter at
    # index in new syllable
    newsylset.append(syllable)
    syllable = ""

# if phoneme value is less than preceding AND following value
# (trough)
elif (i < len(sylset) - 1) and tup[1] < sylset[i + 1][1] and \
        tup[1] < sylset[i - 1][1]:
    # append and break syllable BEFORE appending letter at
    # index in new syllable
    newsylset.append(syllable)
```

```
                syllable = ""
                syllable += tup[0]

            # if phoneme value is less than preceding value AND equal to
            # following value
            elif (i < len(sylset) - 1) and tup[1] == sylset[i + 1][1] and \
                    tup[1] < sylset[i - 1][1]:
                syllable += tup[0]
                # append and break syllable BEFORE appending letter at
                # index in new syllable
                newsylset.append(syllable)
                syllable = ""

    newsylset = no_syll_no_vowel(newsylset)

    return (newsylset)
```

Appendix C

SyllabiPy MHG

```
from __future__ import unicode_literals # for python2 compatibility
# -*- coding: utf-8 -*-
# created at UC Berkeley 2015
# Authors: Christopher Hench, Alex Estes

'''This program contains a function 'syllabipymhg' that syllbifies
Middle High German words for further analysis.
Input is string (word), output is list of strings with syllables.'''

import codecs
import re
import csv # for generating exception file


def syllabipymhg(word):

    def no_syll_no_vowel(ss):
        # no syllable if no vowel
        nss = []
        front = ""
        for i, syll in enumerate(ss):
            # if following syllable doesn't have vowel,
            # add it to the current one
            if not any(char in vowels for char in syll):
                if len(nss) == 0:
                    front += syll
                else:
                    nss = nss[:-1] + [nss[-1] + syll]
            else:
```

```
            if len(nss) == 0:
                nss.append(front + syll)
            else:
                nss.append(syll)

    return nss

# strip extra punctuation and lower case word
from string import punctuation
for c in punctuation:
    word = word.replace(c, "")

word = word.lower()

# THIS SECTION PREPARES ORTHOGRAPHY AND ASSIGNS VALUE

# list of basic sounds
longvowels = "āēīœōūâæêîôû"
vowels = 'āēėęīįœōūůaeiouyàáâäæãåèéêëîïíìôöòóøõûüùúýÿ' # includes long vowels
resonants = "lmnrw"
consonants = "łžbcdgtkpqvxhçsfzj"

# replace single phonemes represented by
# multiple letters with single letter
# THIS IS UNIQUE TO MIDDLE HIGH GERMAN
if "sch" in word:
    word = word.replace("sch", "ç")
if "ch" in word:
    word = word.replace("ch", "ł")
if "ph" in word:
    word = word.replace("ph", "ž")

vowelcount = 0 # if vowel count is 1, syllable is automatically 1
sylset = [] # to collect letters and corresponding values

# cycle through each letter and assign value in SSP (Sonority Sequencing
# Principle) hierarchy creating list of tuples in sylset
for letter in word:
    if letter in vowels:
        sylset.append((letter, 3))
        vowelcount += 1 # to check for monosyllabic words
    if letter in resonants:
        sylset.append((letter, 2))
    if letter in consonants:
```

```
        sylset.append((letter, 1))

# THIS SECTION CREATES SYLLABLE BOUNDARIES

newsylset = []
if vowelcount == 1: # finalize word immediately if monosyllabic
    newsylset.append(word)
if vowelcount != 1:
    syllable = '' # prepare empty syllable to build upon
    for i, tup in enumerate(sylset):
        if i == 0: # if it's the first letter, append automatically
            syllable += tup[0]
        # lengths below are in order to not overshoot index
        # when it looks beyond
        else:
            # add whatever is left at end of word, last letter
            if i == len(sylset) - 1:
                syllable += tup[0]
                newsylset.append(syllable)

            # accounts for ge prefix in MHG
            elif (i < len(sylset) - 1) and syllable == "ge" and \
                    tup[0] in vowels:
                # gei and geu are accepted dipthongs in MHG
                if tup[0] != "i" and tup[0] != "u":
                    # append and break syllable BEFORE appending letter at
                    # index in new syllable
                    newsylset.append(syllable)
                    syllable = ""
                    syllable += tup[0]
                else:
                    syllable += tup[0] # accepting dipthongs

            # breaks syllable on vowels followed by long vowels
            elif (i < len(sylset) - 1) and len(syllable) > 0 and \
                    syllable[-1] in longvowels and tup[0] in vowels:
                # append and break syllable BEFORE appending letter at
                # index in new syllable
                newsylset.append(syllable)
                syllable = ""
                syllable += tup[0]

            # MAIN ALGORITHM BELOW
            # these cases do not trigger syllable breakS
```

```
elif (i < len(sylset) - 1) and tup[1] < sylset[i + 1][1] and \
       tup[1] > sylset[i - 1][1]:
   syllable += tup[0]
elif (i < len(sylset) - 1) and tup[1] > sylset[i + 1][1] and \
       tup[1] < sylset[i - 1][1]:
   syllable += tup[0]
elif (i < len(sylset) - 1) and tup[1] > sylset[i + 1][1] and \
       tup[1] > sylset[i - 1][1]:
   syllable += tup[0]
elif (i < len(sylset) - 1) and tup[1] > sylset[i + 1][1] and \
       tup[1] == sylset[i - 1][1]:
   syllable += tup[0]
elif (i < len(sylset) - 1) and tup[1] == sylset[i + 1][1] and \
       tup[1] > sylset[i - 1][1]:
   syllable += tup[0]
elif (i < len(sylset) - 1) and tup[1] < sylset[i + 1][1] and \
       tup[1] == sylset[i - 1][1]:
   syllable += tup[0]

# these cases DO trigger syllable break
# if phoneme value is equal to value of preceding AND following
# phoneme
elif (i < len(sylset) - 1) and tup[1] == sylset[i + 1][1] and \
       tup[1] == sylset[i - 1][1]:
   syllable += tup[0]
   # append and break syllable BEFORE appending letter at
   # index in new syllable
   newsylset.append(syllable)
   syllable = ""

# if phoneme value is less than preceding AND following value
# (trough)
elif (i < len(sylset) - 1) and tup[1] < sylset[i + 1][1] and \
       tup[1] < sylset[i - 1][1]:
   # append and break syllable BEFORE appending letter at
   # index in new syllable
   newsylset.append(syllable)
   syllable = ""
   syllable += tup[0]

# if phoneme value is less than preceding value AND equal to
# following value
elif (i < len(sylset) - 1) and tup[1] == sylset[i + 1][1] and \
       tup[1] < sylset[i - 1][1]:
```

```
                syllable += tup[0]
                # append and break syllable BEFORE appending letter at
                # index in new syllable
                newsylset.append(syllable)
                syllable = ""

    # THIS SECTION RETURNS ORTHOGRAPHY
    # replace characters treated as one phoneme
    newsylset2 = []
    for syll in newsylset:
        if "ç" in syll:
            syll = syll.replace("ç", "sch")
        if "ł" in syll:
            syll = syll.replace("ł", "ch")
        if "ž" in syll:
            syll = syll.replace("ž", "ph")
        newsylset2.append(syll)

    newsylset = no_syll_no_vowel(newsylset2)

    # read csv of corrections from list generated for compound issues
    # list is generated with onset.py
    tofix = []
    with open("/Users/chench/Box
        Sync/Hench_Dissertation/diss/code/scripts/corrections.csv",
        encoding="utf-8") as f:
        data = [tuple(line)
                for line in csv.reader(f)] # yields tuples with len 4
        for tup in data: # change asterisks to blanks in the tuple
            if tup[0] == "*":
                tofix.append(("", tup[1], tup[2], tup[3]))
            elif tup[1] == "*":
                tofix.append((tup[0], "", tup[2], tup[3]))
            elif tup[2] == "*":
                tofix.append((tup[0], tup[1], "", tup[3]))
            elif tup[3] == "*":
                tofix.append((tup[0], tup[1], tup[2], ""))
            else:
                tofix.append(tup)

    for correction in tofix:
        newsylset2 = []
        for i, syll in enumerate(newsylset):
            # to break up two characters in following syllable, first string of
```

```
            # tuple is blank
            if len(correction[0]) == 0:
                if (i < len(newsylset) - 1) and \
                    newsylset[i + 1][:(len(correction[1]))] == \
                        correction[1]:
                    syll = syll + correction[2]
                    newsylset2.append(syll)
                elif i > 0 and syll[:len(correction[1])] == correction[1]:
                    syll = syll[len(correction[2]):]
                    newsylset2.append(syll)
                else:
                    newsylset2.append(syll)

            # to switch characters between words, first string of tuple not
            # blank
            elif len(correction[0]) > 0:
                if (i < len(newsylset) - 1) and syll[-len(correction[0]):] == \
                        correction[0] and \
                        newsylset[i + 1][:len(correction[1])] == \
                        correction[1]:
                    syll = syll[:-len(correction[0])] + correction[2]
                    newsylset2.append(syll)
                elif i > 0 and newsylset[i - 1][-len(correction[0]):] == \
                        correction[0] and syll[:len(correction[1])] == \
                        correction[1]:
                    syll = correction[3] + syll[len(correction[1]):]
                    newsylset2.append(syll)
                else:
                    newsylset2.append(syll)

            # to break up two characters in current syllable, necessary?
            elif len(correction[1]) == 0:
                if (i < len(newsylset) - 1) and syll[-len(correction[0]):] == \
                        correction[0]:
                    syll = syll[:-len(correction[0])] + correction[2]
                    newsylset2.append(syll)
                elif i > 0 and newsylset[i - 1][-len(correction[0]):] == \
                        correction[0]:
                    syll = syll[len(correction[2]):]
                    newsylset2.append(syll)
                else:
                    newsylset2.append(syll)

        # use new sylset created after each fix, so multiple fixes can be used
```

```
        # on same set
        newsylset = newsylset2

    # handling ch intervocalically c-h
    newsylset2 = []
    for i, syll in enumerate(newsylset):
        if (i < len(newsylset) - 1) and syll[-1] in vowels and \
                newsylset[i + 1][:2] == 'ch' and newsylset[i + 1][2] in vowels:
            syll = syll + 'c'
            newsylset2.append(syll)
        elif (i > 0) and newsylset[i - 1][-1] in vowels and syll[:2] == \
                'ch' and syll[2] in vowels:
            syll = syll[1:]
            newsylset2.append(syll)
        else:
            newsylset2.append(syll)

    newsylset = newsylset2

    # handling sch intervocalically

    newsylset2 = []
    for i, syll in enumerate(newsylset):
        if (i < len(newsylset) - 1) and len(syll) > 2 and syll[-1] in vowels \
            and syll[-2] != 'g' and len(newsylset[i + 1]) > 3 and \
            newsylset[i + 1][:3] == 'sch' and \
                newsylset[i + 1][3] in vowels:
            syll = syll + 's'
            newsylset2.append(syll)
        elif (i > 0) and len(newsylset[i - 1]) > 3 and newsylset[i - 1][-1] \
                in vowels and newsylset[i - 1][-2] != 'g' and len(syll) > 2 \
                and syll[:3] == 'sch' and syll[3] in vowels:
            syll = syll[1:]
            newsylset2.append(syll)
        else:
            newsylset2.append(syll)

    newsylset = newsylset2

    # handling long vowel + short vowel (breaking up)
    newsylset2 = []
    for i, syll in enumerate(newsylset):
        if len(syll) >= 2 and syll[-2] in longvowels and syll[-1] in vowels:
            syll = syll[:-1]
```

```
            newsylset2.append(syll)
        elif (i > 0) and len(newsylset[i - 1]) >= 2 and newsylset[i - 1][-2] \
                in longvowels and newsylset[i - 1][-1] in vowels:
            syll = newsylset[i - 1][-1] + syll
            newsylset2.append(syll)
        else:
            newsylset2.append(syll)

    newsylset = newsylset2

    # handling 'lich'
    lichreg = re.compile(r"l(i|î)ch")
    licreg = re.compile(r"l(i|î)c")

    if re.search(lichreg, ''.join(newsylset)):
        fix = True
        for i, syll in enumerate(newsylset):
            if licreg.match(syll[:3]):
                fix = False
                lic_ind = False
            elif re.search(licreg, syll):
                lic_ind = i
                lic_str_ind = re.search(licreg, syll).start()
            else:
                fix = False
                lic_ind = False

        if fix and lic_ind:
            newsylset2 = newsylset[:lic_ind - 1]
            newsylset2.append(newsylset[lic_ind - 1] +
                              newsylset[lic_ind][:lic_str_ind])
            newsylset2.append(newsylset[lic_ind][lic_str_ind:])
            newsylset = newsylset2 + newsylset[lic_ind + 1:]

    # handling 'heit' at end of word
    newsylset2 = []
    for i, syll in enumerate(newsylset):
        if (i < len(newsylset) - 1) and len(newsylset[i + 1]) > 4 and \
                newsylset[i + 1][-4:] == "heit":
            syll = syll + newsylset[i + 1][:-4]
            newsylset2.append(syll)
        elif syll[-4:] == "heit":
            syll = syll[-4:]
            newsylset2.append(syll)
```

```
        else:
            newsylset2.append(syll)

    newsylset = no_syll_no_vowel(newsylset2)

    return (newsylset)
```

Appendix D

Lyric Texts in Chapter 3

texts
Albrecht von Johansdorf (Minnesangs Frühling)
Albrecht von Raprechtswil
Bernger von Horheim (Minnesangs Frühling)
Bligger von Steinach (Minnesangs Frühling)
Boppe
Bruno von Hornberg
Brunwart von Augheim
Burggraf von Regensburg (Minnesangs Frühling)
Burggraf von Rietenburg (Minnesangs Frühling)
Burkart von Hohenfels
Christan von Hamle
Christan von Luppin
Der Burggraf von Lienz
Der Düring
Der Dürner
Der Jenaer Meißner
Der Junge Meißner
Der Kanzler
Der Kol von Niunzen
Der Markgraf von Hohenburg
Der Marner Teil II (Sangsprüche)
Der Marner Teil III (Meisterlieder)
Der Püller
Der Schenk von Limburg
Der Schulmeister von Esslingen
Der Taler

Der grave von Anhalte
Der tugendhafte Schreiber
Der von Gliers
Der von Kürenberg (Minnesangs Frühling)
Der von Trostberg
Der von Wengen
Der wilde Alexander
Dietmar der Setzer
Dietmar von Eist (Minnesangs Frühling)
Eberhard von Sax
Engelhart von Adelnburg (Minnesangs Frühling)
Frauendienst Lieder (Bechstein)
Frauenlob Teil I (Leiche)
Frauenlob Teil II (Lieder)
Frauenlob Teil III (Supplementband)
Friderich von Liningen
Friedrich der Knecht
Friedrich von Hausen (Minnesangs Frühling)
Gast
Gedrut — Geltar
Goeli
Goesli von Ehenhein
Gottfried von Neifen
Gottfried von Strassburg
Gottfried von Straßburg (Minnesangs Frühling)
Günther von dem Vorste
Hartmann von Aue (Minnesangs Frühling)
Hartmann von Starkenberg
Hartwig von Rute (Minnesangs Frühling)
Hawart
Heinrich Hetzbolt von Wissense
Heinrich Teschler
Heinrich der Rost
Heinrich von Frauenberg
Heinrich von Morungen (Minnesangs Frühling)
Heinrich von Rugge (Minnesangs Frühling)
Heinrich von Sax
Heinrich von Stretelingen
Heinrich von Tettingen
Heinrich von Veldeke (Minnesangs Frühling)
Heinrich von Veldeke (limburg. Fassg.) (Minnes...
Heinrich von der Muore

Herger (Minnesangs Frühling)
Herrand von Wildonie Lieder
Herzog Heinrich von Pressela
Hesso von Rinach
Hiltbolt von Schwangau
Hugo von Montfort
Hugo von Mülndorf
Hugo von Werbenwag
Jakob von Warte
Johann von Ringgenberg
Johannes Hadlaub
Kaiser Heinrich (Minnesangs Frühling)
Konrad von Altstetten
Konrad von Kilchberg
Konrad von Landeck
Konrad von Würzburg Lyrik
Kraft von Toggenburg
König Konrad der Junge
Leuthold von Seven
Markgraf Heinrich von Meißen
Meinloh von Sevelingen (Minnesangs Frühling)
Namenlos
Namenlose Lieder (Minnesangs Frühling)
Neidhart-Lieder Hs. c (Bennewitz)
Niune
Otto von Bottenlouben
Otto von Brandenburg mit dem Pfeil
Otto zum Turm
Pfeffel
Reimar der Fiedler
Reimar der Junge
Reimar von Brenneberg
Reinmar (Minnesangs Frühling)
Rubin
Rubin und Rüdeger
Rudolf der Schreiber
Rudolf von Fenis (Minnesangs Frühling)
Rudolf von Rotenburg
Rumelant v. Sachsen
Spervogel (Minnesangs Frühling)
Steinmar
Süezkint

Süezkint der Jude von Trimperg
Tannhäuser
Ulrich von Baumburg
Ulrich von Gutenburg (Minnesangs Frühling)
Ulrich von Liechtenstein
Ulrich von Singenberg
Ulrich von Winterstetten Leich
Ulrich von Winterstetten Lieder
Von Munegiur
Von Obernburg
Von Sachsendorf
Von Scharfenberg
Von Stadeck
Von Stamheim
Von Suoneck
Von Wissenlo
Wahsmuot von Kunzich
Wahsmuot von Mülnhusen
Walter von Mezze
Walther von Klingen
Walther von Prisach
Walther: Lieder und Sangsprüche
Waltram von Gresten
Wenzel von Beheim
Wernher von Hohenberg
Wernher von Teufen
Willehelm von Heinzenburg
Winli
Wolfram Lieder
Wolfram von Eschenbach (Minnesangs Frühling)
von Buchheim

Table D.1: Lyric texts in Chapter 3

Appendix E

Verse Texts in Chapter 4

texts
Aalener Stadtratsgedicht
Abriss der böhmischen Geschichte
Abriss der böhmischen Geschichte (Kopie)
Albrecht von Johansdorf (Minnesangs Frühling)
Albrecht von Raprechtswil
Alexander (R. v. E.) (Rudolf von Ems)
Alexander (U.v.E) (Ulrich von Eschenbach)
Alexander Anhang
Alexius
Alexius (Fassung K)
Allerheiligenlitanei
Alpharts Tod
Aristoteles und Phyllis
Armer und reicher König
Barlaam und Josaphat
Basler Alexander (Einleitung)
Begrabener Ehemann
Beispiel Salomons
Bernger von Horheim (Minnesangs Frühling)
Biterolf und Dietleib
Bitte um mildes Gericht
Bligger von Steinach (Minnesangs Frühling)
Boppe
Bremse im Blütenhaus
Bruder Wernher
Bruno von Hornberg

Brunwart von Augheim
Buch von Akkon (Steirische Reimchronik, Aussch...
Buch von Troja
Burggraf von Regensburg (Minnesangs Frühling)
Burggraf von Rietenburg (Minnesangs Frühling)
Burkart von Hohenfels
Bußgebet
Christan von Hamle
Christan von Luppin
Christi Hort
Christliche Lehre
Christus eine gebärende Frau
Commemoratio pro defunctis et vivis
Daniel von dem blühenden Tal
Das Annolied
Das Beispiel vom Hasen
Das Bild
Das Buch der Natur (Buch 1)
Das Gebot der Ehre
Das Glück im Traum
Das Gänslein
Das Häslein
Das Katzenauge
Das Marienbild in Konstantinopel
Das Narrenschiff
Das Nonnenturnier
Das Rolandslied
Das Rädlein
Das Schneekind (Hs.A)
Das Schneekind (Hs.B)
Das Säcklein Witz
Das Turnier von Nantes
Das Wildpret
Das entweihte Gotteshaus
Das geschändete Sakrament
Das heisse Eisen
Das heiße Eisen
Das stinkende Haus
Das wilde Roß
Der Abgott der Juden
Der Bergmann
Der Bildschnitzer von Würzburg

Der Burggraf von Lienz
Der Düring
Der Dürner
Der Esel
Der Gast und die Wirtin
Der Gärtner
Der Hahn und die Perle
Der Hase
Der Heilige Ulrich
Der Heller der armen Frau
Der Hofhund
Der Hort
Der Hund am Wasser
Der Hund und der Stein
Der Jenaer Meißner
Der Jesusknabe als Geisel (Die Witwe und ihr S...
Der Juden Abgott
Der Judenknabe
Der Junge Meißner
Der Jüngere Sigenot
Der Jüngere Titurel
Der Jüngling
Der Kanzler
Der Kater als Freier
Der Kirchtag
Der Knecht in Herrenkleidern
Der Koch
Der Kol von Niunzen
Der Krämer
Der Kummer
Der Käfer im Rosenhaus
Der Markgraf von Hohenburg
Der Marktdieb
Der Marner Teil I (Lieder)
Der Marner Teil II (Sangsprüche)
Der Marner Teil III (Meisterlieder)
Der Mönch als Liebesbote (A)
Der Münchner Oswald
Der Ochse und die Maus
Der Pfaffe im Käsekorb
Der Pfaffe mit der Schnur
Der Pfaffen Leben

Der Püller
Der Rabe mit den Pfauenfedern
Der Ratgeber
Der Renner
Der Richter und der Teufel
Der Ring (Bechstein)
Der Ring (Wiesner)
Der Ritter unter dem Zuber
Der Ritter von Staufenberg
Der Ritterspiegel
Der Roßtäuscher
Der Salamander
Der Schalk und die beiden Könige
Der Schenk von Limburg
Der Schlegel
Der Scholar und das Marienbild
Der Schulmeister von Esslingen
Der Schwanritter
Der Schüler von Paris
Der Sperber
Der Spiegel
Der Stein philosophorum
Der Sünder und der Einsiedel
Der Taler
Der Taugenichts
Der Teufel als Kämmerer
Der Teufel und die Seele
Der Tor und Feuer
Der Tor und das Feuer
Der Traum des Scholaren (Der Schüler aus Sizil...
Der Traum des Sünders
Der Trojanische Krieg
Der Tropfen auf dem Stein
Der Tugendspiegel oder der Meizoge
Der Turse
Der Vogel und der Sperber
Der Waldschrat
Der Weidemann
Der Welsche Gast
Der Welt Lohn
Der Wernigeroder Alexander
Der Wolf als Fischer

Der Wolf und das Weib
Der Wolf und der Bauer
Der Wolf und der Biber
Der Wolf und der Hund
Der Wolf und die Geige
Der Wolf und die Gänse
Der Wolf und sein Sohn
Der Wucherer
Der Wunderer
Der Wälsche Gast (Inhaltsangabe)
Der altgewordene Sünder
Der arme Heinrich
Der arme Lazarus
Der arme und der reiche König
Der begrabene Ehemann
Der beichtende Student
Der blinde Dieb
Der blinde Führer
Der dankbare Lindwurm
Der durstige Einsiedel
Der eigensinnige Spötter
Der ernsthafte König
Der ertrunkene Gärtner
Der falsche Blinde
Der feige Ehemann
Der fünfmal getötete Pfarrer
Der gefangene Räuber
Der geäffte Pfaffe
Der grave von Anhalte
Der guote Gêrhart
Der heilige Georg
Der hässliche Pfaffe
Der junge Baum
Der junge Ratgeber
Der kluge Knecht
Der milde König
Der nackte Bote
Der nackte Ritter
Der tugendhafte Schreiber
Der unbelehrbare Zecher
Der unfruchtbare Baum
Der ungeratene Sohn

Der ungetreue Knecht
Der verflogene Falke
Der von Gliers
Der von Kürenberg (Minnesangs Frühling)
Der von Trostberg
Der von Wengen
Der wahre Freund
Der wilde Alexander
Der wunderbare Stein
Des Ehemanns Rat
Des Königs alte Kleider
Des Muses Lehre
Des Mönches Not
Des Teufels Ammen
Di tutsch kronik von Behem lant
Die Ave Maria-Lilie (Von einem Edelmann)
Die Bischöfe
Die Blume im Munde (Von einem sündigen Schüler)
Die Buhlschaft auf dem Baume (Fassung A)
Die Ehre und das Seelenheil
Die Eule und der Habicht
Die Frauenehre
Die Frauentreue
Die Geistlichen
Die Goldene Schmiede
Die Gäuhühner
Die Heidin (Fassung B)
Die Herren zu Österreich
Die Katze
Die Katze als Nonne
Die Klage
Die Klage der Kunst
Die Königin vom Mohrenland
Die Mahnung zur rechtzeitigen Buße
Die Martinsnacht
Die Maße
Die Messe
Die Milch und die Fliegen
Die Minneburg
Die Männler
Die Pfaffendirne
Die Rabenschlacht

Die Rache des Ehemannes
Die Rebhühner
Die Rettung des gehängten Diebes
Die Schlange ohne Gift
Die Sommerlatte
Die Spieler
Die Suche nach dem glücklichen Ehepaar
Die Sünderin
Die Tinte
Die Tochter und der Hund
Die Weisheit Salomons
Die beiden Knappen
Die beiden Knechte
Die beiden Königinnen
Die beiden Zimmerleute
Die böse Adelheit
Die drei Gott verhaßtesten Dinge
Die drei Mönche zu Kolmar
Die drei Waffen
Die drei Wünsche
Die eingemauerte Frau
Die ewige Verdammnis
Die falsche und die rechte Freigebigkeit
Die feisten Jagdvögel
Die freigebige Königin
Die geliehenen Kleider
Die gepfefferte Speise
Die geschwätzigen Mönche (Die Mönche und der T...
Die gestohlene Monstranz
Die getreue Gattin
Die geweihten Toren
Die halbe Birne
Die jüngere Judith
Die milde Königin
Die reiche Stadt
Die schreiende Klage
Die sechs Teufelsscharen
Die sechs Versuchungen
Die sieben himmlischen Gaben
Die tumben Pfaffen
Die törichten Pfaffen
Die undankbaren Gäste

Die ungehorsamen Juden
Die unschuldige Mörderin
Die verlorenen Christen
Die vier Evangelisten
Die wandelbaren Juden
Die zwei Brote
Die zwei Herren
Die Äffin und die Nuß
Die Äffin und ihre Kinder
Dietmar der Setzer
Dietmar von Eist (Minnesangs Frühling)
Dietrich und Wenezlan
Dietrichs Flucht
Diocletianus
Disputation mit einem Juden
Diu Crone
Diu urstende
Drei Wünsche
Drei listige Frauen
Drei Äpfel zur Warnung (Ein Ritter wird Einsie...
Eberhard von Sax
Eckenlied (Fassung L.)
Edelmann und Pferdehändler
Ehmanns Rat
Ehre und Seelenheil
Ein Beispiel Salomos
Ein Scholar, Marias Bräutigam
Ein Sohn beißt dem Vater die Nase ab
Ein böses Weib scheidet eine Ehe
Eneide
Engelhard
Engelhart von Adelnburg (Minnesangs Frühling)
Engeltaler Schwesternbuch
Erec
Ernsthafte König
Erscheinung am Lichtmesstage
Esel
Eule und Habicht
Ezzolied (Hs. V)
Falke und Eule
Falsche und rechte Milte
Fliege und Kahlkopf

Flore und Blanscheflur
Flugschrift Donnerstein (Deutsche Fassung)
Frau Ehre und Frau Schande
Frauenbuch (Lachmann)
Frauenbuch (Spechtler)
Frauendienst (Bechst.) (Epik, Bechstein)
Frauendienst (Briefe)
Frauendienst (Büech.) (Büechlîn, Bechstein)
Frauendienst (Spechtler) (Epik, Spechtler)
Frauendienst Lieder (Bechstein)
Frauenehre
Frauenleben und Pfaffenleben
Frauenlob Teil I (Leiche)
Frauenlob Teil II (Lieder)
Frauenlob Teil III (Supplementband)
Friderich von Liningen
Friedrich der Knecht
Friedrich von Hausen (Minnesangs Frühling)
Friedrich von Sonnenburg (Die Sprüche)
Fuchs und Wolf im Eimer
Garel von dem blüenden Tal
Gast
Gaude Maria Virgo (Ein Blindgeborener wird seh...
Gauriel von Muntabel
Gebet von den Freuden Marias
Gebet zum Meßopfer
Gebet zum Schutzengel
Gedrut — Geltar
Gegen Gleichgeschlechtlichkeit
Geliehene Kleider
Goeli
Goesli von Ehenhein
Goldemar
Gott ist Vater, Herr und Bruder
Gottes Zukunft
Gottfried von Neifen
Gottfried von Strassburg
Gottfried von Straßburg (Minnesangs Frühling)
Graf Rudolf
Gregorius
Gärtner
Gäuhühner

Günther von dem Vorste
Hartmann von Aue (Minnesangs Frühling)
Hartmann von Starkenberg
Hartwig von Rute (Minnesangs Frühling)
Hase und Löwe
Hawart
Heinrich Hetzbolt von Wissense
Heinrich Teschler
Heinrich der Rost
Heinrich von Frauenberg
Heinrich von Kempten
Heinrich von Morungen (Minnesangs Frühling)
Heinrich von Rugge (Minnesangs Frühling)
Heinrich von Sax
Heinrich von Stretelingen
Heinrich von Tettingen
Heinrich von Veldeke (Minnesangs Frühling)
Heinrich von Veldeke (limburg. Fassg.) (Minnes...
Heinrich von der Muore
Helmbrecht
Herger (Minnesangs Frühling)
Herrand von Wildonie Lieder
Herren zu Österreich
Herzmaere
Herzog Ernst (Hs. B)
Herzog Ernst (Hs.D, strophig)
Herzog Heinrich von Pressela
Hesso von Rinach
Hilfe in Seelennot (Maria gebietet dem Meeress...
Hiltbolt von Schwangau
Hofhund und Jagdhund
Hort
Hugo von Montfort
Hugo von Mülndorf
Hugo von Werbenwag
Hund und Stein
Iwein
Jakob von Warte
Johann von Ringgenberg
Johannes Hadlaub
Kaiser Heinrich (Minnesangs Frühling)
Kaiserchronik

Kaiserchronik (Anhang 1) (Bairische Fortsetzung)
Kaiserchronik (Anhang 2) (Schwäbische Fortsetz...
Karl der Grosse
Kater als Freier
Katzenauge
Kirchtag
Kluger Knecht
Koninc Ermenrîkes Dôt
Konrad von Altstetten
Konrad von Kilchberg
Konrad von Landeck
Konrad von Würzburg Lyrik
Kraft von Toggenburg
Krämer
Kudrun
Käfer im Rosenhaus
Kölner Fechtbuch
König Konrad der Junge
König Rother
Königin vom Mohrenland
Lambrechts Alexander (Strassburger Hs.)
Lamprechts Alexander (Vorauer Hs.)
Lannzilet
Lanzelet
Laurin
Leich (W.v.V.)
Leuthold von Seven
Liet von Troye
Lohengrin
Löwe und Maus
Mahnung zu rechtzeitiger Buße
Mai und Beaflor
Maria im Turnier (Maria und der Ritter)
Maria rettet einen Maler (Der Maler und der Te...
Maria rettet einen Ritter um seiner Frau Wille...
Marias Fürbitte für einen Ritter (Der Ritter u...
Marien Rosenkranz (Der Mönch und die Rosenkränze)
Mariengruß
Marienlitanei
Markgraf Heinrich von Meißen
Marktdieb
Meinloh von Sevelingen (Minnesangs Frühling)

Meleranz
Moriz von Craûn
Mönch von Salzburg (Weltliche Lieder)
Mönch von Salzburg (Geistliche Lieder)
Nackte Ritter
Namenlos
Namenlose Lieder (Minnesangs Frühling)
Neidhart-Lieder Hs. C (SNE)
Neidhart-Lieder Hs. R (SNE)
Neidhart-Lieder Hs. c (Bennewitz)
Neidhart-Lieder Hs. c (SNE)
Nibelungenlied (B/C) (Bartsch/deBoor)
Nibelungenlied (C) (Hs. C)
Nibelungenlied (Hs. A) (Batts)
Nibelungenlied (Hs. B) (nach Batts)
Niune
Ortnit
Oswald Lieder
Otto von Bottenlouben
Otto von Brandenburg mit dem Pfeil
Otto zum Turm
Pantaleon
Partonopier und Meliur
Parzival
Paränese zum Mariengruß
Passionsgebet
Pfaffe Amis
Pfeffel
Predigtmärlein vom brüllenden Löwen
Processus Luciferi
Pyramus und Thisbe
Reimar der Fiedler
Reimar der Junge
Reimar von Brenneberg
Reimchronik der Stadt Cöln
Reinfried von Braunschweig
Reinhart Fuchs
Reinmar (Minnesangs Frühling)
Rennewart
Respice Finem
Ritter Beringer
Rosengarten (Hs. A)

Rosengarten (Hs.D)
Rubin
Rubin und Rüdeger
Rudolf der Schreiber
Rudolf von Fenis (Minnesangs Frühling)
Rudolf von Rotenburg
Rumelant v. Sachsen
Salman und Morolf
Salman und Morolf (Forts.) (Fortsetzung)
Salve Sancta Parens (Der einfältige Pfarrer)
Sanct Franzisken Leben
Schrätel und Wasserbär
Seifrits Alexander
Sigenot
Silvester
Spervogel (Minnesangs Frühling)
St. Petrus und der Holzhacker
Steinmar
Steirische Reimchronik
Süezkint
Süezkint der Jude von Trimperg
Sünder und Einsiedel
Tandareis und Flordibel
Tannhäuser
Theophilus
Titurel
Tochter Syon
Tristan (H.v.F.)
Tristan (Ulrich v. Türheim)
Tristan als Mönch
Tristan und Isold
Tristrant
Tristrant (Hs. D)
Tristrant (Hs. H)
Ulrich von Baumburg
Ulrich von Gutenburg (Minnesangs Frühling)
Ulrich von Liechtenstein
Ulrich von Singenberg
Ulrich von Winterstetten Leich
Ulrich von Winterstetten Lieder
Unser vrouwen hinvart
Van der Girheit

Vaterunser
Verflogener Falke
Verlorenen Christen
Veronica
Vespasianus
Virginal
Vom Geburtsfest Marias
Vom Tode
Vom heiligen Geist
Vom jüngsten Tage
Von Edelsteinen
Von Munegiur
Von Obernburg
Von Sachsendorf
Von Scharfenberg
Von Stadeck
Von Stamheim
Von Suoneck
Von Wissenlo
Von bösen Frauen
Von christlichen Werken
Von der Hochfahrt
Von der Hoffahrt
Von der Messe
Von der besten Frau
Von drei Freunden
Von einem varnden Schuler
Von üblen Weibern
Vorrede
Wahsmuot von Kunzich
Wahsmuot von Mülnhusen
Walberan
Waldschrat
Walter von Mezze
Walther von Klingen
Walther von Prisach
Walther: Lieder und Sangsprüche
Waltram von Gresten
Warum Gott sein Haupt neiget
Weidemann
Weltchronik
Wenzel von Beheim

Werner von Elmendorf
Wernher von Hohenberg
Wernher von Teufen
Wigalois, der Ritter mit dem Rade
Wigamur
Wildes Roß
Wildpret
Wilhelm von Wenden
Wilhelm von Österreich
Willehalm (U.v.T.)
Willehalm (Wolfram)
Willehelm von Heinzenburg
Winli
Wolfdietrich (Hs. A)
Wolfdietrich (Hs. B.)
Wolfdietrich (Hs. D.)
Wolfram Lieder
Wolfram von Eschenbach (Minnesangs Frühling)
Wucherer
Zwei Herren
Zweierlei Bettzeug
Zwischenrede
von Buchheim

Table E.1: Verse texts in Chapter 4