

A COMPUTATIONAL ANALYSIS OF POETIC CRAFT IN
CONTEMPORARY PROFESSIONAL AND AMATEUR POETRY

A THESIS
SUBMITTED TO THE DEPARTMENT OF SYMBOLIC SYSTEMS
OF STANFORD UNIVERSITY
IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE DEGREE OF
BACHELORS OF SCIENCE WITH HONORS

Justine T. Kao
June 2011

Justine T. Kao

I certify that I have read this thesis and that, in my opinion, it is fully adequate in scope and quality as a thesis for the degree of Bachelors of Science with Honors.

(Daniel Jurafsky
Computer Science and Linguistics) Principal Adviser

I certify that I have read this thesis and that, in my opinion, it is fully adequate in scope and quality as a thesis for the degree of Bachelors of Science with Honors.

(Christopher Potts
Linguistics)

Abstract

What makes a poem beautiful? To answer this question, this thesis uses computational methods to examine stylistic and content features employed by professional and amateur poets. I propose and test the hypotheses below: (1) Beautiful poems contain more sophisticated diction and poetic sound devices (2) Beautiful poems are more emotional (3) Beautiful poems contain more concrete details and imagery. Building upon existing techniques designed to quantitatively analyze poetic style and sentiment in texts, I identify and analyze factors testing these hypotheses. Results showed no significant difference between the difficulty levels of professional and amateur poets' vocabulary. Contrary to prediction, professional poets use significantly less sound devices such as rhyme and alliteration than amateur poets. Professional poems were also found to contain significantly fewer emotional words than amateur poems. The most important mark of good poetry was instead concrete and specific imagery. Professional poets refer to specific objects more often and employ a greater variety of words. This suggests that professional poets are able to evoke sentiment using concrete and specific imagery instead of abstract or explicitly emotional words. These results challenge and confirm several established doctrines in creative writing and poetic theory, suggesting that methods from computational linguistics may help support the analysis of beauty in verbal art.

*Ink runs from the corners of my mouth.
There is no happiness like mine.
I have been eating poetry.*

—From *Eating Poetry*, by Mark Strand

Acknowledgements

I would like to thank my advisor Professor Dan Jurafsky for his guidance and encouragement. This thesis would not have been possible without his experience, patience, and wonderful enthusiasm. I would like to thank Professors Christopher Potts, Lera Boroditsky, Todd Davies, and Michael Frank for their insightful comments in the early stages of the project. Thank you also to David Kaplan for his generosity in sharing the code for some of the features used in this thesis.

Thank you to every teacher, friend, and writer who has ever inspired me to read or write. An enormous thanks to my friends for always being there: Aya Inamori for cookies, love, and caffeine; Viria Vichit-Vadakan for merit and laughter; Johaina Crisostomo, the English major, for tolerating (and inspiring) my computational spin on literature; Michael Ding for his affection and support.

Finally, love and gratitude goes to Joann Kao for putting up with a nerdy sister for fifteen years. And most of all to my parents, who gave me both my languages and everything else that is beautiful in my life.

Contents

Abstract

Acknowledgements

1	Introduction	1
1.1	What makes a poem beautiful?	2
1.2	Computational linguistics and poetry	4
1.3	Roadmap	5
2	Theoretical components	6
2.1	Elements of Craft	6
2.2	Style: The Poem as a Verbal Object	7
2.2.1	Diction	7
2.2.2	Sound Device	9
2.3	Content: the poem as a unique perspective	12
2.3.1	Emotion	12
2.3.2	Imagery	14
3	Computational components	16
3.1	Computational Aesthetics	16
3.2	Computational Analysis of Poetic Style	18
3.3	Sentiment and Content Analysis	19
3.4	Integration	21
4	Materials and Methods	22
4.1	Poetry Corpus	22

4.1.1	Professional Poets	22
4.1.2	Amateur Poets	23
4.2	Features	24
4.2.1	Style	24
4.2.2	Emotion	26
4.2.3	Imagery	27
5	Analysis	29
5.1	Logistic Regression Model	29
5.2	Results	29
5.3	Analysis of content features	31
5.3.1	Emotion and sentiment	32
5.3.2	Imagery	34
6	Discussion	36
6.1	Style	37
6.2	Emotion	38
6.3	Imagery	39
6.4	Summary	41
7	Future Directions	42
7.1	General strengths and weaknesses	42
7.2	Future Extensions	43
	Bibliography	46

List of Tables

4.1	List of “generalization” words	28
5.1	Model formula	30
5.2	Odds ratios and p values of significant predictors of professional poetry	31
5.3	POSITIVE OUTLOOK (Positiv) words in professional and amateur poetry	32
5.4	NEGATIVE OUTLOOK (Negativ) words in professional and amateur poetry	33
5.5	POSITIVE EMOTION (posEmo) words in professional and amateur poetry	33
5.6	NEGATIVE EMOTION (negEmo) words in professional and amateur poetry	34
5.7	CONCRETE (Object) words in professional and amateur poetry . . .	35
5.8	ABSTRACT (ABS) words in professional and amateur poetry	35

Chapter 1

Introduction

Imagine that you are trying to write a poem for your loved one. After several hours of hard work, you are dissatisfied with your creation and feel that your affection simply cannot be expressed by words. You browse the Internet for inspiration using the keywords “love poem” and find the following two poems.

Poem A

*I love the way you look at me,
Your eyes so bright and blue.
I love the way you kiss me,
Your lips so soft and smooth.*

*I love the way you make me so happy,
And the ways you show you care.
I love the way you say, "I Love You,"
And the way you're always there.*

Poem B

*My words rained over you, stroking you.
A long time I have loved the sunned mother-of-pearl of your body.
I go so far as to think that you own the universe.
I will bring you happy flowers from the mountains, bluebells,
dark hazels, and rustic baskets of kisses.*

*I want
to do with you what spring does with the cherry trees.*

If you had to guess, which poem do you think was written by Pablo Neruda, winner of the Nobel Prize for Literature and one of the greatest writers of love poetry in the 20th century? Even given the subjectivity inherent in aesthetic judgements of poetry, the answer should be apparent. Indeed, 85.7% of untrained subjects recruited on Amazon Mechanical Turk were able to identify the second poem as the one written by a professional poet, and 74.3% subjects found the second poem to be more beautiful.

What factors contribute to these intuitive judgments of sophistication and poetic beauty? What literary devices do professional poets use to distinguish their work from mediocre poems? In this thesis, I use modern computational techniques to help answer questions that have intrigued poets, literary scholars, and philosophers throughout the ages.

1.1 What makes a poem beautiful?

Poetry is language at its most distilled and most powerful.

—Rita Dove (1952–), Poet Laureate and winner of the Poetry Pulitzer Prize

The quest to understand poetic beauty dates back over two thousand years to Aristotle’s *Poetics* (fourth century BC) and continues in literature classrooms and creative writing workshops today. While literary critics are interested in analyzing defining characteristics of great works of literature, the goal of creative writers is to take those characteristics and apply them to their own writing in innovative ways. Analyzing aesthetic language serves the dual purposes of illuminating humanities’ literary lineage and helping current writers perfect their art according to the aesthetic ideals of their time.

Every question concerned with the evaluation of art faces the challenge of subjectivity. Is there such a thing, after all, as the idea of one poem being absolutely more beautiful than another? While beauty may be in the eye of the reader, the fact that publishers, critics, and readers have preferences and that these preferences often have patterns suggests that there exist certain general rules about what makes a poem good. The creative writing workshop, a fairly recent invention, was founded

on the idea that writers can benefit from learning more about basic elements of craft (Earnshaw, 2007). In a well-known poetry instruction book, “The Poet’s Companion: A Guide to the Pleasures of Writing Poetry,” Addonizio & Laux (1997) write: “Craft provides the tools: knowing how to make a successful metaphor, when to break a line, how to revise and rewrite—these are some of the techniques the aspiring poet must master.” The value placed on the elements of craft in creative writing workshops where the chief goal is to produce better works of art suggests that these elements are key to writing beautiful and successful poetry.

What are some possible characteristics of poetic craft that may be unique to good poetry? One common perception of good poets is that they are skilled with words and have a larger vocabulary at their disposal. This suggests that good poetry is perhaps more likely to contain sophisticated words that demonstrate the poet’s mastery of language. Literary critics and practicing poets also value poetry’s musical quality and often view it as one of the defining characteristics of superior craft. Another perception of poetry is that beautiful poems tend to be highly evocative and emotional. Lastly, good poetry should “show, not tell,” and contain concrete and specific imagery that form a deeper meaning in the poem (Earnshaw, 2007). These common theories about the elements of craft can be summarized into the following three hypotheses:

Hypothesis 1: Good poetry contains more sophisticated diction and richer sound devices.

Hypothesis 2: Good poetry is more emotional.

Hypothesis 3: Good poetry contains more concrete imagery.

The following chapters explain the theoretical motivations for these three hypotheses in more detail and test them using computational methods. It is important to note that since standards and trends for what is considered “good poetry” evolve through time, this project focuses on contemporary poetry and aims to understand what distinguishes a good poem from a bad one to the mind of a modern poet and reader.

1.2 Computational linguistics and poetry

Why should we use methods from computational linguistics to analyze poetic beauty? First of all, literary writing is an important subset of language, one that is appreciated and celebrated by a substantial portion of the language community (Nell, 1988; Zwicky, 2000). While language has long been a major area of research in cognitive psychology (Gentner, 2010; Best, 1992), the majority of the language community seems to take for granted the kind of language studied by the majority of the science community. The ubiquity of language and the efficiency of our linguistic faculties render language almost invisible, and we are as oblivious to its wonders as a fish is to water. As American poet and critic James Russell Lowell once argued in favor of the establishment of the academic study of literature, “The blossoms of language have certainly as much value as its roots, for if the roots secrete food and thereby transmit life to the plant, yet the joyous consummation of that life is in the blossoms, which alone bear the seeds that distribute and renew it in other growths” (Earnshaw, 2007). It is “the blossoms of language,” frivolous, impractical, fragrant, and colorful, that surface from a river of linguistic input and make the average language user think, “Isn’t language wonderful?”

While the sensation of having understood and resonated with a beautiful poem is not a rare phenomenon, little empirical research has been done to examine the textual features or mental processes that engender this sensation. This does not mean that aesthetics is altogether too “subjective” an area to study empirically. Previous research has examined the aesthetics of faces, visual art, and music using scientific methods, and results from these studies reveal important features of beauty that inform our knowledge of human evolution and perception (Fink & Penton-Voak, 2002; Zaidel & Cohen, 2005; Rigau et al., 2008; Reber et al., 2004; Grinde, 1996; Smith & Melara, 1990; Smith & Melar, 1990). Poetry engages many of the same mental faculties used to process ordinary language, such as our ability to make inferences, recognize patterns, and resolve ambiguity. I argue that the scientific study of aesthetic language can also contribute to our understanding of the origins and mechanisms of language and linguistic faculties.

While there is much to explore in the cognitive processes that support the appreciation of aesthetic language, this thesis focuses on analyzing textual features that may be responsible for generating sensations of poetic beauty. Understanding the

textual features of aesthetic language is the first step to gaining insight on how these features may evoke aesthetic sensations. By comparing the poetic devices used in poems written by professional and amateur poets, I hope to reveal interesting ways in which skilled poets convey meaning and beauty through language.

This project aims to accomplish several different but complementing goals. From the perspective of natural language processing and digital humanities, I hope to develop computational methods that test existing theories in literature and poetry. From the point of view of cognitive psychology, I hope to examine the defining features of aesthetic language in order to understand how they might engender sensations of beauty in readers. Finally, as a lover of language and literature, I hope to help promote the interdisciplinary study of language and contribute new methods to the analysis and appreciation of verbal art.

It is important to note that this project does not attempt to reduce the wisdom and skill of poets to hard-and-fast rules and numbers. Rather, its aim is to develop new methods that celebrate and carry on an age-old conversation about humankind's relationship with language and beauty. Throughout this thesis, I will pepper quotes about poetry written by famous poets to create a sense of dialogue between our current project and practicing poets throughout the ages. My hope is to make a small contribution to the rich and fruitful tradition of qualitative research in literature using computational techniques.

1.3 Roadmap

The rest of the thesis is organized as follows. Chapter 2 provides a theoretical background for possible features of good poetry. Chapter 3 reviews areas of research in natural language processing that provide the computational basis of this thesis. Chapter 4 describes the poetry corpus and introduces the feature set. Results and analyses are given in Chapter 5. Chapter 6 discusses the implications of these results as well as how they compare to the theories on poetic craft. Chapter 7 proposes possible improvements and directions for future research.

Chapter 2

Theoretical components

One demands two things of a poem. Firstly, it must be a well-made verbal object that does honor to the language in which it is written. Secondly, it must say something significant about a reality common to us all, but perceived from a unique perspective. What the poet says has never been said before, but, once he has said it, his readers recognize its validity for themselves.

—W. H. Auden (1907 - 1973)

u

2.1 Elements of Craft

The quote from W. H. Auden, a poet noted for his stylistic and technical achievements (Academy of American Poets, 2007), represents two types of criteria for a good poem. The requirement that a poem must be a “well-made verbal object” is a criteria about style; the demand for a good poem to describe familiar issues “from a unique perspective” is concerned with issues of content.

In this section, I will review several elements of style and content that most serious poets, regardless of era and agenda, must consider when they write. Each poet may have a different way of interpreting and realizing these elements of craft, but it is a common language that they share in much the same way that experimental scientists, regardless of their particular fields of research, follow established principles of experimental design. Since these elements of craft are repeatedly emphasized in

creative writing settings and literary criticism, they are also fundamental to how most trained readers and writers approach, analyze, and evaluate poetry.

2.2 Style: The Poem as a Verbal Object

2.2.1 Diction

Prose is words in their best order; poetry is the best words in their best order.

—Samuel Taylor Coleridge (1772 – 1834)

Aristotle was one of the first to point out that writers should pay close attention to diction (Earnshaw, 2007). According to Aristotle, the key to good diction is the balance between ordinary words that make the writing comprehensible and strange words that make the writing distinguished. This balance allows creative writers to be innovative and clear at the same time (Aristotle, 1998).

A few hundred years later in Greece, first century AD, a teacher of rhetoric named Longinus wrote an essay describing the sources of aesthetic experience, which he called “the sublime.” Longinus included “noble diction and elevated word arrangement” as one of the primary sources of aesthetic language (Earnshaw, 2007; Longinus, 2001). These doctrines proposed by the earliest scholars of poetic craft suggest that poetic beauty may stem from the level of individual words.

The value of diction in aesthetic writing has carried through to modern ideas on poetic craft. In her creative writing textbook titled, “Imaginative Writing: The Elements of Craft,” Burroway (2007) describes poetry as a high-density form of language. The density of poetry is not simply a consequence of most poems being shorter than most prose writing; poetic language is usually intentionally ambiguous and strives to convey several meanings at once (Burroway, 2007; Addonizio & Laux, 1997). As a result, each word in a poem carries an especially heavy weight. Creative writing instructors often remind students to choose their words with care, and literary critics have been known to spend pages dissecting a single pivotal word. Based on these ideas, good poetry is perhaps that which uses more sophisticated vocabulary.

Word frequency

The importance of diction in poetic craft has interesting quantifiable implications in terms of word frequency. Word frequency has been shown to be strongly correlated with word difficulty (Breland, 1996; Graves et al., 1987), a characteristic of language that allows educational psychologists, linguists, and testing agencies to estimate word difficulty based on word frequency (Tomayo, 1987; Marks, Carolyn B. et al., 1974). Word frequency is also used as a feature in several well-performing automatic essay grading systems (Ben-Simon & Bennett, 2007; Burstein et al., 2004), which suggests that the tendency to use lower-frequency words is an indicator of superior language skills. It is thus reasonable to predict that poems written by professional poets might have lower average word frequencies than poems written by amateur poets.

Another reason why good poetry may contain lower-frequency words is related to word specificity. High-frequency words are usually distributed more generally across different contexts than low-frequency words. For example, many of the highest-frequency words in English are function words that are used across all topics, domains, and styles of writing. High-frequency words tend to have broader connotations and associations and are less able to convey meaning with the specificity that a poem requires. As a result, one might predict that good poetry often employs a set of lower-frequency words used only in specific contexts.

Consider the first two lines from a poem by Emily Dickinson:

*Hope is the thing with feathers—
that perches in the soul—*

If we replace the word *perches* (359 counts on the Corpus of Contemporary American English (COCA)) with the much more common hypernym *stands* (27,548 counts on COCA), or *resides* (1,459 counts on COCA), one gets a very different image of hope. Dickinson's word choice reinforces the idea of hope as a delicate, light, and birdlike entity, a connotation that would be lost if replaced by the other two word choices.

Based on the relationships between word frequency, vocabulary difficulty, and specificity, I predict that the average word frequencies of good poems should be lower than the average word frequencies of mediocre poems.

Type-token ratio

Readability measures and automatic essay grading systems often use the ratio of total word types to total number of words as an additional feature of vocabulary sophistication and complexity (Ben-Simon & Bennett, 2007; Pitler & Nenkova, 2008). Similarly, professional poets might also utilize a larger and more varied vocabulary by avoiding using the same word several times throughout a poem. If this is the case, then we would predict that poems written by professional poets would have a larger number of distinct word types than poems written by amateur poets.

2.2.2 Sound Device

Poetry is the rhythmical creation of beauty in words.

—Edgar Allan Poe (1809 – 1849)

Poetry has a rich oral tradition that predates literacy, and many of the earliest poems were recited or sung (Rubin, 1995). One of the most defining features that distinguish poetry from prose is that poetry is meant to be heard aloud. Traces of this aspect of poetic history can be found in sound devices such as rhyme, repetition, and meter, which are literary devices that are still the most closely associated with the poetic form. How a poem sounds is critical to how it is perceived, understood, and remembered. Burroway (2007) writes, “W. H. Auden once observed that the best poets often start out with a passion, not for ideas or people, but for the possibilities of sound.” Earnshaw (2007) also states that, “Writing poetry is very much about engaging the senses, especially the ear.” Indeed, most contemporary creative writing handbooks devote sections to defining various sound devices and analyzing notable poetry according to interesting patterns of sound (Burroway, 2007; Addonizio & Laux, 1997; Earnshaw, 2007).

Rhyme

Rhyme is one of the most well-known and popular sound devices in poetry. The earliest poets used strict rhyme schemes as a mnemonic device to help them memorize and recite long poems, and the tradition carried on to poetic forms such as sonnets and villanelles. Rubin (1995) wrote that “repeating patterns of sound in the form of

rhyme and alliteration cue memory more broadly and in less time than either imagery or meaning.”

Children’s nursery rhymes today are still full of words with identical end vowel sounds, and we are tickled by silly rhymes in Dr. Seuss books without really knowing why. Addonizio & Laux (1997) writes, ”We delight in language not only for its necessary function of helping us communicate, but for the sheer physical satisfaction of saying things.” Rhymes create an alluring echo and rhythm in poetry that draws our attention towards the surface features of the spoken word. Moreover, a word that rhymes with others is often given more emphasis because the repetition of sounds serves as a constant reminder of its presence. In the poem “One Art” by Elizabeth Bishop, the consistent rhyme scheme causes the word *disaster* to leave a pervasive impression throughout the poem.

*The art of losing isn’t hard to master;
So many things seem filled with the intent
to be lost that their loss is no disaster,*

*Lose something every day. Accept the fluster
of lost door keys, the hour badly spent.
The art of losing isn’t hard to master.*

Research in psychology confirms poets’ intuitions about the powerful effects of rhyme on perception and learning. For example, aphorisms that include rhymes are more likely to be perceived as true even though the rhymes do not contribute to the meaning of the messages (McGlone & Tofigbakhsh, 2000). Furthermore, exposure to rhymes has been shown to enhance phonological awareness in young children and can lead to better reading performances (Bryant et al., 1990).

However, most contemporary poetry no longer follows the long-standing tradition of rhyme schemes. Instead, free verse has become the pervasive form of modern poetry. This literary paradigm shift began with Ezra Pound and other Imagist poets who broke away from traditional forms to “create a new rhythm—as the expression of new moods” (Lowell, 1917).

While strict rhyme schemes involving identical ending vowel sounds are no longer the norm in modern poetry, repetition of sounds is still considered an important device and often manifests itself in different ways. Burroway (2007) writes that “poets

had tired of [perfect end rhymes] and launched into slant rhymes, also called off rhymes, thus extending the range of effects.” Slant rhymes are rhymes in which either the stressed vowels are identical, or the phonemes following the stressed vowels are identical, but not both. Here is an example of an off rhyme from a poem by Lucille Clifton, titled “scar” (1999).

*i will call you
ribbon of hunger
and desire
empty pocket flap
edge of before and after.*

The stressed vowel sounds of *hunger*, *desire*, and *after* are not identical, but the ending phonemes /er/ are the same. The effect of the rhyme is more subtle and varied than the perfect end rhymes in Robert Frost’s poem “The Road Not Taken” written earlier in the century (1915), where the rhyming words have identical stressed vowel sounds as well as identical phonemes that follow.

*Two roads diverged in a yellow wood,
And sorry I could not travel both
And be one traveler, long I stood
And looked down one as far as I could
To where it bent in the undergrowth.*

Given the shift from perfect end rhymes to slant rhymes noted in modern poetry, I predict that good modern poetry would have less perfect end rhymes but more slant rhymes than poems written by amateurs.

Alliteration and Consonance

Alliteration is the repetition of consonant sounds at the beginning of words, and consonance is the repetition of consonant sounds elsewhere. In addition to rhyme, alliteration was used as a powerful mnemonic device in ancient epic poetry (Rubin, 1995). Researchers in psychology and discourse analysis have shown that alliteration reactivates readers’ memories for previous information that was phonologically similar to the cue (Lea et al., 2008). In poetry, alliteration creates a schema of sounds that

evokes certain sensations. Here is an example from a line from Edgar Allan Poe's poem, titled "The Bells."

What a tale of terror, now, their turbulency tells!"

The repetition of the phoneme /t/ forms a grouping of voiceless plosive sounds and creates an effect of urgency and breathlessness. As Burroway (2007) writes, "Plosives explode, sibilants hiss, and gutturals growl." Alliteration and consonants allow poets to repeat distinctive sound clusters to evoke meaningful associations. I predict that poems written by skilled poets would contain more alliteration and consonance.

Assonance

Assonance is the repetition of vowel sounds. Similar to consonants, different vowel sounds also have their own characteristics and effects. Long vowels take longer to utter and draw out the rhythm and pacing of the line, while short vowels feel brief and urgent (Burroway, 2007). Here is an example from a love poem by Marvin Bell, titled "To Dorothy":

*So close, in the personal quiet
Of a windy night, it brushes the wall
And sweeps away the day till we sleep.*

The long /ee/ vowel is repeated in *sweep* and *sleep*, and the long /ae/ vowel is repeated in *away* and *day*. These repetitions of long vowels form a sense of lazy tenderness that echo the softness of something brushing against the wall. Based on the striking effects that vowel repetition can create in poetry, I predict that poems written by skilled poets would contain more assonance.

Since sound has long been a critical component of poetry, it is likely that good poets employ sound devices in particular ways that contribute to the beauty and musical quality of a poem. As a result, I predict that professional poets use richer sound devices than amateur poets, particularly subtler and more sophisticated effects such as slant rhyme, alliteration, consonance, and assonance.

2.3 Content: the poem as a unique perspective

2.3.1 Emotion

A poem begins with a lump in the throat, a home-sickness or a love-sickness. It is a reaching-out toward expression; an effort to find fulfillment. A complete poem is one where the emotion has found its thought and the thought has found the words.

—Robert Frost (1874 – 1963)

The emotional intensity of poetry and literary art has long been an object of both celebration and fear. While Longinus identified “strong and inspired emotion” as an important source of sublimity (Longinus, 2001), Plato criticized poetry and wished to ban it from his utopia on the grounds that poetry provokes emotions when we should strive to be calm and controlled (Plato, 2003).

These two divergent views on the emotional intensity of poetry are reflected in a modern debate on whether poetry should be utilized in therapeutic settings. Some studies have shown that the imagery, sound effects, and concern about human consciousness expressed in poetry allow patients to explore and reinterpret their own emotions in useful ways (Mcardle & Byrt, 2001). By writing their own poems, patients are also able to freely express their thoughts without the constraints of form and logic (Harrower, 1972; Bowman & Halfacre, 1994).

On the other hand, critics of poetry therapy have suggested that writing poetry may be harmful to psychological health because it allows the poet to immerse herself in an inexplicable emotion without having to make sense or order out of it (Stirman & Pennebaker, 2001). For example, Silverman & Will (1986) claimed that Sylvia Plath’s poetry may have undermined her control mechanisms and contributed to her death.

Whichever the case, the debate on poetry therapy centers on the idea that interaction with poetry is inherently emotional. Reading beautiful poetry is often an intensely emotional and cathartic experience, and good poems should be able to resonate strongly with readers. Do skilled poets use more or less emotional words to accomplish this effect? Freud once wrote, “Not I, but the poet discovered the unconscious.” If reading good poetry is found to be cathartic and therapeutic, do skilled poets make more references to psychological states and explore the mental and emotional world with more depth and clarity?

Moreover, what sorts of emotions do professional poets express in their writing? Great works of literature are often perceived as being dark or tragic, and indeed it is a difficult task to write a masterpiece with a happy ending. Consistent with the stereotype of the tortured and suffering poet, suicide rates are much higher among professional poets than among other creative writers or the general population (Jamison, 1993). If great poetry arises from “a lump in the throat,” do professional poets write more often about negative emotions than people who write poems recreationally?

2.3.2 Imagery

Poetry is fact given over to imagery.

—Rod McKuen (1933–), bestselling American poet, composer, and singer

Ezra Pound and other Imagist poets in the early twentieth century began the Imagism movement to counter the overly sentimental and moralizing tradition of Romantic and Victorian poetry (Blakeney Willaims, 2002). In an essay in the *Poetry* magazine that officially launched the movement, Pound defined an image as “that which presents an intellectual and emotional complex in an instant of time” (Lowell, 1917). The goal of the Imagist poet was to reveal the essence of objects using clear language and precise imagery.

Although the Imagist movement was cut short by World War I, Imagism has had a strong influence on Modernist poetry and beyond (Perkins, 1987). In most creative writing workshops today, one of the single most important and oft-repeated piece of advice for writers is this: “Show, don’t tell.” Burroway (2007) interprets this as meaning: “Use concrete, significant details that address the senses.” Effective imagery allows readers to bring in their own associations to understand and truly experience a new emotion, and skilled poets and writers are able to pick out specific sensory details that evoke larger abstractions and generalizations.

The appeal of concrete imagery may have roots in processes that facilitate learning and memory. Previous research has shown that concrete noun pairs are easier to memorize than abstract noun pairs, which suggests that imagery can enhance the learning of word pairings (Paivio et al., 1966). Other studies have shown that mental

imagery facilitates relational association between concepts (Bower, 1970). Furthermore, Jessen et al. (2000) found neural correlates that suggest that concrete are processed differently in the brain than abstract nouns. One of the reasons why we find poetic imagery striking may be due to the fact that imagery is a psychologically powerful tool that evokes rich associations formed by culture and personal experience.

Another reason why imagery is an essential element of poetic craft is that it allows writers to avoid falling into cliché, which is the bane of the creative writer's existence. Burroway (2007) writes, "flat writing is... full of abstractions, generalizations, and judgments. When these are replaced with nouns that call up a sense image and with verbs that represent actions we can visualize, the writing comes alive." Many abstract and common concepts can be embodied or evoked by surprising imagery. Consider the first two stanzas of a poem by Stephen Dobyns, titled "Fragments," written for his friend whose daughter had just passed away.

*Now there is a slit in the blue fabric of air.
His house spins faster. He holds down books,
chairs; his life and its objects fly upward:
vanishing black specks in the indifferent sky.*

*The sky is a torn piece of blue paper.
He tries to repair it, but the memory
of death is like paste on his fingers
and certain days stick like dead flies.*

These two stanzas convey a profound sense of grief and disintegration through a series of concrete imagery. Instead of using the cliché, "his life was falling apart," Dobyns shows us specific objects that are moving away from him—books, chairs, and even a house. Instead of explicitly telling us how difficult it is for the man to forget his grief, Dobyns compares the memory of death to paste on his fingers on which dead flies stick. This striking imagery allows us to visualize an emotion that is otherwise abstract and indescribable.

I predict that skilled poets are more likely to describe concrete objects and less likely to reference abstract concepts or make generalizations.

Based on the teachings of poetic craft passed down from poets and philosophers as

well as previous research in literary theory and psychology, I predict that beautiful poetry should possess the following characteristics: elevated diction, varied word types, more sophisticated sound devices, stronger emotions, and more concrete imagery.

Chapter 3

Computational components

In this chapter, I review studies whose methods have influenced this work and provide a brief description of how they intersect to form the project’s computational backbone. I then discuss how techniques from natural language processing can be integrated to empirically examine the hypotheses posed in the previous chapter.

3.1 Computational Aesthetics

The word “Verse” is used here as the term most convenient for expressing, and without pedantry, all that is involved in the consideration of rhythm, rhyme, meter, and versification... the subject is exceedingly simple; one tenth of it, possibly may be called ethical; nine tenths, however, appertains to the mathematics.

—Edgar Allan Poe (1809 – 1849)

In 1933, mathematician G.D. Birkhoff proposed a formula for beauty called the aesthetic measure, which he characterized as a ratio between order and complexity. Birkhoff (1933) compared the measure against humans’ aesthetic judgments of shapes, sounds, and poems and found interesting (though not always consistent) correlations between the measure and people’s ratings. While several of his colleagues regarded this venture into the arts as an indication that Birkhoff’s serious mathematical work had ended, Birkhoff’s aesthetic measure has become one of the earliest and most influential attempts at analyzing beauty mathematically. In more recent years, the aesthetic measure has been conceptualized within the framework of information

theory and applied to computer graphics and the analysis of visual art (Rigau et al., 2008; Hekkert, 2006).

While the aesthetic measure has enjoyed some success in the domain of visual arts, it runs into problems of semantics, or “meaning,” when applied to language. Birkhoff’s aesthetic measure judges a poem’s beauty based solely on phonemic features. The orderliness of a poem is defined as the sum of its number of alliterations and assonance, rhymes, and musical vowels, subtracted by alliterative and assonantal excess and consonantal excess. This formula clearly does not capture the subtlety of word choice and richness of meaning in poetry. Indeed, Beebe-Center & Pratt (1937) found that people’s ratings for meaningful lines of poetry do not correspond with the scores given by the aesthetic measure. However, when asked to rate non-sense sentences such as *salanta monaralume oarimely loase; bered ak filner dinstem jeebenot*, people’s ratings correlated more strongly with the measure. This suggests, rather unsurprisingly, that the beauty of a poem is shaped in part by its meaning. Since Birkhoff’s measure only considers phonetic features, it fails to fully quantify the aesthetic value of meaningful poetic texts.

A more recent attempt to capture poetic beauty using computational methods focused entirely on the semantic aspects of language. Liu & Maes (2005) developed a computer program called Aesthetiscope that maps impressions of literary narratives onto grids of colors to visualize the aesthetic experience of reading. Their goal is to design a machine that can translate aesthetic works into different modalities, and thus create art that people find psychologically meaningful. This process of translation involves modeling the various impressions and semantic associations generated by words and in turn mapping those associations onto colors. The colors are then arranged into an artistic grid, forming a machine-made abstract artwork “inspired” by the input narrative. To evaluate how well the color grids match the input narratives, Liu and Maes randomly paired color grid A with another grid B and asked subjects to guess which grid was inspired by the input narrative of A. These at-a-glance judgments had an accuracy rate of 75.2%, which is well above chance, suggesting that the Aesthetiscope can capture aesthetic impressions evoked by poetic narratives in a meaningful way simply by incorporating semantic associations.

Liu and Maes’ (2006) work shows that the semantic content of poetic language can be extracted and translated into tangible colors that evoke similar impressions.

However, one major weakness is that all stylistic information from the input narrative is lost in the translation process. Each input text is treated as a bag of words, and features such as “rhythm, rhyme, meter, and versification”—things that usually make a verse a verse—are no longer considered.

The fact that these computational approaches are able to quantify some aspects of verbal beauty suggests that the aesthetic experience may be more universal and patterned than we may think. The strengths and weaknesses of these two studies suggest that a more realistic approach would be to combine both stylistic and semantic elements to more accurately capture important aspects of poetry.

3.2 Computational Analysis of Poetic Style

Poetry without rules is like tennis without a net.

—Robert Frost (1874 – 1963)

Poetry is as precise as geometry.

—Gustave Flaubert (1821 – 1880)

The quantitative analysis of poetic style dates back to the 1940’s, when poet and literary critic Josephine Miles began her extensive work analyzing the surface statistics of poetry across time (Miles, 1946, 1967). While Miles’ work was influential in establishing a statistical framework for thinking about poetry, it was done largely by hand and thus limited in scope and size (Kaplan, 2006).

More recently, Hayward (1996)’s connectionist model of poetic meter incorporated more sophisticated and varied features in the analysis. For every feature considered, including prosody, meter, and syntax, Hayward hand-assigned numeric scores to each syllable in ten samples of poetry. Analyzing these scores allowed him to identify unique patterns for each poet and to note similarities within each time period. However, this analysis required Hayward’s personal assessment of the poems as well as manual assignment of feature scores. Since it is unfeasible to apply this method to a large set of poems, Hayward’s model also faces limitations in size and generalizability (Kaplan, 2006).

One of the most thorough and sophisticated computational analyses of poems to date of is the *PoetryAnalyzer*, Kaplan & Blei (2007)’s work on visualized comparison

of style in American poetry. Modern statistical and computational tools allowed the authors to integrate more features to analyze a large set of poems in an automated manner. The authors mapped poems from different poets and eras into a vector space based on three types of stylistic elements—orthographic, syntactic, and phonemic — in order to find stylistic similarities among poems. Since poetry differs from prose in a variety of ways, the authors integrated several features traditionally used by literary critics to analyze poetry and style. For example, on the orthographic level, features such as number of lines, number of stanzas, and average line length were considered. On the syntactic level, frequencies of different parts of speech were measured to capture linguistic formality. On the auditory level, the authors also included phonemic features such as rhyme, alliteration, assonance, and consonance.

After building feature vectors for a collection of poems, the authors projected them onto a two-dimensional space using PCA to allow for a visualized comparison of style for different poems. They showed that poems by the same poet generally clustered in this space, and that poems with qualitatively different styles tend to be further apart. This suggests that the features they implemented are important indicators of poetic style and can capture significant aspects of a poem’s form and character. Since the experience of reading a poem is very much influenced by its orthographic, syntactic, and phonemic features, these features may be highly relevant to understanding a poems aesthetic appeal as well.

3.3 Sentiment and Content Analysis

Poetry is the revelation of a feeling that the poet believes to be interior and personal which the reader recognizes as his own.

—Salvatore Quasimodo (1901 - 1968), Italian poet and winner of the Nobel Prize
for Literature

The advent of the Internet and digital social media and networks has sparked a growing need to extract social information from large bodies of text. Researchers in natural language processing are beginning to address the issue of understanding how language conveys sentiment, and sentiment analysis may provide useful tools for the analysis of emotional content in poetry.

A growing body of literature focuses on extracting sentiment and content information from texts using sentiment lexicons (Pang & Lee, 2008). These lexicons can map words in large quantities of texts to patterns of positive or negative sentiment, allowing researchers to develop methods for sentiment classification and opinion mining. Several different sets of sentiment lexicons are currently available for sentiment analysis research, such as WordNet (Fellbaum, 1998), SentiWordNet (Baccianella et al., 2010), Linguistic Inquiry and Word Count (LIWC) (Pennebaker et al., 2001), and the Harvard General Inquirer (Stone et al., 1966). These lexicons are composed of a number of valence or semantic-related categories. For a given text, one can analyze the distribution of words in each category and infer the general sentiment the text expresses. These lexicons have been used to help classify customer reviews of products based on positive or negative sentiment (Pang & Lee, 2008; Pang et al., 2002; Kennedy & Inkpen, 2006), analyze trends in public opinion (Thomas et al., 2006; O'Connor et al., 2010), understand social interaction (Ranganath et al., 2009; Jurafsky et al., 2009), and applied for medical and psychological purposes (Campbell & Pennebaker, 2003; Ramirez-Esparza et al., 2008).

One interesting study applies sentiment analysis to poetry by comparing the linguistic features of poems written by suicidal and non-suicidal poets (Stirman & Pennebaker, 2001). The authors analyzed 300 poems written by nine suicidal and nine non-suicidal professional poets matched in demographics and time period. Using LIWC (Pennebaker et al., 2001), the authors identified words in the poems pertaining to certain sentiment categories. Results showed that poems by suicidal poets contained more first-person singular self-references, less words referencing communication with others, and more words associated with death. However, the authors did not find significant differences between the suicidal and non-suicidal poets' use of positive and negative emotion words.

While it is somewhat surprising that suicidal and non-suicidal poets did not differ significantly in their use of emotional words, one possible explanation is that good poetry is inherently emotional, and thus the emotional salience of poetry does not vary significantly based on the poets' emotional stability. Stirman & Pennebaker (2001)'s results suggest that sentiment analysis techniques may help illuminate important patterns and psychological implications in the emotional content of poetry. While their work focused primarily on how language reflects mental health and personality,

this thesis is interested in how good poetry evokes emotional reactions in readers, a critical aspect that has largely been unexplored in most computational analyses of poetry.

3.4 Integration

Poetry is ordinary language raised to the Nth power. Poetry is boned with ideas, nerved and blooded with emotions, all held together by the delicate, tough skin of words.

—Paul Engle (1908 -1991), poet and director of the Iowa Writers’ Workshop

The studies reviewed above inspired and influenced the efforts put forth into this thesis. The empirical study of aesthetics is not itself a novel idea, and many of the tools used here have been researched and implemented in various areas of natural language processing. Research from computational aesthetics provides a background for the kinds of methodology and theories related to the empirical study of beauty. It also establishes our research question as a legitimate one to ask at this stage in our understanding of language. Previous work on the computational analysis of poetic style sheds interesting insight on important elements of poetry. The growing field of sentiment analysis in natural language processing also provides existing tools to extract sentiment from texts. Together, these domains help answer a question about poetry that combines computational methods with the analysis of aesthetics, style, and content.

The innovations that this research brings to the field of computational linguistics are two-fold. First, it seeks to combine both stylistic and content features in the analysis of poetry. This integration of style and sentiment analysis provides a more thorough view of the artistic elements in poetry. Second, the main goal of the thesis is not to identify stylistic trends or to cluster poems according to authorship. The question is not, “How can we use computational methods to determine poetic style?” but rather, “How can we use computational methods to examine what makes a poem beautiful?” By seeking to understand poetic features that contribute to the beauty and sophistication of a poem, this thesis integrates tools in computational linguistics to examine important questions of craft and aesthetics in contemporary poetry.

Chapter 4

Materials and Methods

4.1 Poetry Corpus

In order to analyze the defining features of good poetry, the first step is to build a corpus containing poems that vary in quality and “goodness” by some established standard. One way to do this would be to randomly sample poems from various sources and ask experts to rate them for quality and beauty. However, such a method can be expensive and time-consuming. A more efficient way of achieving a similar effect is to sample poems from pre-existing categories, such as poems written by skilled professional poets versus poems written by amateur poets. We assume that award-winning poets produce poems that experts would consider “better” and more beautiful than poetry written by amateurs. Although there might be exceptions, for example experts may consider some poems written by amateur poets to be very beautiful and sophisticated, these pre-existing categories for the most part should be a good approximation of expert opinions.

4.1.1 Professional Poets

One hundred poems were selected from sixty-seven professional poets whose work was included in a collection of Contemporary American Poetry (Poulin & Waters, 2006). The poets produced most of their work towards the middle and end of the 20th century and are considered contemporary poets. All of the poets are listed in the website of the Academy of American Poets (Academy of American Poets). Many of

the poets in this collection have won prestigious awards for their poems, such as the National Book Award for Poetry (Mark Doty (2008), Lucille Clifton (2000), Mary Oliver (1992), Robert Bly (1968), etc) and the Pulitzer Prize for Poetry (Stephen Dunn (2001), Mark Strand (1999), Rita Dove (1987), Anne Sexton (1957), etc). This serves as confirmation that the poets in this collection are widely acclaimed and that their craft is acknowledged and celebrated by poetry experts and literary critics.

After confirming that the poets in the collection are widely recognized for their writerly prowess, I randomly selected one to three poems from each poet, proportionate to the number of poems each poet had in the collection of Contemporary American Poetry (Poulin & Waters, 2006). When an excessively long poem (over 500 words) was selected, I removed it and replaced it with a different poem from the same poet. This served as a rough control for the length of the poems in the corpus. The final selection of one hundred professional poems ranged from 33 to 371 words in length with an average length of 175 words. I believe that these poems are a good representation of work produced by the best and most celebrated poets of our time.

4.1.2 Amateur Poets

One hundred poems were selected from amateur poets who submitted their work anonymously to a free and open-to-all website, aptly called “Amateur Writing” (www.amateur-writing.com). At the time of selection, the website had over 2500 amateur poem submissions by registered users. There is no screening process for the creative work submitted. As a result, the website contains a diverse set of poems submitted by amateur writers with a wide range of experience and skill levels.

I randomly selected one hundred poems from the website and did not encounter any poems that were over 500 words long. I corrected for mispellings and obvious grammatical errors in the poems to control for the effect of basic language skills on the differences between poems written by professional and amateur poets. The final selection of amateur poems ranged from 21 to 348 words in length with an average length of 135.8 words.

Sampling poems from two categories of contemporary poets—professional and amateur—yielded a poetry corpus of 200 poems that contain both beautiful and mediocre poetry.

4.2 Features

I implemented 17 features to test the hypotheses concerning the distinguishing characteristics of beautiful poetry.

Hypothesis 1: Good poetry contains more sophisticated diction.

Hypothesis 2: Good poetry is more emotional than bad poetry.

Hypothesis 3: Good poetry contains more concrete imagery than bad poetry.

The features described below target three separate domains: style, emotion, and imagery.

4.2.1 Style

The stylistic features described below were selected and designed with the purpose of capturing the elements of craft reviewed in the previous chapter.

Word frequency

As discussed in the previous chapter, good poems may contain lower-frequency words that reflect the poet's superior language skills. One simple stylistic feature that tests this hypothesis is the average word frequency of each poem. If poems written by professional poets have lower average word frequencies, then this may suggest that skilled poets employ more sophisticated vocabulary.

We tested the word frequency hypothesis using word frequency data from the Corpus of Contemporary American English (COCA) (Davies, 2011). A word list was downloaded containing 500,000 entries of the most common word forms in the COCA corpus, their frequencies, and part of speech. An **average log word frequency (aveLogFreq)** was obtained for each poem by looking up each word in the poem in the word list and summing up the log word frequencies. If a word does not appear in the wordlist, we assume that the word appears at least 10 times in the English language and add 1 to the total log word frequency. The total log frequency was then divided by the number of words in the poem.

Type-token ratio

To test whether or not professional poets have a more diverse vocabulary, a **type-token ratio (typeToken)** feature was calculated for each poem by counting all the separate instances of words and dividing that number by the number of words in the poem. This feature captures one aspect of the richness and complexity of vocabulary employed in professional and amateur poems.

Perfect end rhyme

The sound device features described below were computed using Kaplan (2006)'s *PoetryAnalyzer*. *PoetryAnalyzer* utilizes the Carnegie Mellon Pronouncing Dictionary to obtain pronunciations of words in each poem. Although the pronunciations were not perfect (for example, the noun “record” and the verb “record” are sometimes resolved incorrectly), Kaplan (2006) found that it was a sufficient tool for identifying most instances of poetic sound device.

An end rhyme is identified by examining the phoneme sequences at the end of lines. A window of four line endings is analyzed at a time. If two words in the window have different initial consonants but identical phoneme sequences from the stressed vowel phoneme onward, then an instance of a perfect end rhyme instance is recorded. The final count of perfect end rhymes in a poem is normalized by the total number of words.

Slant end rhyme

If two words in the window of four line endings have the same stressed vowel but different phonemes following the stressed vowel, then an instance of a slant end rhyme is recorded. The final count of slant end rhymes in a poem is normalized by the total number of words.

Alliteration

If the initial phoneme of two consecutive words are identical consonants, the alliteration count is incremented. The total count is then divided by the total number of words to obtain a **alliterationFreq** score for each poem.

Consonance

If there are at least two identical consonant phonemes in a window of nine syllables, the consonance count is incremented. The count is divided by the total number of words in a poem to obtain a **consonanceFreq** score.

Assonance

An **assonanceFreq** score was obtained for each poem in the same fashion as consonanceFreq, except the target phonemes are vowels instead of consonants.

4.2.2 Emotion

Several existing sentiment lexicons are available to use for sentiment analysis. Six categories from the Harvard General Inquirer (Stone et al., 1966) and two categories from LIWC (Pennebaker et al., 2001) were selected because they are most suitable for our purpose of analyzing elements of poetic craft.

The General Inquirer has 182 categories, including basic sentiment categories, categories for concrete objects, and categories for abstract concepts (Inquirer, 2011). (A full description of the 182 categories and samples of the words in each category can be found at <http://www.wjh.harvard.edu/inquirer/homecat.htm>.) While the Harvard General Inquirer was designed for content analysis, LIWC was designed to facilitate the understanding of individuals' cognitive and emotional states through text analysis. As a result, most of the categories in LIWC are concerned with mental activity and is composed of almost 4,500 words and word stems related to affective, social, and cognitive processes.

We focus on the following types of sentiment and content tags in order to understand the ways in which professional and amateur poets evoke emotion in poetry.

Positiv (Inquirer): This category contains 1915 words with a positive outlook, such as *able*, *love*, and *friend*. This category does not identify explicit references to positive emotion, but rather expressions concerning positive sentiment in general.

Negativ (Inquirer): This category contains 2219 words with a negative outlook, such as *abandon*, *death*, and *enemy*. This category also does not only

identify explicit references to negative emotion, but rather expressions concerning negative sentiment in general.

posEmo (LIWC): This category contains 408 word stems in regular expressions that are explicit references to positive emotion, such as *happiness*, *love*, and *hope*.

negEmo (LIWC): This category contains 499 word stems in regular expressions that are explicit references to negative emotion, such as *fury*, *sorrow*, and *grief*.

WlbPhys (Inquirer): This category contains 226 words concerned with physical well-being, such as *alive*, *dead*, *eat*, *young*, and *sleep*. It also contains words pertaining to body parts, such as *breast* and *lips*.

WlbPsyc (Inquirer): This category contains 138 words concerned with psychological well-being, such as *adjust*, *calm*, and *afraid*.

Positiv and **Negativ** were selected to examine and compare the general tone and outlook of professional versus amateur poems. **posEmo** and **negEmo** were selected in addition to the two main valence categories to examine whether professional poets use more or less explicitly emotional words than amateur poets. **WlbPhys** and **WlbPsyc** were used to examine the degree to which poems written by professional and amateur writers are concerned with physical versus mental states. More instances of words in the latter category in professional poems would indicate that professional poets make more explicit insights into human psychology.

Together, these categories from the General Inquirer and LIWC represent the emotional and sentiment-related elements in contemporary professional and amateur poetry.

4.2.3 Imagery

Most unforgettable poems engage in common and highly abstract issues of human experience such as love, death, grief, loneliness, and existence. However, skilled poets should discuss these broad concepts in concrete detail and imagery that make the “language come alive” (Burroway, 2007). The categories below were used to examine how professional and amateur poets employ concrete objects, abstractions, and generalizations to express poetic material.

Table 4.1: List of “generalization” words

all	none
every	everyone
always	never
everything	nothing
everybody	nobody
everywhere	nowhere

Object (Inquirer): This category contains 661 words that reference concrete objects, such as *boat*, *leaf*, *book*, and *doll*. The **Object** category is further divided into subcategories such as **Tool**, **Food**, and **NatObj**, but since all of the words in the subcategories are considered “concrete,” only the main category **Object** is considered. Since most poetic images involve concrete objects that the reader can visualize, this category is used as an approximation to identifying instances of specific imagery in poems.

ABS (Inquirer): This category contains 276 words related to abstract concepts, such as *causation*, *day*, *past*, and *destiny*.

generalization: This category was manually added in order to approximate the number of times generalizations appear in each poem by counting the number of words often used when making generalizing statements. The 13 words in this category are listed above.

For each category taken from the Inquirer, scores were calculated using the General Inquirer system available on a server (Inquirer, 2011). A score for a certain category is the number of words in a poem that appear in the category normalized by the length of the poem. For the two categories taken from LIWC, scores were calculated by counting the number of words in each poem that match a word stem in the LIWC dictionary and dividing it by the total number of words.

Chapter 5

Analysis

5.1 Logistic Regression Model

A score for each of the features was derived for every poem in the poetry corpus. The scores were normalized to have zero mean and unit variance.

To measure the effect of each variable on the likelihood of a poem being written by a professional or an amateur poet, we constructed a logistic regression model in R (R: A Language and Environment for Statistical Computing). For model selection, we used the step-wise backward elimination method. This method begins by building a model using all 17 feature variables. It then recursively eliminates variables that do not significantly contribute to explaining the variance in the data and stops when further eliminating a variable would significantly reduce the model fit. The final logistic regression model for the predictors of professional versus amateur poetry is summarized in the formula below (Table 5.1).

5.2 Results

Results show that poem type (professional or amateur) is significantly predicted by eight different variables ($p < 0.05$): type token ratio, perfect end rhyme frequency, alliteration frequency, positive outlook words, negative emotional words, concrete object words, abstract concept words, and generalization words. The other nine variables: average log word frequency, slant end rhyme frequency, assonance, consonance,

Table 5.1: Model formula

Probability(poem type = professional | X), where

$X\beta - 0.6071$	=		
-0.5039	*	average log word frequency	+
0.6646	*	type token ratio	+
0.4602	*	slant end rhyme frequency	+
-2.1	*	perfect end rhyme frequency	+
-0.6326	*	alliteration frequency	+
-1.0701	*	positive outlook words (Positiv)	+
-0.7861	*	negative emotional words (negEmo)	+
-0.5227	*	psychological words (WlbPsyc)	+
1.3124	*	concrete object words (Object)	+
-1.2633	*	abstract concept words (ABS)	+
-0.836	*	generalization words	

negative outlook words, positive emotional words, physical well-being words, and psychological words did not have significant predictive value. While the variables positive outlook (**Positiv**) and positive emotion (**posEmo**) were highly correlated ($r = 0.54$) and negative outlook (**Negativ**) and negative emotion (**negEmo**) are highly correlated ($r = 0.53$), there was no collinearity among the variables in the final logistic regression model selected by the backward elimination method.

The model predicts the likelihood of the poem type (professional or amateur) using the formula. The influence of each feature is represented by the coefficient β for each variable. A positive value for a coefficient increases the likelihood of a poem being written by a professional. For example, type token ratio and concrete object words have positive coefficient values; thus higher type token ratios and more concrete object words increase the likelihood of a poem being a professional poem. A negative value for a coefficient decreases the likelihood of a poem being written by a professional. For example, perfect end rhyme frequency has a negative coefficient value, and thus higher perfect end rhyme frequencies decrease the likelihood of a poem being written by a professional poet.

Coefficients in logistic regression models are often reported as odds ratios, which indicate the relative probabilities that a poem is written by a professional poet. If an odds ratio is greater than 1, it favors the outcome that a poem was written by a professional. The greater it is than 1, the more it favors the outcome. If an odds

Table 5.2: Odds ratios and p values of significant predictors of professional poetry

Feature variable	Odds	p -value
type token ratio	1.94	0.0308
perfect end rhyme frequency	0.12	$5.06e^{-7}$
alliteration frequency	0.53	0.0188
positive outlook words	0.34	0.0130
negative emotional words	0.46	0.0244
concrete object words	3.72	0.0002
abstract concept words	0.28	0.0027
generalization words	0.43	0.0035

ratio is smaller than 1, it favors the outcome that a poem was written by an amateur. The closer the ratio is to 0, the more it favors the outcome that a poem was written by an amateur. The relative odds and p -values of each significant predictor variable are in Table 5.2 below.

In summary, professional poems have higher type-token ratios, contain fewer perfect end rhymes, fewer instances of alliteration, fewer positive outlook words, fewer negative emotional words, more references to concrete objects, less references to abstract concepts, and fewer generalizations. From the odds ratios, we can see that the most significant predictors of professional poetry are fewer perfect end rhymes and more references to concrete objects.

5.3 Analysis of content features

In order to gain deeper insight on the results of content analysis using semantic categories from the Harvard General Inquirer and LIWC, the most frequent words in each of the significant categories were tallied for professional and amateur poems and listed in the tables below.

The lists of words from the Harvard General Inquirer categories **Positiv** and **Negativ** give us insight into the general outlook, tone, and topics that appear in professional and amateur poetry. Most of the frequent Positiv words in professional poems that do not appear very often in amateur poems concern positive physical or sensory attributes, such as *light*, *open*, and *bright*. Most of the frequent **Positiv** words in amateur poems that do not appear in professional poems are psychological

Table 5.3: POSITIVE OUTLOOK (Positiv) words in professional and amateur poetry

Professional (Total words: 17,055)		Amateur (Total words: 13,795)	
Word instances	527	Word instances	730
Word types	237	Word types	202
Word	Count	Word	Count
LIGHT	36	LOVE	134
LOVE	27	TRUE	31
GIVE	15	BEAUTY	25
FRIEND	14	GIVE	22
OPEN	12	SMILE	19
CLEAR	9	CARE	16
CLEAN	9	REAL	14
SWEET	8	GOOD	14
BRIGHT	8	FRIEND	13
LIKE	8	HAPPY	13

or abstract attributes such as *true*, *good*, and *happy*. While words like *love*, *give*, and *friend* seem to be prevalent in both professional and amateur poems, amateur poets use *love* much more often than profession poets.

Professional and amateur poets also seem to differ in the types of **Negativ** words they use. Professional poets tend to use more words related to physical states, such as *dead* and *cold*, while amateur poets are more likely to mention negative emotional states such as *sorrow* and *hate*.

5.3.1 Emotion and sentiment

Table 5.5 lists the top 10 most frequent words in professional and amateur poetry in the **posEmo** category. As shown in the table, amateur poets make more references to *love* than professional poets. Professional poets make more references to positive sensory words such as *warm*, *bright*, and *sweet*.

Table 5.6 lists the top 10 most frequent words in professional and amateur poetry in the **negEmo** category. Amateur poets use a greater number of explicitly negative emotion words than professional poets (275 counts versus 167). The list of most frequent negative emotion words in amateur poems also contain more references to expressions of negative emotion, such as *cry*, *crying*, and *tears*.

Table 5.4: NEGATIVE OUTLOOK (Negativ) words in professional and amateur poetry

Professional (Total words: 17,055)		Amateur (Total words: 13,795)	
Word instances	680	Word instances	567
Word types	330	Word types	248
Word	Count	Word	Count
NO	32	NO	42
DEAD	20	PAIN	16
DARK	18	NEED	16
DEATH	11	DARK	15
DIRT	11	LIE	14
LOST	10	HURT	13
COLD	9	SORROW	10
BURN	8	SAD	8
AGAINST	8	LOST	8
PASSE	8	HATE	8

Table 5.5: POSITIVE EMOTION (posEmo) words in professional and amateur poetry

Professional (Total words: 17,055)		Amateur (Total words: 13,795)	
Word instances	322	Word instances	522
Word types	88	Word types	104
Word	Count	Word	Count
LIKE	97	LOVE	119
LOVE	23	LIKE	70
FRIEND	11	CARE	15
WELL	11	GOOD	14
WARM	8	FREE	14
GOOD	8	HAPPY	14
BRIGHT	8	FUN	12
KISS	7	KISS	12
SWEET	7	BETTER	10
GREAT	7	FRIEND	10

Table 5.6: NEGATIVE EMOTION (negEmo) words in professional and amateur poetry

Professional (Total words: 17,055)		Amateur (Total words: 13,795)	
Word instances	167	Word instances	275
Word types	96	Word types	98
Word	Count	Word	Count
LOST	10	PAIN	18
BROKE	8	CRY	15
ALONE	7	TEARS	12
LIES	6	LIES	10
STRANGE	5	ALONE	9
SUFFERING	5	HURT	9
WEEP	4	LOST	8
FEAR	4	SAD	8
CRY	4	SORROW	8
GRAVE	3	CRYING	7

5.3.2 Imagery

As shown in Table 5.7, professional poets not only use more object words than amateur poets (698 counts versus 205), but they also use a larger and more diverse set of object words (250 types versus 85). Professional poets seem to reference natural objects very often, such as *tree*, *grass*, and *flower*. The most frequent Object word in amateur poems, on the other hand, is the extremely vague word *thing*. This suggests that even when amateur poets reference concrete objects, they do not use words that provide specific sensory details.

Table 5.7: CONCRETE (Object) words in professional and amateur poetry

Professional (Total words: 17,055)		Amateur (Total words: 13,795)	
Word instances	698	Word instances	205
Word types	250	Word types	85
Word	Count	Word	Count
TREE	29	THING	40
ROOM	20	WALL	12
THING	18	BED	11
GRASS	17	CLOCK	7
WALL	14	ROOM	7
FLOWER	13	TREE	6
GLASS	13	LEAVE	6
FLOOR	13	GIFT	5
CAR	12	MIRROR	4
DIRT	11	FLOWER	4

Table 5.8: ABSTRACT (ABS) words in professional and amateur poetry

Professional (Total words: 17,055)		Amateur (Total words: 13,795)	
Word instances	304	Word instances	345
Word types	82	Word types	75
Word	Count	Word	Count
DAY	40	DAY	54
NIGHT	31	TIME	33
YEAR	25	BEAUTY	25
TIME	20	SOUL	16
DEATH	11	NIGHT	15
NEW	9	NEW	14
MORNING	8	MOMENT	13
CHILDHOOD	7	CHRISTMAS	12
HOURL	7	THINK	11
AFTERNOON	7	FUTURE	9

Chapter 6

Discussion

*so much depends
upon*

*a red wheel
barrow*

*glazed with rain
water*

*beside the white
chickens.*

—*The Red Wheelbarrow*, by William Carlos Williams (1883—1963), Modernist and Imagist poet

What is it that skilled poets are doing differently from amateurs that allows them to write beautiful poetry?

Influential theories in literary tradition and poetic craft propose three hypothesis. First, professional poems were predicted to contain more sophisticated and varied vocabulary and richer sound devices. They were also expected to be more emotional, more negative, and more interested in psychological states. Lastly, I predicted that professional poets use more concrete imagery and less abstractions and generalizations than amateurs. This chapter discusses the results from the previous chapter and their implications on poetic craft and computational aesthetics.

6.1 Style

Poetry is a deal of joy and pain and wonder, with a dash of the dictionary.

—Kahlil Gibran (1883—1931), Lebanese American poet

Do professional poets use more sophisticated elements of craft such as diction and sound devices?

Based on the results from our regression model, it appears that professional poets do not use particularly sophisticated diction. The words in professional poetry are not significantly more unusual or difficult than words used by amateur writers. Instead, professional and amateur poets seem to use vocabulary with similar levels of difficulty. This suggests that contemporary poets are not interested in flowery diction or obscure words, but are focused on using ordinary words to create extraordinary effects. If the poetry corpus was seen as two bags of words, one professional and one amateur, one bag of words would not be significantly more difficult or “elevated” than the other.

However, one bag of words does contain more distinct word types. In the 100 poems written by professional poets, there are a total of 18,304 words and 4,315 distinct word types (23.57%). In the 100 poems written by amateur poets, there are a total of 14,046 words and 2,367 distinct word types (16.85%), a much smaller portion. In aggregate, professional poets have a larger and more varied vocabulary than amateur poets. Furthermore, professional poets use a significantly larger number of word types within each poem. Although professional poets do not use more difficult and unusual words, high type-token ratio is a significant predictor of professional poetry, suggesting that professional poems may be distinguished by a richer set of words.

The results on sound devices provide interesting insights into the current stylistic trends of contemporary professional poetry. While sound devices have a long history in poetry and are considered to be a significant part of poetic beauty, contemporary professional poets now use these devices much less often than amateur poets. Sound devices that traditionally were important in poetry for mnemonic purposes, such as rhyme and alliteration, are more prevalent in amateur poems. Even subtle and sophisticated sound devices like slant rhyme, consonance, and assonance are not significant indicators of professional poetry. These results suggest that the repetition of sound is

becoming a less aesthetically significant poetic device among contemporary masters of poetry.

Such a finding may have several implications on the role of sound in musicality in sophisticated poetry. For one, perfect end rhyme and alliteration seem to have fallen out of high art and have become second-hand devices that amateurs use in great quantities. Professional poets may be using other sound devices that our features do not capture such as meter and rhythm. This indicates that conventional perceptions of the musical and repetitive quality of poetry may no longer fit the styles and aspirations of contemporary professional poets.

6.2 Emotion

Poetry is the spontaneous overflow of powerful feelings: it takes its origin from emotion recollected in tranquility.

—William Wordsworth (1770—1850)

Do professional poems contain “stronger and more inspired emotion” than amateur poems? Are they more concerned with negative emotions, and do they make more references to psychological states?

Our results suggest that poems by professional poets are not more negatively emotional—at least not explicitly. On the contrary, amateur poets are significantly more likely to reference and describe negative emotions than professional poets. Our results reveal an interesting distinction between words with positive and negative outlooks and connotations (**Positiv/Negativ**) versus words that reference positive and negative emotions (**posEmo/negEmo**). While the two pairs of categories are strongly correlated, they capture different aspects of a text’s emotional content. **Positiv** and **Negativ** contain many words that are not emotions but may evoke certain emotional attitudes, such as *clean* and *death*. The fact that professional poets are significantly less likely to use explicitly negative emotion words than amateur poets, but not significantly less likely to use negatively connotative words, suggests that professional poets may evoke more negative emotion through connotation rather than explicit descriptions.

In terms of positive sentiment, professional poets are significantly less likely to make references to positive outlook words than amateur poets, but not significantly

less likely to use explicitly positive emotional words. This result may be due to the fact that amateur poets make many more references to the positive outlook word *love* than professional poets (134 counts versus 36). This suggests that amateur poets tend to write more often about positive topics, such as love, but describe more negative emotions, such as pain. It is possible that many amateurs write for therapeutic reasons and are inspired to produce poems only when they are experiencing especially strong negative emotions. Serious professional poets who make a career out of writing are likely to engage in a more diverse set of topics. Furthermore, while amateur poets often write for personal reasons, professional poets write with the intention of getting published. They are thus more likely to address issues that would interest their reader—politics, racial issues, and mortality—instead of personal accounts of romantic relationships and heartbreak.

One important implication of these results on sentiment analysis is that professional poems are emotional in ways that are not easily captured using existing sentiment analysis techniques. Professional poems may evoke strong emotions, but they do not explicitly describe emotions. As a result, sentiment lexicon categories, both connotative and explicitly emotional, do not show professional poems to be significantly more emotional than amateur poems. Instead, the emotions in professional poetry may have been “recollected into tranquility” and can only be recovered through a deeper understanding of the poem. Sentiment lexicons are not able to capture these subtle aspects of emotionally evocative texts. Other factors must be integrated into such an analysis, such as tone, imagery, and even pragmatics, such as the ability to infer that the poet would only use image *X* and avoid using word *Y* if he or she intended to express emotion *Z*.

Overall, our results show that amateur poets write more about positive topics than professional poets and are more likely to make explicit references to negative emotions and their expression.

6.3 Imagery

The job of the poet is to render the world—to see it and report it without loss, without perversion. No poet ever talks about feelings. Only sentimental people do.

—Mark Van Doren (1894—1972), winner of the Pulitzer Prize for Poetry

If professional poets are not using elevated diction, sophisticated sound devices, or even more emotionally salient words, what, then, allows poems by professional poets to leave such a deep aesthetic impression on readers?

The answer lies in concrete and specific imagery. As predicted, poems written by professional poets contain significantly more words that reference objects and significantly less words about abstract concepts and generalizations. This result suggests that professional poets follow the sacred rule of “show, don’t tell” and let images instead of words convey emotions, concepts, and experiences that stick to readers’ minds.

Imagism has strongly influenced the ways in which modern poets and literary critics think about literary writing. Literary critic I.A. Richards argued that image clusters and patterns of imagery are keys to deeper meaning in literary works, and that critics should pay close attention to these patterns in order to understand “the language of art” beneath the surface ordinary language (Richards, 1893). Not only are concrete images able to “render the world” in spectacular detail, they also provide windows into particular experiences on which readers can project their own perceptions and interpretations.

The types of objects that professional poets write about tend to be commonplace in everyday life. The word *tree* appeared the most often in professional poems (29 times), followed by *room* (20 times). Professional poets also tend to mention objects that one does not expect to see in poetry, such as *manure*, *rag*, and *corn*. In the famous poem *The Red Wheelbarrow* included at the beginning of this chapter, Imagist poet William Carlos Williams presents nothing but a clear image using the simplest language and sharpest detail. Yet “so much depends upon” these series of images that are both striking and ambiguous in their clarity.

Consistent with our predictions and with the aesthetic ideals of Imagism, professional poets also make significantly fewer direct references to abstract and intangible concepts. If the deeper meaning of a poem is conveyed through imagery, abstract words are no longer needed to reference concepts and experiences explicitly. Grief becomes dead flies sticking to a man’s fingers; love is what spring does with the cherry trees. Through imagery and metaphor, the skilled contemporary poet is able to pack large concepts into tiny objects to insert into the reader’s mind.

On the other hand, amateur poets use significantly more words concerned with generalizations. Amateur poets have the same impulse to make generalizations as most human beings. The skilled poet must repress this impulse and learn to report each object as it is—one-of-a-kind, with unique details that single out each object and experience from the rest.

Overall, professional poets are more likely to show, while amateur poets have a tendency to tell. This difference marks the most significant distinction between contemporary professional and amateur poetry and may be an essential aspect of craft and poetic beauty.

6.4 Summary

The key to good poetry is not fancy diction, sophisticated use of sound devices, or emotional salience. Rather, it is about using concrete and visualizable images to convey deeper concepts and emotions. While beautiful poetry is often associated with memorable patterns of sound, our analysis suggests that good contemporary poetry is more concerned with striking patterns of images. This marks a critical shift in a common conception of poetry and provides empirical evidence that establishes imagery as one of the most distinguishing elements of craft in contemporary poetry, as well as the most appealing to the aesthetic of our time.

Chapter 7

Future Directions

You can tear a poem apart to see what makes it tick. . . You're back with the mystery of having been moved by words. The best craftsmanship always leaves holes and gaps. . . so that something that is not in the poem can creep, crawl, flash or thunder in.

—Dylan Thomas (1914 – 1953), *Poetic Manifesto*

7.1 General strengths and weaknesses

This thesis presents a novel computational framework for analyzing elements of craft commonly associated with poetic beauty. By applying both stylistic and content analyses to the quantitative assessment of contemporary poetry, we were able to test and challenge common perceptions of poetic craft on a representative set of poems.

This thesis makes several contributions to natural language processing and creative writing. It is a simple but innovative attempt to computationally analyze beauty in poetic language, an area in which very little empirical work has been done. The application of sentiment analysis tools to poetry was a novel idea that revealed the limitations of sentiment lexicons when dealing with literary texts, which are often rich with evocative metaphor and imagery.

In terms of contributions to creative writing and literary criticism, this thesis provides an automated way to analyze the quality of poetry based on critical elements of craft. It helps identify important features of good contemporary poetry and facilitates the understanding of current standards and criteria for poetic beauty. In

the spirit of digital humanities, a growing interdisciplinary field aimed at using computational methods to analyze traditional disciplines in the humanities, this thesis serves as an additional tool to support the analysis and understanding of verbal art.

This analysis can be further improved and augmented in many ways. For example, word frequencies and type token ratios may not be sufficient measures of diction sophistication in poetry. More complex features may be necessary, such as bigram or trigram frequencies that capture not only what words poets use, but also how they use them to create surprising phrasings and sentence structures. Parse features could also be incorporated to measure linguistic and syntactic complexity, such as average noun and verb phrases and parse tree height (Pitler & Nenkova, 2008; Feng et al., 2010). However, implementing these features may be challenging because poetry often breaks away from ordinary syntactic rules.

In terms of sound devices, meter was not included as a feature. It may, however, be a critical part of poetic craft that distinguishes professional poets' use of sound from that of amateur poets. Given the challenges of designing automated ways to mark and quantify poetic meter (Hayward, 1996), such a feature was not implemented in this thesis but could be included in future extensions.

The importance and structure of poetic imagery can also be further explored. Since most concrete objects are highly imageable, we used the **Object** feature to approximate imageability. However, an imageability measure such as that provided by the MRC Psycholinguistic Database (Wilson et al., 1987) would more directly capture the presence of imagery in poems. Furthermore, instead of looking only at frequencies of references to concrete imagery, it may be helpful to model the patterns in which these images appear in order to reveal interesting ways in which skilled poets set up image structures and themes.

7.2 Future Extensions

In addition to improving the feature set in the ways described above, other extensions may help increase the scope and influence of the work presented here.

In order to test the robustness and generalizability of our results, the features could be tested on a larger set of poems. It would be interesting to build a classifier based on the logistic regression model that can automatically distinguish between a

separate set of professional and amateur poems. Not only would this help validate the features and model proposed in this thesis, such a classifier could also be useful for sorting through large amounts of amateur poetry and identifying ones that may reflect superior poetic skill. Since online fiction and poetry writing forums are becoming an increasingly popular medium for aspiring writers and poets (Lawrence & Schraefel, 2006; Bode, 2008), an automated way to identify beautiful poems may help improve the visibility of unknown but talented writers.

Another extension would be to ask people to rate a set of poems on measures of beauty. Based on their ratings, we could build a regression model using the features we implemented to predict reader preferences. Mixed-effect models could then be used to test additional effects, for example whether or not people find certain poets' work more appealing than others.

Furthermore, literary critics and laypeople may have very different tastes for poetry. Features that predict expert and laypeople's ratings of poems could be compared in order to identify the different ways in which they form judgments of poetic beauty. This could contribute to our understanding of how expert knowledge affects perception and appreciation of art (Hekkert & Wieringen, 1996; Kao et al., 2010).

In terms of using this tool to complement the qualitative analysis of poetic beauty, future work could analyze professional poetry in different eras and track how standards and criteria for good poetry have changed through time. Modern amateur poetry could be compared to past professional poetry to examine whether or not current amateur poems are in fact more similar to professional poetry in the past than they are to contemporary professional poetry. This would suggest that while contemporary professional poets have broken away from traditional forms and created new aesthetic ideals, amateur poets are a few steps behind and tend to follow literary traditions that have turned into clichés.

An application of this work to the psychological study of aesthetics would be to examine why certain features contribute to the perception of verbal beauty. Why are concrete images appealing to readers? What sorts of criteria must be present in order for an image to be psychologically meaningful and evocative? Answering these questions can help us understand the cognitive mechanisms responsible for the appeal of important and distinguishing literary tropes.

This thesis is a first step towards examining the aesthetic functions of language using computational methods. I have torn the poem apart to see what makes it tick, and some of the pieces are here before you. Although this is a humble contribution to understanding language's aesthetic capacity, I present it with the hope that something will creep or thunder in to fill in the gaps so that we may be able to understand the workings of language at its most distilled and most powerful.

Bibliography

- Academy of American Poets (2011). Poets, <http://www.poets.org>.
- Addonizio, K., & Laux, D. (1997). *The Poet's Companion: A guide to the pleasures of writing poetry*. W. W. Norton and Company.
- Aristotle (1998). Poetics. *The Critical Tradition: Classical Texts and Contemporary Trends*.
- Baccianella, S., Esuli, A., & Sebastiani, F. (2010). Sentiwordnet 3.0: An enhanced lexical resource for sentiment analysis and opinion mining. In *Seventh Conference on International Language Resources and Evaluation*, (pp. 2200–2204).
- Beebe-Center, J. G., & Pratt, C. C. (1937). A test of birkhoff's aesthetic measure. *The Journal of General Psychology*, 17(2), 339–353.
- Ben-Simon, A., & Bennett, R. E. (2007). Toward more substantively meaningful automated essay scoring. *The journal of Technology, Learning, and Assessment*.
- Best, J. (1992). *Cognitive Psychology*. West Publishing Company.
- Birkhoff, G. (1933). *Aesthetic Measure*. Kessinger Publishing.
- Blakeney Willaims, L. (2002). *Modernism and the Ideology of History: Literature, Politics, and the Past*. Cambridge University Press.
- Bode, D. (2008). And now, a word from the amateurs. *Transformative works and cultures*, 1, 89–92.
- Bower, G. (1970). Imagery as a relational organizer in associative learning. *Journal of Verbal Learning and Verbal Behavior*, 9(5), 529–533.

- Bowman, D. O., & Halfacre, D. L. (1994). Poetry therapy with the sexually abused adolescent: A case study. *The Arts in Psychotherapy*, 21(1), 11–16.
- Breland, H. M. (1996). Word frequency and word difficulty: A comparison of counts in four corpora. *Psychological Science*, 7(2), pp. 96–99.
- Bryant, P., Maclean, M., Bradley, L., & Crossland, J. (1990). Rhyme and alliteration, phoneme detection, and learning to read. *Developmental Psychology*, 26(3).
- Burroway, J. (2007). *Imaginative Writing: The Elements of Craft*. Pearson, 2 ed.
- Burstein, J., Chodorow, M., & Leacock, C. (2004). Automated essay evaluation: The criterion online writing service. *AI Magazine*, 25(3).
- Campbell, S. R., & Pennebaker, J. W. (2003). The secret life of pronouns: flexibility in writing style and physical health. *Psychological Science*, 14(1), 6065.
- Davies, M. (2011). Word frequency data from the corpus of contemporary american english (coca). Downloaded from <http://www.wordfrequency.info> on May 10, 2011.
- Earnshaw, S. (Ed.) (2007). *The Handbook of Creative Writing*. Edinburgh University Press.
- Fellbaum, C. (1998). Wordnet: An electronic lexical database. *Bradford Books*.
- Feng, L., Jansche, M., Huenerfauth, M., & Elhadad, N. (2010). A comparison of features for automatic readability assessment. In *Proceedings of the 23rd International Conference on Computational Linguistics*, (pp. 276–284).
- Fink, B., & Penton-Voak, I. (2002). Evolutionary psychology of facial attractiveness. *Current Directions in Psychological Science*, 11(5), 154–158.
- Gentner, D. (2010). Psychology in cognitive science: 1978-2038. *Topics in Cognitive Science*, 2(3), 328–344.
- Graves, M. F., Ryder, R. J., Slater, W. H., & Calfee, R. C. (1987). The relationship between word frequency and reading vocabulary using six metrics of frequency. *The Journal of Educational Research*, 81(2), pp. 81–90.

- Grinde, B. (1996). The biology of visual aesthetics. *Journal of Social and Evolutionary Systems*, 19(1), 31–40.
- Harrower, M. (1972). *The therapy of poetry*. Oryx, London.
- Hayward, M. (1996). Analysis of a corpus of poetry by a connectionist model of poetic meter. *Poetics*, 24(1), 1–11.
- Hekkert, P. (2006). Design aesthetics: principles of pleasure in design. *Psychology Science*, 48, 157–172.
- Hekkert, P., & Wieringen, P. C. W. V. (1996). Beauty in the eye of expert and nonexpert beholders: A study in the appraisal of art. *The American Journal of Psychology*, 109(3), pp. 389–407.
- Inquirer (2011). How the general inquirer is used and a comparison of general inquirer with other text-analysis procedures.
URL <http://www.wjh.harvard.edu/inquirer/3JMoreInfo.html>
- Jamison, K. (1993). *Touched with fire: manic-depressive illness and the artistic temperament*. New York: Free Press.
- Jessen, F., Heun, R., Erb, M., Granath, D. O., Klose, U., Papassotiropoulos, A., & Grodd, W. (2000). The concreteness effect: Evidence for dual coding and context availability. *Brain and Language*, 74(1), 103 – 112.
- Jurafsky, D., Ranganath, R., & McFarland, D. (2009). Extracting social meaning: Identifying interactional style in spoken conversation. In *Proceedings of NAACL HLT*.
- Kao, J., Ryan, R., Dye, M., & Ramscar, M. (2010). An acquired taste: how reading literature affects sensitivity to word distributions when judging literary texts. In *Annual meeting of the Cognitive Science Society*, (pp. 990–995).
- Kaplan, D. (2006). Computational analysis and visualized comparison of style in american poetry. Unpublished undergraduate thesis.
- Kaplan, D., & Blei, D. (2007). A computational approach to style in american poetry. In *IEEE Conference on Data Mining*.

- Kennedy, A., & Inkpen, D. (2006). Sentiment classification of movie reviews using contextual valence shifters. *Computational Intelligence*, 22, 110–125.
- Lawrence, K. F., & Schraefel, M. C. (2006). Web based semantic communities: Who, how and why we might want them in the first place. In *Web Based Communities 2006*.
- Lea, R., Rapp, D., Elfenbein, A., Mitchel, A., & Romine, R. (2008). Sweet silent thought: Alliteration and resonance in poetry comprehension. *Psychological Science*, 19(709).
- Liu, H., & Maes, P. (2005). The aesthetoscope: visualizing aesthetic readings of text in color space. *International Journal on Artificial Intelligence*.
- Longinus (2001). On sublimity. *The Norton Anthology of Theory and Criticism*.
- Lowell, A. (1917). *Tendencies in Modern American Poetry*. New York: Macmillan Company.
- Marks, Carolyn B., Doctorow, Marleen J., & Wittrock, M. C. (1974). Word frequency and reading comprehension. *The Journal of Educational Research*, 67(6), 259–262.
- Mcardle, S., & Byrt, R. (2001). Fiction, poetry and mental health: expressive and therapeutic uses of literature. *Journal of Psychiatric and Mental Health Nursing*, 8(6), 517–524.
- McGlone, M., & Tofighbakhsh, J. (2000). Birds of a feather flock conjointly (?): Rhyme as reason in aphorisms. *Psychological Science*, 11, 424–428.
- Miles, J. (1946). Major adjectives in english poetry: From wyatt to auden. *University of California Publications in English*, 12(3), 305–426.
- Miles, J. (1967). *Style and Proportion: The Language of Prose and Poetry*. Little, Brown and Co., Boston.
- Nell, V. (1988). *Lost in a book: The psychology of reading for pleasure*. New Haven, CT. US: Yale University Press.

- O'Connor, B., Balasubramanyan, R., Routledge, B., & Smith, N. (2010). From tweets to polls: Linking text sentiment to public opinion time series. In *International AAAI Conference on Weblogs and Social Media*.
- Paivio, A., Yuille, J., & Smythe, P. (1966). Stimulus and response abstractness, imagery, and meaningfulness, and reported mediators in paired-associate learning. *Canadian Journal of Psychology*, 20(4).
- Pang, B., & Lee, L. (2008). Opinion mining and sentiment analysis. *Foundations and trends in information retrieval*, 2(1), 1 – 135.
- Pang, B., Lee, L., & Vaithyanathan, S. (2002). Thumbs up? sentiment classification using machine learning techniques. In *Proceedings of the Conference on Empirical Methods in Natural Language Processing (EMNLP)*, (pp. 79–86).
- Pennebaker, J., Francis, M., & Booth, R. J. (2001). *Linguistic Inquiry and Word Count (LIWC): LIWC2001*. Mahwah, NJ: Erlbaum.
- Perkins, D. (1987). *A history of modern poetry: modernism and after*. Harvard University Press.
- Pitler, E., & Nenkova, A. (2008). Revisiting readability: A unified framework for predicting text quality. In *Empirical Methods in Natural Language Processing*, (pp. 186–195).
- Plato (2003). *Republic*. London: Penguin.
- Poulin, A., & Waters, M. (2006). *Contemporary American Poetry*. Houghton Mifflin Company, eighth ed.
- Ramirez-Esparza, N., Chung, C., Kacewicz, E., & Pennebaker, J. (2008). The psychology of word use in depression forums in english and in spanish: Testing two text analytic approaches. In *Int'l AAAI Conference on Weblogs and Social Media (ICWSM) 2008*.
- Ranganath, R., Jurafsky, D., & McFarland, D. (2009). It's not you, it's me: Detecting flirting and its misperception in speed-dates. In *Proceedings of EMNLP 2009*.

- Reber, R., Schwarz, N., & Winkielman, P. (2004). Processing fluency and aesthetic pleasure: is beauty in the perceiver's processing experience? *Personality and Social Psychology Review*, 8(4).
- Richards, I. (1893). *Practical criticism: a study of literary judgment*. Transaction Publishers.
- Rigau, J., Feixas, M., & Sbert, M. (2008). Informational aesthetics measures. In *IEEE Computer Graphics and Applications*.
- Rubin, D. (1995). *Memory in oral traditions: The cognitive psychology of epic, ballads, and counting-out rhymes*. New York: Oxford University Press.
- Silverman, M., & Will, N. (1986). Sylvia plath and the failure of emotional self-repair through poetry. *Psychoanal Q*, 55, 99–129.
- Smith, J., & Melar, R. (1990). Aesthetic preference and syntactic prototypicality in music: 'tis the gift to be simple. *Cognition*, 34(3).
- Smith, J. D., & Melara, R. J. (1990). Aesthetic preference and syntactic prototypicality in music: 'tis the gift to be simple. *Cognition*, 34(3), 279 – 298.
- Stirman, S. W., & Pennebaker, J. (2001). Word use in the poetry of suicidal and nonsuicidal poets. *Psychosomatic Medicine*, 63(4), 517–22.
- Stone, P., Dunphy, D., Smith, M., & Ogilvie, D. (1966). *The General Inquirer: A Computer Approach to Content Analysis*. Cambridge, MA: MIT Press.
- Thomas, M., Pang, B., & Lee, L. (2006). Get out the vote: determining support or opposition from congressional floor-debate transcripts. In *Proceedings of EMNLP*, (pp. 327–335).
- Tomayo, J. (1987). Frequency of use as a measure of word difficulty in bilingual vocabulary test construction and translation. *Educational and Psychological Measurement*, 47, 893–902.
- Wilson, M., Wilson, M., Qx, O. O., & Quinlan, P. (1987). Mrc psycholinguistic database: Machine usable dictionary, version 2.00.

Zaidel, D., & Cohen, J. (2005). The face, beauty, and symmetry: perceiving asymmetry in beautiful faces. *International Journal of Neuroscience*, 115(8), 1165–1173.

Zwicky, J. (2000). *Wisdom and metaphor*. Kentville, NS: Gaspereau Press.