

Electronic Thesis and Dissertation Repository

---

2-11-2021 12:15 PM

## Speaking Songs: Music-Analytical Approaches to Spoken Word

Chantal D. Lemire, *The University of Western Ontario*

Supervisor: Cuciurean, John, *The University of Western Ontario*

A thesis submitted in partial fulfillment of the requirements for the Doctor of Philosophy degree in Music

© Chantal D. Lemire 2021

Follow this and additional works at: <https://ir.lib.uwo.ca/etd>



Part of the [Music Theory Commons](#)

---

### Recommended Citation

Lemire, Chantal D., "Speaking Songs: Music-Analytical Approaches to Spoken Word" (2021). *Electronic Thesis and Dissertation Repository*. 7670.

<https://ir.lib.uwo.ca/etd/7670>

This Dissertation/Thesis is brought to you for free and open access by Scholarship@Western. It has been accepted for inclusion in Electronic Thesis and Dissertation Repository by an authorized administrator of Scholarship@Western. For more information, please contact [wlsadmin@uwo.ca](mailto:wlsadmin@uwo.ca).

## Abstract

If we can conceive of music as performance—indeed, if we tend to agree with director and drama theorist Richard Schechner that nearly anything can be studied ‘as’ performance—then it follows that nearly any mode of performance might also be studied for its musicality. Of course, some modes of performance are more conducive to musical study than others. The present work concerns a particularly responsive mode of performance through which the categorical divisions between language and music begin to dissolve: *spoken word*.

Spoken word, in its diversity of forms, traditions, and styles, exists not simply on the fringes of any single scholarly field but at the intersections of multiple artistic disciplines. Despite the many interconnections between spoken-word traditions and musical practices, the former traditionally belong to the domains of literary or cultural studies, dominated by discussions of cultural significance or ambiguities pertaining to its literary status. While scholars of a variety of spoken-word forms ardently contend that the work *is* its performance, few offer insights on how to integrate the sound structures realized through performance into analyses of these works.

Focusing on three distinct forms of spoken word—spoken-word song, slam poetry, and sound poetry—I investigate the rhythmic structures, melodic patterns, formal design, and communicative strategies of selected spoken-word performances using musically centred analytical methods, including recording analysis, formal functions, and motivic analysis. Featuring spoken-word song performances by American singer-songwriter Tom Waits, slam poetry by the founder of the modern poetry slam, Marc Kelly Smith, and the sound



poetry of avant-garde artist Kurt Schwitters, this dissertation endeavors not only to offer an array of analytical approaches and listening strategies that can provide a basis for future investigations of speech-based contemporary art music; it also strives to recognize the categorical overlap that exists among Spoken Word and music, and to bring Spoken Word further into the view of mainstream musicological research and analysis.

## Keywords

Performance Poetry, Performance Analysis, Music and Poetry, Tom Waits, Marc Kelly Smith, Kurt Schwitters.

## Summary for Lay Audience

If we can conceive of music as performance, then it is possible that any mode of performance might also be studied for its musicality. While some modes of performance are more conducive to musical study than others, the present work concerns a particularly responsive mode of performance through which the boundaries between language and music begin to blur, namely, *Spoken Word*. Spoken Word bridges multiple artistic disciplines, yet scholarly research on Spoken Word traditionally occurs within the domains of literary or cultural studies, dominated by discussions of cultural significance or literary merit. And while scholars of spoken-word forms ardently contend that the work *is* its performance, few offer insights on how to integrate the sound structures realized through performance into analyses of these works.

Focusing on three distinct forms of spoken word—spoken-word song, slam poetry, and sound poetry—I explore the musical properties of spoken word. I home in on the rhythmic structures, melodic patterns, formal design, and communicative strategies of selected spoken-word song performances by American singer-songwriter Tom Waits, selected slam poetry performances by the founder of the modern poetry slam, Marc Kelly Smith, and a particularly famous sound poem by avant-garde artist Kurt Schwitters. This dissertation offers an array of analytical approaches and listening strategies that can provide a basis for future investigations of speech-based contemporary art music; it also strives to recognize the categorical overlap that exists among Spoken Word and music, and to bring Spoken Word further into the view of mainstream research and analysis.

## Acknowledgments

This dissertation was written in several locations across the Indigenous lands of Turtle Island and I wish to acknowledge and pay respect to the traditional owners of the land on which I conducted this research. The majority was written at Western University, located on the traditional lands of the Anishinaabek, Haudenosaunee, Lūnaapéewak, and Attawandaron people. During the portions written in and around Owen Sound, Ontario, I fell in love with the awesome beauty of the shores of Lake Huron and am so grateful to have been a guest in traditional Ojibway territory. I offer my gratitude and respect to the communities of the Chippewas of Saugeen First Nation and Chippewas of Nawash First Nation for their continued stewardship of this land. Other portions of this dissertation were composed in Montreal, Quebec, unceded Indigenous lands historically known as Tiohtià:ke, a gathering place for many First Nations. Many Indigenous peoples, communities, and Nations helped to build the foundation of Tiohtià:ke. In particular, the Kanien'kehá:ka have a strong historic and ongoing presence in the territory with two communities bordering Montréal: Kahnawà:ke and Kanehsatá:ke. These are the original, rightful carers and protectors of the lands and waters we now share and enjoy today. Still other portions of this dissertation were written in in my hometown of Langley, British Columbia. It was an extraordinary privilege to grow up in the region south of the Fraser River, which overlaps with the unceded traditional and ancestral lands of the Kwantlen, Snokonomish, Musqueam, Katzie, Semiahmoo, Tsawwassen, Qayqayt and Kwikwetlem peoples.

My sincere thanks to Dr. John Cuciurean, whose advice and encouragement along the way gave me the confidence to explore new and different analytical approaches and helped

me to draw together my eclectic interests and ideas. I am grateful also to Dr. Catherine Nolan for her guidance and support, and to Dr. Alan Dodson for teaching me that you can't rush genius. I am especially grateful to Harald Krebs, Gary Geddes, and Betty Anne Younker for all the inspiration they provided me along the way. Thanks to Tom Adam from Copyright at Western for helping me understand fair dealing. Special thanks to Western Libraries for waiving so many late fines.

The following people were essential to the formation of this work and their love and support mean the world to me: Martin Ross, Kristen Wallentinsen, Margie Bernal, Grant Sawatzky, Elizabeth Mitchell, Mary Blake Bonn, April Morris, Sebastian Ostertag—you were with me for every word in this dissertation. To the Bruno family, the Ostertag-Purdon family, the Dowling family, and the Dykstra family, thank you all for your love and cheerleading, which was always loudest when I needed it most. I am so grateful to my sisters, Sarah and Christina, for being the best human beings on earth, and to all the other sisters I've collected along the way: Catherine, Elena, Katie, Isla, Yolanda, Adriana, Sonya, Roxi, Gloria, Danielle, Nancy, Robin, Tara, Sue, Vonica, Diana, Kelsey, Al, Shrub, Sarah, Sam, and all the other incredible women that continue to inspire me.

Finally, my deepest and most heartfelt thanks to my mother, Judi Paget. I am humbled by her enormous generosity and unwavering faith in me, and immensely grateful for her love and support. In every way, this dissertation truly belongs to her.

# Table of Contents

Abstract.....	ii
Summary for Lay Audience .....	iv
Acknowledgments .....	v
Table of Contents .....	vii
List of Figures.....	ix
List of Appendices.....	xiv
Epigraph .....	xv
Chapter 1 .....	1
1 Introduction to Spoken Word .....	1
1.1 Definitions .....	2
1.1.1 Voiced Texts and Oral Performance .....	4
1.1.2 Spoken Word and/as Musical Material .....	5
1.2 Genres, Styles, and Types .....	7
1.3 Speakers, Slams, and Sounds .....	15
Chapter 2 .....	22
2 The Singer Speaks .....	22
2.1 Tom Waits’s Spoken-Word Songs .....	23
2.2 Recording Analysis .....	26
2.3 Speaking in/of Time .....	31
2.4 Speaking in/a Tune .....	50
2.5 Speaking Live.....	70
2.6 Conclusions .....	75
Chapter 3 .....	79
3 The Poet Performs .....	79
3.1 Slam Poetry and/as Music .....	80
3.2 The Ethnopoetic Approach.....	86
3.3 Reforming Formal Functions .....	89
3.4 Formal Structure of a Slam Poem .....	98
3.5 Speaking of/in Sonata Form .....	106
3.6 Conclusions .....	124
Chapter 4 .....	128

4	The Artist Composes .....	128
4.1	Kurt Schwitters .....	130
4.2	<i>Die Ursonate</i> (1922–1932) .....	132
4.2.1	Introduction and Rondo, 1 <sup>st</sup> Movement .....	133
4.2.2	Largo, 2 <sup>nd</sup> Movement .....	134
4.2.3	Scherzo and Trio, 3 <sup>rd</sup> Movement .....	136
4.2.4	Sonata (presto) with Cadenza, 4 <sup>th</sup> Movement .....	137
4.3	Die Scherzo (und trio) .....	140
4.3.1	Musical Form .....	140
4.3.2	Motivic Fragments .....	144
4.3.3	A Visual Collage Score .....	153
4.3.4	An Audio Collage .....	159
4.4	Conclusions .....	166
	Chapter 5 .....	169
5	Last Words .....	169
5.1	Implications for Cognitive Research .....	170
5.2	Avenues for Further Research and Analysis .....	173
5.2.1	Spoken-Word Songs .....	174
5.2.2	Performance and Slam Poetry .....	178
5.2.3	Sound Poetries .....	179
5.3	Last Last Words .....	182
	Appendices .....	185
	Bibliography .....	193
	Discography .....	205
	Curriculum Vitae .....	206

## List of Figures

Figure 1.1 Table of media categories for oral poetry by John Miles Foley. ....	5
Figure 1.2 Relational diagram of voiced text genres.....	8
Figure 1.3 Hugo Ball’s <i>Karawane</i> (1917).....	14
Figure 2.1 Log-frequency spectrograms generated by Sonic Visualiser for an excerpt from “Dave Brubeck” (1959) by Jack Kerouac and Steve Allen. ....	28
Figure 2.2 Peak frequency spectrogram generated by Sonic Visualiser for an excerpt from “Dave Brubeck” (1959) by Jack Kerouac and Steve Allen.....	29
Figure 2.3 Text and musical sections for "Watch Her Disappear" by Tom Waits.....	34
Figure 2.4 Two instances of word painting in "Watch Her Disappear" by Tom Waits. ...	35
Figure 2.5 Onset time-point graph and alignment discrepancies in line 7 from “Watch Her Disappear” by Tom Waits (0:37–0:43). ....	37
Figure 2.6 Onset time-point graph and alignment discrepancies in line 1 from “Watch Her Disappear” by Tom Waits (0:08–0:15). ....	38
Figure 2.7 Onset time-point graph and alignment discrepancies in line 2 from "Watch Her Disappear" by Tom Waits (0:15–0:21).....	40
Figure 2.8 Normative metrical stress grid for line 2 from “Watch Her Disappear” by Tom Waits.....	43
Figure 2.9 Alternate metrical stress grid with stress shifts for line 2 from “Watch Her Disappear” by Tom Waits. ....	43
Figure 2.10. “Rubin’s Vase,” afbildn 3 by Edgar Rubin.....	44
Figure 2.11 Comparison of speech rhythm and melodic rhythm for lines 14-15 from “Watch Her Disappear” by Tom Waits (1:15–1:20). ....	46

Figure 2.12 Onset time-point graph and alignment discrepancies in lines 14–15 from “Watch Her Disappear” by Tom Waits (1:15–1:20). .....	46
Figure 2.13. <i>Turning the Tables</i> (1990) by Roger Shepard.....	48
Figure 2.14 Annotated spectrogram and melodic/harmonic reduction for mm. 12–18 from “Crossroads” by Tom Waits (0:19–0:31). .....	53
Figure 2.15 Summary of phrase structure for “The Ghosts of Saturday Night” by Tom Waits.....	57
Figure 2.16 Table showing text and musical form for "The Ghosts of Saturday Night (After Hours at Napoleone's Pizza House)" by Tom Waits. ....	59
Figure 2.17 Annotated spectrogram for verse 1 from “The Ghosts of Saturday Night” by Tom Waits (0:19–0:37). ....	61
Figure 2.18 Reduction of opening verse from “The Ghosts of Saturday Night” by Tom Waits.....	62
Figure 2.19 Spectrogram and harmonic reduction of verse two, "The Ghosts of Saturday Night" by Tom Waits (0:36–0:56).....	63
Figure 2.20 Spectrogram and melodic/harmonic reduction for “The Ghosts of Saturday Night,” verse 3A (0:57–1:09). ....	65
Figure 2.21 Spectrogram and melodic/harmonic reduction for verse 3B from "Ghosts of Saturday Night" by Tom Waits (1:20–1:39). ....	67
Figure 2.22 Full melodic/harmonic reduction of “Ghosts of Saturday Night,” studio recording.....	69
Figure 2.23 Melodic similarities between line 11 from live and studio recordings of “The Ghosts of Saturday Night” by Tom Waits.....	71



Figure 2.24 Melodic similarities between line 2 from live and studio recordings of “The Ghosts of Saturday Night” by Tom Waits.....	72
Figure 2.25 Melodic similarities between lines 7–8 from live and studio recordings of “The Ghosts of Saturday Night” by Tom Waits.....	73
Figure 2.26 Longer excerpt comparing melodic inflections in lines 12-13 from live and studio recordings of “The Ghosts of Saturday Night” by Tom Waits.....	74
Figure 3.1 Comparison of musical and poetic sentences. ....	94
Figure 3.2 Table of Generalized Definitions for Formal Functions. ....	97
Figure 3.3 Opening two verses of Smith's "Kiss It," print format.....	99
Figure 3.4 Form-functional analysis of verse 1 from “Kiss It” by Marc Kelly Smith. ....	99
Figure 3.5 Form-functional analysis of verse 2 from “Kiss It” by Marc Kelly Smith. ...	101
Figure 3.6 Final verses of “Kiss It” by Marc Kelly Smith, print format. ....	102
Figure 3.7 Form-functional analysis of verse 3 from “Kiss It” by Marc Kelly Smith. ...	103
Figure 3.8 Full binary form analysis for "Kiss It" by Marc Kelly Smith. ....	105
Figure 3.9 Main Theme of “I Wear My Father’s Coat” by Marc Kelly Smith. ....	107
Figure 3.10 Transition of “I Wear My Father’s Coat” by Marc Kelly Smith. ....	109
Figure 3.11 Subordinate theme from “I Wear My Father’s Coat” by Marc Kelly Smith. ....	112
Figure 3.12 Development from “I Wear My Father’s Coat” by Marc Kelly Smith.....	116
Figure 3.13 Comparison of transition material from “I Wear My Father’s Coat” by Marc Kelly Smith.....	121

Figure 3.14 Recapitulation (with MT/trans. Fused) from “I Wear My Father’s Coat” by Marc Kelly Smith. ....	122
Figure 3.15 Truncated recapitulation from “I Wear My Father’s Coat” by Marc Kelly Smith.....	123
Figure 4.1 Annotated score of <i>Die Ursonate</i> , 1st movement, with colors indicating motivic segments of themes 1–4. ....	134
Figure 4.2 Example of Kurt Schwitters’s musical notations for lines 1–4 of <i>Die Ursonate</i> , 2nd movement. ....	135
Figure 4.3 Musical interpretation of Schwitters's rhythmic notation for <i>Die Ursonate</i> , 2nd movement. ....	136
Figure 4.4 Kurt Schwitters, <i>Die Ursonate</i> , 4th movement, coda.....	138
Figure 4.5 Annotated score of the scherzo from <i>Die Ursonate</i> , 3 <sup>rd</sup> movement, indicating the formal structure.....	142
Figure 4.6 Annotated print format showing formal structure of the trio from <i>Die Ursonate</i> , 3 <sup>rd</sup> movement. ....	143
Figure 4.7 Notated rhythm of motivic fragment III. ....	144
Figure 4.8 Beethoven Piano Sonata No. 29, Op. 106, ii, mm. 1–14. ....	145
Figure 4.9 Notated rhythms of motivic fragments 8a and 9a. ....	146
Figure 4.10 Chopin Scherzo No. 3, op. 39, mm. 1–4. ....	146
Figure 4.11 Notated rhythms of motivic fragments 8b and 9b from <i>Die Ursonate</i> , 3 <sup>rd</sup> movement. ....	147
Figure 4.12 Beethoven Piano Sonata No. 15, Op. 28, iii, mm.1–16. ....	147

Figure 4.13 Notated rhythms of motivic fragments III and 4a from <i>Die Ursonate</i> , 3 <sup>rd</sup> movement. ....	148
Figure 4.14 Beethoven Piano Sonata No. 29, Op. 106, ii, mm. 5–7 and mm. 12–14. ....	149
Figure 4.15 Notated rhythm of motivic fragment 3, “ziiuu lenn trll?” from <i>Die Ursonate</i> , 3 <sup>rd</sup> movement. ....	149
Figure 4.16 Chopin Scherzo No. 4, Op. 54, mm.1–25. ....	150
Figure 4.17 Notated rhythm of Motive 10, “Lümpff tümpff trll” from <i>Die Ursonate</i> , 3 <sup>rd</sup> movement. ....	151
Figure 4.18 Chopin Scherzo No. 4, Op. 54, mm. 197–201. ....	151
Figure 4.19 Notated pitch inflections for motive 4, “Rrumpff tillff too” from <i>Die Ursonate</i> , 3 <sup>rd</sup> movement, as recited by E. Schwitters 1954. ....	152
Figure 4.20 Visual representation of musical collage allusions in <i>Scherzo</i> , overlaying digital fragments of paper materials used in the 1922 <i>Merz</i> collage, <i>Er</i> (He), by Kurt Schwitters with score fragments from Beethoven Piano Sonata No. 15, Op. 28, iii and No. 29, Op. 106, ii; and Chopin Scherzo No. 3, op. 39 and No. 4, Op. 54. ....	154
Figure 4.21 <i>Er</i> (He) by Kurt Schwitters, 1922. ....	157
Figure 4.22 Audio Collage of <i>Die Ursonate</i> , 3 <sup>rd</sup> movement, Scherzo A section. ....	162
Figure 4.23 Audio Collage of <i>Die Ursonate</i> , 3 <sup>rd</sup> movement, Scherzo B section. ....	164

## List of Appendices

Appendix A. Table of selected spoken-word songs, focusing on songs by popular music artists for further research and analysis. ....	185
Appendix B. Table of selected slam and performance poems of interest for further research and analysis. ....	189
Appendix C. Table of selected sound poems and text-sound compositions for further research and analysis. ....	191

## Epigraph

I wish that I had spoken only of it all.

*Stanzas in Meditations*  
*Part V, Stanza LXIII*

(1932)

Gertrude Stein (1874–1946)

## Chapter 1

### 1 Introduction to Spoken Word

This dissertation comprises the very thing that it is about: words. Why so many words about words, the reader may wonder, when this dissertation's purpose is to fulfill not the requirements of a doctoral degree in literature or linguistics, but in music theory? Singer-songwriter Tom Waits provides the best answer I can think of because, he says, "words are music really. I mean, people ask me, 'do you write music, or do you write words?' but it's all one thing at its best."<sup>1</sup> This dissertation explores how a handful of musicians, poets, and artists make music out of words (or in one case, word-like sounds). Inspiration to write on this subject emerged from a game I have long enjoyed playing where, upon hearing a string of sounds I ask myself, "what if I were to call this music?" I find this game intellectually and philosophically rewarding when played with nearly any sonic stimulus but my favourite always seems to be words. I particularly enjoy, for example, playing this game at the bar while listening to a friend tell a story. I like to imagine I am audience to a live improvisatory word concerto—my friend the soloist and the blending din of conversation, music, and clinking glassware the unsuspecting instruments of a symphony orchestra unlike any other.

Another inspiration (or perhaps better put, motivation) for writing about words has to do with being a music theorist. What seems to me to be the main instrument music theorists

---

<sup>1</sup> Amanda Petrusich and Tom Waits, "Interview: Tom Waits," in *Tom Waits on Tom Waits: Interviews and Encounters*, ed. Paul Jr. Maher (Chicago: Chicago Review Press, 2011), 418.

play is, in fact, words. We write them, speak them, define and redefine them, and we are constantly making them up. So, it interests me professionally to explore how artists of various disciplines make music out of words because I might learn something to make my words about music more musical themselves.

## 1.1 Definitions

When I say that this dissertation is about words, I ought to clarify that it is not—at least not focally—about how words are written or appear on a page. Rather, this dissertation is about how words sound. Not unlike those fluent in reading musical notation, those fluent in a language hear words sounding even when the sounds themselves are not made manifest. Comedian, poet, and actor Bo Burnham captures this experience beautifully in his poem, *Magic*:

Read this to yourself. Read it silently.  
 Don't move your lips. Don't make a sound.  
 Listen to yourself. Listen without hearing anything.  
 What a wonderfully weird thing, huh?

NOW MAKE THIS PART LOUD!  
 SCREAM IT IN YOUR MIND!  
 DROWN EVERYTHING OUT.  
 Now, her a whisper. A tiny whisper.

Now, read this next line with your best crotchety old-man voice:  
 "Hello there, sonny. Does your town have a post office?"  
 Awesome! Who was that? Whose voice was that?

It sure wasn't yours!  
 How do you do that?  
 How?!

Must be magic.<sup>2</sup>

---

<sup>2</sup> Bo Burnham, *Egghead: Or, You Can't Survive on Ideas Alone* (New York: Grand Central Publishing, 2013), 1.

Burnham calls it magic; I want to call it music. While Burnham's poem certainly exploits the way words appear on the page to influence how a reader imagines them as they read, what interests me about this poem is that it is written to be read but is *about* how it sounds. Many writers share Burnham's interest in how writing evokes sonic characteristics. Fictional novelist Stephen King alludes not just to the sound of his writing but the music of it when he discusses his editing process. In *On Writing*, King tells his audience that the big question he asks himself as he edits is, "Is this story coherent and, if it is, what will turn coherence into a song? What are the recurring elements? Do they entwine to make a theme?... What I want, most of all, is resonance."<sup>3</sup> While King's use of the word "song" is largely metaphorical, the questions that follow and his goal of resonance suggest that the sonic properties of words, even when only imagined in the reader's mind, are important to the clarity and artistry of the story.

Again, when I say that this dissertation is about how words sound, I should clarify once more that it is not—at least not focally—about how words sound in our mind when we read; rather, it is about how words sound when spoken aloud. More specific still, at its core this dissertation is about how words sound when spoken aloud *for the purposes of artistic creation and expression*. For me, the purposeful act of creation and expression is what differentiates Spoken Word from words spoken. Subtle though it may seem, the requisite of artistic and expressive intent does narrow my focus to some degree, such that the reader will not find analyses of the Word Concerto for Friend at the Bar in this dissertation (though I will confess I find the challenge enticing). Nor will the reader find

---

<sup>3</sup> Stephen King, *On Writing: A Memoir of the Craft* (New York: Scribner, 2000), 173.



analyses of Stephen King's novels or Bo Burnham's poetry. Still, the general definition of Spoken Word allows a certain freedom to explore artistic forms and genres not typically thought of as music or as in the purview of music theory.

### 1.1.1 Voiced Texts and Oral Performance

Spoken Word defined thus far is something of a catch-all term for any word-based art form performed orally by an artist and received aurally by an audience. Oral poetries reach an audience through a diverse array of media and by recognizing that diversity we can distinguish between various kinds of oral poetries. I will further refine my scope by distinguishing between two kinds of Spoken Word using terms borrowed from comparative oral traditions scholar John Miles Foley: *voiced texts* and *oral performances*, two different types of oral poetries.

In his 2002 book, *How to Read an Oral Poem*, Foley proposes a general system of media categories that takes into consideration how a work is composed (in real time or premeditated); how it is performed (on the page or on the stage); and how it is received (heard or read). The model, Foley hopes, is "just comprehensive enough to demonstrate some overall unities and offer some practical bases for comparison, and just flexible enough to accommodate the natural diversity of human expression (Figure 1.1).<sup>4</sup> A voiced text in Foley's model is one in which the work is composed in a written medium, performed orally, and received aurally.

---

<sup>4</sup> John Miles Foley, *How to Read an Oral Poem* (Urbana: University of Illinois Press, 2002), 38–40.

**Figure 1.1 Table of media categories for oral poetry by John Miles Foley.**

	<b>Composition</b>	<b>Performance</b>	<b>Reception</b>
<b>Oral Performance</b>	Oral	Oral	Aural
<b>Voiced Texts</b>	Written	Oral	Aural
<b>Voices from the Past</b>	Oral/Written	Oral/Written	Aural/Written
<b>Written Oral Poems</b>	Written	Written	Written

While Foley offers slam poetry as an exemplar of voiced texts, one could argue that the category also accurately describes most Western classical music. Just as a voiced text is composed in a written medium, performed orally, and received aurally, the composition of a musical work begets a score that is sounded in performance and appreciated by a listening audience. Indeed, twentieth- and twenty-first century artists across many disciplines seem particularly interested in exploring the relations between spoken and musical sound systems, and an understanding of their work as related within the broader media category of voiced texts offers a new entry point for analytical investigation.

### 1.1.2 Spoken Word and/as Musical Material

While voiced texts and Western classical music may occupy the same media category, we tend to think of their compositional material in very different terms. And despite the strong temptation to argue for their synonymy by invoking John Cage's precept that Spoken Word and Western classical music both simply comprise sounds organized over time, I can hardly deny that the sounds organized over time in Spoken Word engage linguistic systems and cognitive processes that are experientially (and, many would

argue, evolutionarily) distinct from musical systems and cognitive processes.<sup>5</sup> How might one reconcile these differences? I propose one must begin—as so many good things do—with the Word.

When comparative oral traditions, such as that of Foley, address the concept of a “word” as it pertains to Spoken Word, they invoke a more abstract concept than simply the things one finds in a dictionary. Foley, among others interested in oral and performance poetry traditions, tend to prefer to think of a word as a speech act. Inspired by Serbo-Croatian oral epic bards (*guslari*), Foley suggests that when examining verbal art we do better to adopt the concept of *reč* to talk about words.<sup>6</sup> A word is not, Foley explains, “a string of black letters bounded by white spaces or something enshrined in a dictionary, but rather a unified utterance, complete enough to have idiomatic force as a speech act.”<sup>7</sup> Words, in this sense, can share the same bounds as lexical words but they also can (and just as often do) extend beyond those bounds.

Words also need not have any lexical correlate at all: laughter, groans, hisses, and nonsensical syllabic formations can all constitute a word, should their immanent context merit. Spoken Word comprises words in this sense, which renders comparison to music somewhat easier than if one were to try to map a one-to-one relationship between

---

<sup>5</sup> The reader may share in my admiration for John Cage and his ability to solve so complex a problem as the relationship between speech and music with such elegance.

<sup>6</sup> The Serbo-Croatian word *reč* technically translates to “word” in English, but its definition is more nuanced. A more accurate translation might be “sound-byte,” or a distinct and integral unit of expression within one’s particular way of speaking.

<sup>7</sup> Foley, *How to Read an Oral Poem*, 17. Also John Miles. Foley, “Word-Power, Performance, and Tradition,” *The Journal of American Folklore* 105 (1992): 279.

phonemes and notes or lexemes and chords. By thinking of the word concept qua *reč*, words are much more like musical motives—salient recurring figures or succession of sounds that have special importance in or are characteristic of a composition.

## 1.2 Genres, Styles, and Types

Given the import and popularity of rap and hip-hop in our current culture and society, as well as the historical relationship between rap and spoken-word poetry, it would seem natural for a study of the musical qualities of speech in the context of spoken-word art forms to focus on these genres. But rap, while among those art forms that constitute voiced texts, is far from the only artistic tradition to explore and exploit the bounds of language and music and the present work focuses instead on spoken-word genres that are not only distinct from rap or hip-hop, but often go overlooked by analytical theorists in both musical and literary disciplines.<sup>8</sup>

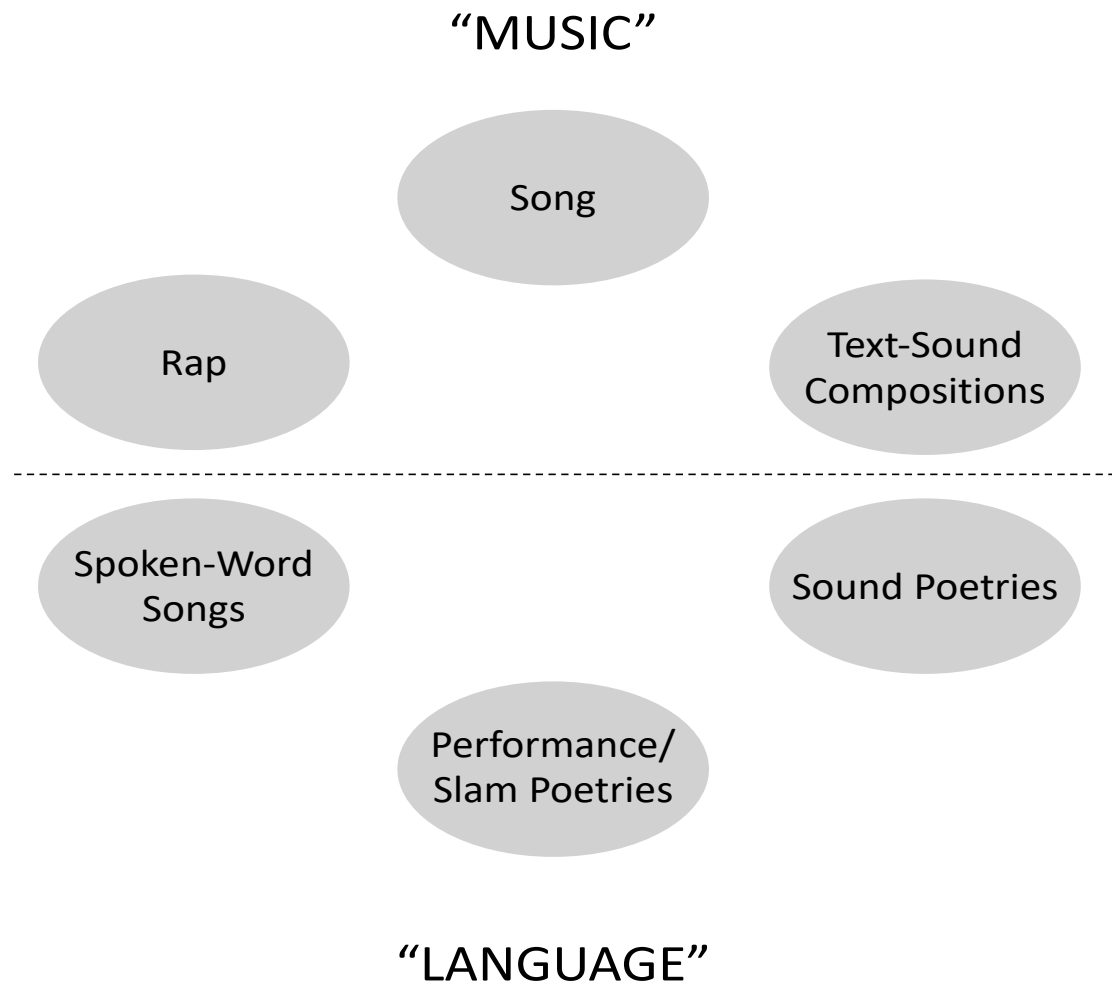
The categorical concept of voiced texts establishes a familial relationship among musical and poetic genres that would otherwise appear to have little in common. Some genres in this voiced text family are already considered musical, while others are not. A large part of the project at hand endeavors to address this. I argue that a driving factor for treating

---

<sup>8</sup> I am not the first to recognize the necessity of this distinction: Price-Styles similarly discusses the interplay between spoken-word poetry and hip-hop, focusing on the impact and influence that politically charged spoken-word poetry has had on hip-hop and how the progressive nature of rap as an oral form relates to spoken-word poetry. While Price-Styles recognizes the similarities and overlaps between spoken-word and rap, she ultimately argues that the two forms are distinct, writing that “in comparing the two forms of rap and spoken word, the merits and achievements of both are illuminated and the two can be appreciated as important and equal entities” (20). Alice Price-Styles, “MC Origins: Rap and Spoken Word Poetry,” in *The Cambridge Companion to Hip-Hop*, ed. Justin A. Williams (Cambridge: Cambridge University Press, 2015), 11–21.

some voiced texts as more musical than others is a product of accepting, a priori, some dichotomy between language and music. Within the voiced text family, six genres stand in relation to one and other. A coherent graphic representation of this family, oriented along a North-South axis where North is the land of music and South is the land of language, shows how they relate to each other and how a dichotomous view of language and music results in the privileging of certain genres over others in musicological, music-analytical, and music-theoretical discourse (Figure 1.2).

**Figure 1.2 Relational diagram of voiced text genres.**



The genre of song (art or pop) sits squarely at the top, as a type of voiced text that is undeniably music. Opposite song and at the South pole of the hexagon sits performance poetry (which encompasses the genre of slam poetry), closest to the realm of language.<sup>9</sup> Four intermedial genres sit somewhere between. Rap and Spoken-Word Song group together because they both involve speech utterances in conjunction with non-vocal musical instruments. In these genres, Spoken Word becomes an instrumental line in itself, and the artistic virtuosity of rappers and spoken-word songsters emerges primarily in their manipulation of phonological features of speech – emphasizing the musical potential inherent in the timing and tuning of words.

While their affinities would group rap and spoken-word song together, they do not sit at the same point in relation to the North-South dichotomy of music and language. A rapper's flow is heavily influenced by the poetic structure and literary techniques employed within the lyrics of a rap. Despite a wide range of styles within the broad genre of rap, there is a commonality of rhyme and poetic meter—another way of saying a regularly recurring syllable count and stress pattern. Rap ubiquitously exploits numerous literary devices like rhyme, alliteration, and many rely on consistent metrical patterns. The degree to which an artist implements these various devices, and the degree to which the performance of a text aligns with and reinforces an underlying musical setting, will influence the perceptual qualities of a rapper's flow.

---

<sup>9</sup> The term performance poetry originates from an early press release describing the 1980s performance poet Hedwig Gorski, whose audio recordings achieved success on Spoken Word radio programs around the world. The term attempts to distinguish her text-based vocal performances from performance art, especially the work of performance artists such as Laurie Anderson, who worked with music at the time.

One of the most notable distinctions between rap and spoken-word song relates to the kinds of flow these artists implement in their word. Rap scholars tend to agree that the distinctive styles of certain rappers are reflections of their different kinds of *flow*, or particular way of particular way of composing and orally delivering a text. The concept of flow, popularized in Adam Krims's 2000 book, *Rap Music and the Poetics of Identity*, refers to the rhythmic structure that arises from the interaction between the rapper's words and the musical rhythms of those words, such that both what is said and how it is said contribute to the kinds of flow that emerge in a particular rap performance. Krims offers three kinds of flow that rappers may use: Sung rhythmic style (used by Too Short, Grandmaster Flash and the Furious Five, and the Beastie Boys); Percussive-effusive style (used by B-Real of Cypress Hill); and Speech-effusive style (used by Big Pun).<sup>10</sup>

While there is room within Krims's speech-effusive style of flow to include more flexible rhythmic structures in rap, few examples come close to the kind of flow one hears in a spoken-word song, where the relationship between the words and the musical rhythms of those words is often considerably more ambiguous. Spoken-word songs have their own kind of flow that one might add to Krims's list of flow styles: a spoken-word style.<sup>11</sup> In other words, spoken-word song is not just effusive of speech, but rather it is very naturally spoken.<sup>12</sup> All this is to say that rap sits perhaps further South than song on the

---

<sup>10</sup> Adam Krims, *Rap Music and the Poetics of Identity* (Cambridge: Cambridge University Press, 2000), 50–51.

<sup>11</sup> And this is a style that is, on occasion, even used by rappers, primarily on introductory tracks on albums, or as segues between songs in live performance.

<sup>12</sup> It may be the case that part of why this flow is so uniquely different to the flow styles of rap is because spoken-word song does not share the same foundational goals of rap, at least those suggested by the origin story that relates rap to MCing the dance party.

North-South dichotomy of music and language, and spoken-word song sits further North than performance poetry, but neither move so far as to intersect or overlap.

Text-Sound compositions and Sound poetry similarly group together and sit in comparable relation to the North-South dichotomy of music and language. By text-sound compositions, I mean not only those works by the Swedish artists Lars-Gunnar Bodin, Bengt Emil Johnson and Åke Hodell, for and by whom the term was originally coined, but works that more generally engage with the linguistic dichotomies of written versus spoken word and text versus the sound of its utterance in relation to both poetry and music.<sup>13</sup>

Emerging from electronic experimental music, text-sound compositions include works by many well-known composers including (but by no means not limited to) Karlheinz Stockhausen, Luciano Berio, György Ligeti, Milton Babbitt, John Cage, Steve Reich, Gavin Bryers, Charles Amirkhanian, Charles Dodge, Georges Aperghis, and Paul Lansky.<sup>14</sup> While text-sound compositions push the boundaries of the language-music

---

Spoken-word songs, on the other hand, emerge out of the Beat generation's jazz-poetry fusion movement, and the philosophical vision to enhance the experience of oral poetic traditions driving the movement.

<sup>13</sup> Larry Wendt, "Sound Poetry: I. History of Electro-Acoustic Approaches II. Connections to Advanced Electronic Technologies," *Leonardo* 18, no. 1 (1985): 11–23. More about the compositions and biographical details of Lars-Gunnar Bodin, Bengt Emil Johnson and Åke Hodell can be found in Jesper Olsson, "Kneaded Language: Concrete Poetry and New Media in the Swedish 1960s," *Modernism - Modernity*, 2011, <https://doi.org/10.1353/mod.2011.0038>. Accessed November 30, 2020. See also Per Bäckström, "The Intermedial Cluster: Åke Hodell's Lågsniff," *Acta Universitatis Sapientiae, Film and Media Studies* 10, no. 1 (2016), <https://doi.org/10.1515/ausfm-2015-0025>. Accessed November 30, 2020.

<sup>14</sup> A list of works by these and other composers that would fall into the text-sound composition type of voiced texts can be found in Appendix C.



dichotomy, the list of practitioners above indicates a primary involvement of composers associated with musical artistic practices reaching towards language as a compositional medium. Like rap, text-sound compositions sit perhaps further South than song on the North-South dichotomy of music and language. The inverse genre is often called Sound Poetry, which sits further North than performance poetry, but not so far North as to quite intersect or overlap with text-sound compositions.

Sound Poetry, also called sonorist rhythms, phonetic poetry, or *poésie sonore*, engages with the relationship between sound and sense in poetry such that the phonetic and aural properties of language become of greater import than the semantic or syntactic logic that would typically drive the structure of a work. Sound poems challenge the limits of natural language, with some poems attempting to generate natural signifying relationships between sound and meaning through phonetic symbolism while others use sound as antagonistic or indifferent toward meaning.

The practice of sound poetry has fairly distinct origins primarily in the Dadaist and Futurist movements emerging around the turn of the twentieth century and maturing through the inter-World War years.<sup>15</sup> On the Futurist front, an early example of sound poetry is Filippo Achille Emilio Marinetti's sound poem *Zang Tumb Tuuum* (1914), which uses *Parole in libertà* (words in freedom) and other poetic impressions of the events of the battle, including the sounds of gunfire and explosions.

---

<sup>15</sup> Henry Feinsod, "Sound Poetry," in *The Princeton Encyclopedia of Poetry and Poetics*, ed. Roland Greene et al. (Princeton, N.J.: Princeton University Press, 2012), 1327–29.

On the Dadaist front, Hugo Ball (1886-1927) was an essential pioneer in the development both of the Dada movement itself and of sound poetry.<sup>16</sup> Hugo Ball's major contribution to Dada evenings, and perhaps the most innovative experiment in Zurich Dada was the purely abstract sound poem, of which Ball's *Karawane* (Caravan) is best known (Figure 1.3). In *Karawane*, and in other works of its kind, Ball attempted to emancipate words from functional and syntactic control by inventing new word-sounds that resonated and half alluded to meaningful linguistic concepts.

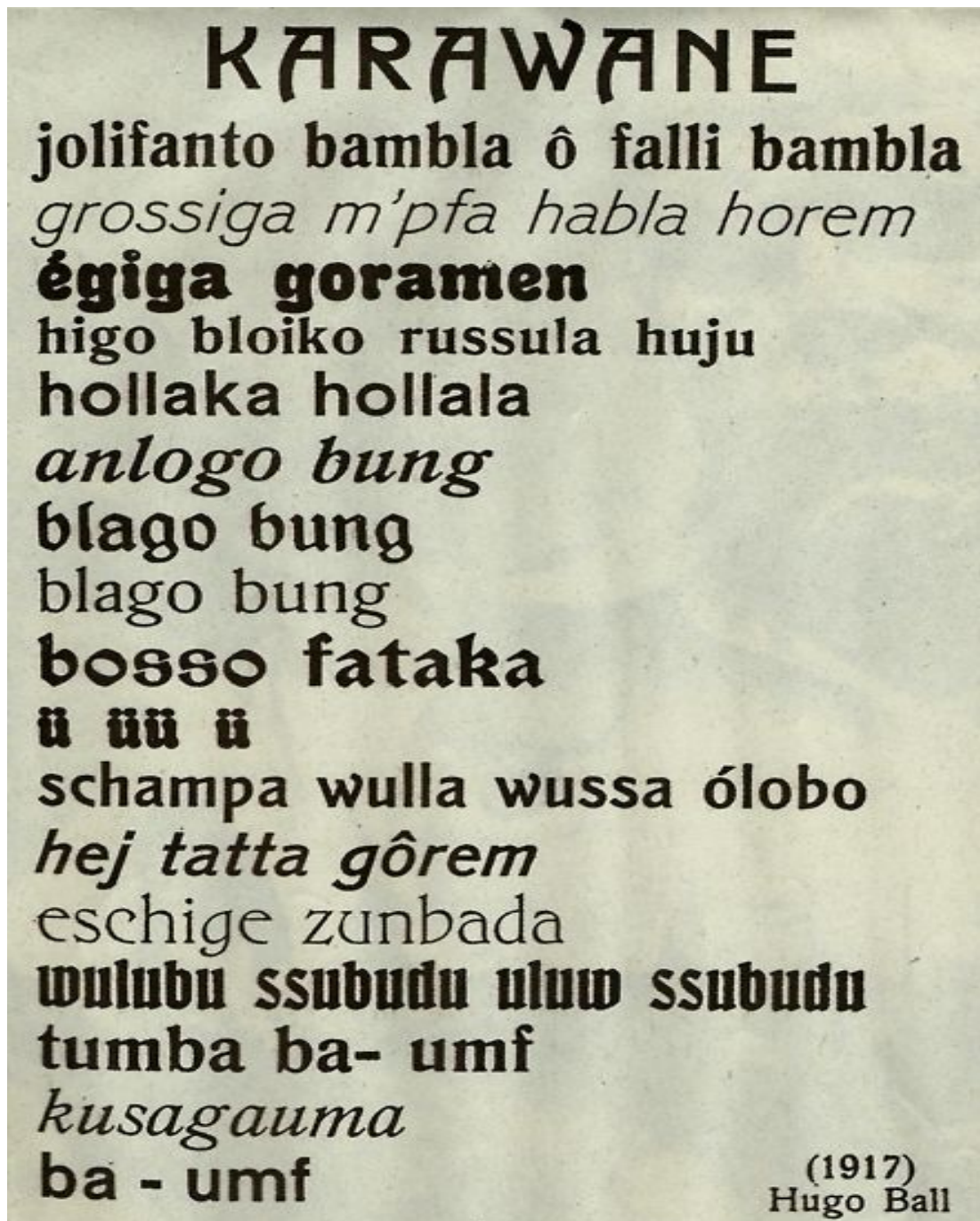
Other Dadaist or Dada-affiliated artists whose sound poetry paved the way for subsequent generations to explore phonetic collage, experimental visual and concrete poetry, and performance poetry include Tristan Tzara (1896-1963), Romanian writer, promoter and poet; Richard Huelsenbeck (1892-1974), German writer, editor and publisher; Elsa von Freytag-Loringhoven (1874-1927), German avant-garde, Dadaist artist, and poet; Raoul Hausmann (1886-1971), Austrian artist and writer; and Kurt Schwitters (1887-1948), German artist, poet, and publisher.<sup>17</sup>

---

<sup>16</sup> Marc Dachy, *The Dada Movement, 1915-1923* (Geneva; New York: Skira; Rizzoli, 1990).

<sup>17</sup> For the interested reader, performances of sound poems by these artists can be found at <https://writing.upenn.edu/pennsound/x/Dada-Sounds.php>.

Figure 1.3 Hugo Ball's *Karawane* (1917).<sup>18</sup>



<sup>18</sup> Hugo Ball, *Karawane* (1917). In Richard Huelsenbeck, Malcom Green, and Alastair Brotchie, eds., *The Dada Almanac*, 2nd ed. (London: Atlas Press, 1993), 53.

### 1.3 Speakers, Slams, and Sounds

Of the six genres or types of voiced texts discussed (song, rap, sound-text compositions, spoken-word songs, performance/slam poetry, and sound poetry), those three occupying the Northern hemisphere of Figure 1.2 have, to varying extents, found some reasonable degree of footing in contemporary musicological and music theoretical/analytical discourses. But for those sitting closer to the South, few have acknowledged the familial relationship to musical practice.<sup>19</sup> Fewer still treat the works of these genres as sonic compositions in the same vein as music proper and explore what insights might materialize by adopting a music-analytical lens. Literary scholar Corey Frost touches upon the challenges brought about by the liminality of Spoken Word when he calls it “a form that overlaps with many categories but doesn’t quite belong to any.”<sup>20</sup> In my dissertation, I take on this challenge by playing the game “what if I were to call this music?” Focusing on representative works from each of the southern genres presented in Figure 1.2, the subsequent analytical chapters address spoken-word songs, slam poetry, and sound poetry in turn.

---

<sup>19</sup> Sound poetry is, perhaps the exception, as scholars like Marjorie and Nancy Perloff, Craig Dworkin, Christian Bök, Larry Wendt, and other have made strides to establish a space within musical discourses for sound and abstract poetic traditions. Marjorie Perloff and Craig Dworkin, eds., *The Sound of Poetry/The Poetry of Sound* (Chicago: University of Chicago Press, 2009). Within the spoken-word song genre, a recent and notable exception is Marian Wilson Kimber, *The Elocutionists: Women, Music, and the Spoken Word* (Urbana: University of Illinois Press, 2017).

<sup>20</sup> Corey Frost, “The Omnidirectional Microphone: Performance Literature as Social Project” (Ph.D. diss., City University of New York, 2010), 5.

Chapter 2, “The Singer Speaks,” examines the rhythmic and melodic interplay among speech and music that make American singer-songwriter Tom Waits’s spoken-word songs so uniquely musical. With more than fifty spoken-word songs recorded over five decades, Waits’s contributions to the unique genre provide a rich repertoire worthy of closer study. In various interviews about his spoken-word songs, Waits acknowledges the influence of Beat generation poets such as Allen Ginsberg, Jack Kerouac, William S. Burroughs, Ken Nordine, Charles Bukowski, to name only a few, and frequently cites one album in particular as a model for much of his own spoken-word works. But critics and fans alike have recognized a crucial difference between Beat generation spoken-word song and Waits’s unique contributions to the development of the form. Robert Ward writes that “Waits has an advantage over his Beat predecessors, because he is first a songwriter. He makes the words and music fit... Waits is not merely cashing in on nostalgia—he is improving on the form.”<sup>21</sup> In close examinations of selected spoken-word songs, I explore how Waits fits speech and music together, paying particular attention to the rhythmic and melodic elements of Waits’s speech in relation to the metric and harmonic material with which his speech is combined. Using recording analysis software Sonic Visualiser, complemented by a range of algorithmic plugins designed to identify changes in spectral energy, I extract precise measurements of Waits’s speech timing and intonation and propose listening strategies to hear Waits’s speech in musically satisfying ways. Analyses focus on Waits’s 2002 recording of “Watch Her Disappear” from the studio album *Alice*, as well as his 1974 studio recording of “The Ghosts of

---

<sup>21</sup> Robert Ward and Tom Waits, “Play It Again, Tom,” in *Tom Waits on Tom Waits: Interviews and Encounters*, ed. Paul Jr. Maher (Chicago: Chicago Review Press, 2011), 55.

Saturday Night (After Hours at Napoleone's Pizza House)" from the album *The Heart of Saturday Night* versus a live performance from the same year. I argue throughout that while Waits is undoubtedly speaking these songs, he does so both in time and in tune. These songs are not simply stories told over an independent and unrelated musical accompaniment; rather they are integral parts of a musical whole.

Chapter 3, "The Poet Performs," transitions from spoken-word songs into slam poetry and considers what it means to think of Spoken Word from a musical perspective in the absence of any unequivocally musical sounds. Whereas the previous chapter focused on precise timing and frequency measurements, this chapter considers the formal similarities between slam poetry and music. The vast majority of slam poems will invariably fall into the nebulous poetic category called "free verse," a kind of poetic form that, for better or worse, exonerates the analyst from considering formal structure in any serious way. In this chapter, I explore how theories of form developed for musical study can apply to and shed light on slam poetry performances. By adopting the musical theory of formal functions, poems lacking consistent meter or rhyme can still be thought of as having distinguishable forms that are analogous to familiar musical structures. Although the specific features of tonal music do not properly occur in speech (the medium for slam poetry), it is possible to draw parallels between functions in music and the parts of a poetic verse, if such functions are sufficiently generalized and redefined. The first step towards generalizing formal functions for application outside the specialized language of the high Viennese school is to reformulate the notion of compositional conventions as tradition-bound formulae in order to distinguish the principles of form-functionality from the specific practices of a specialized language. Through a generalized form-functional

analysis of slam founder Marc Kelly Smith's 2011 performance of the work, "Kiss It," and his 2011 performance of "I wear my father's coat," I demonstrate how re-envisioning the emergent design of a slam performance as a musical form can ultimately communicate more about the formal structure than a textual representation of a slam poem might suggest.

Chapter 4, "The Artist Composes," extends the analytical task one step further, considering Spoken Word from a musical perspective in the absence not only of unequivocal musical sounds but now, also, in the absence of semantic meaning. I offer a new approach to understanding the somewhat enigmatic *Ursonate* by collage artist Kurt Schwitters's, treating the work as a kind of musical collage. Although literary critics often cite Schwitters's 1919 poem, "An Anna Blume," as his most important literary work, music scholars quite naturally gravitate towards his epic sound poem, *Die Ursonate*, (sometimes referred to as *lautsonaten* or *Die Sonate im Urlauten*), composed over the decade from 1922–1932. Scholars from diverse disciplines have engaged with *Die Ursonate*, and the diversity of research interest is matched by the diversity of methodologies employed, which range from historical accounts of the decade-long compositional process to attempts at translations of its German nonsense syllable elements, reinterpretations of its design in new artistic mediums, evaluations of artistic print editions made after Schwitters's death, critiques of the work's artistic merit, at least nine recordings in full or in part, various versions for ensemble and even for robot, and at least two music-analytical studies.<sup>22</sup> Extant musical analyses of *Die Ursonate* are

---

<sup>22</sup> Nancy Perloff, 'Schwitters Redesigned: A Post-War *Ursonate* from the Getty Archives', *Journal of Design History* 23, no. 2 (2010): 195–203. Dary John Mizelle,

primarily concerned with thematic development and form, likely because these features represent the main and most obvious connections between the sound poem and music (other than the performance directive). But when we take into consideration both the efforts Schwitters made to render *Die Ursonate* musically (in notation and in performance) and the particularities of his artistic philosophy, *Merz*, such considerations reveal new and surprising ways in which *Die Ursonate* sews together poetic and musical sound, particularly in the formally straightforward third movement, *Scherzo*.<sup>23</sup> My analysis of the third movement is informed by the style and technique of Schwitters as an artist during the mid-1920s, including his collage techniques and assemblage works (in which he transformed appropriated imagery and texts from print media into dynamic and layered compositions), his short-lived jigsaw techniques, and his attitude towards the use and metamorphosis of found-objects in his artistic works. With these techniques and the *Merz* philosophy in mind, I present an analysis that treats the *Scherzo* as a found-object

---

‘Kurt Schwitters’s *Ursonate* and *Quanta* and *Hymn to Matter*’ (PhD dissertation, University of California, 1977). Sabine Lichtenstein, “Listening to Kurt Schwitters’ *Ursonate*” 38 (2003): 276–85. Jack Ox, “Creating a Visual Translation of Kurt Schwitters’s *Ursonate*,” *Leonardo Music Journal* 3, no. 1993 (1993): 59–61. The most comprehensive list of recordings of *Die Ursonate* can be found through the [PennSound Centre for Programs in Contemporary Writing](#) at the University of Pennsylvania.

<sup>23</sup> While Schwitters is commonly associated with the Dada movement in Germany, he did not consider himself nor his work Dadaist in nature and thus adopted a different descriptor. The word *Merz*—a syllabic extraction from a found print caption “Kommerz – und Privatbank”—applies not only to his paintings, pictures, poetry, and constructions, but also to his person and philosophical outlook on abstract art. The particularities of *Merz* as an artistic philosophy are discussed in detail in Chapter 4.



collage that uses similar jigsaw/assemblage techniques, pasting together incipits from pre-existing musical pieces to create, as Schwitters himself puts it, “a genuine joke.”<sup>24</sup>

Finally, Chapter 5, “Last Words,” addresses the implications of a speech-song (or language-music) dichotomy for studies of cognition and perception, exploring areas of consideration for empirical studies that endeavor to examine the relationship between cognitive processing of music and language. This chapter also acknowledges the woefully narrow sliver of repertoire and artists represented in the central analytical chapters and provides a selected list with brief descriptions of other spoken-word songs, slam poems, and sound poems that the reader may wish to explore and through which the reader may find themselves inspired to play the game “what if I called this music?”

Each of the three kinds of Spoken Word that feature in this dissertation—spoken-word song, slam poetry, and sound poetry—exist not simply on the fringes of a single field but at the intersections of multiple artistic disciplines. This work endeavors not only to offer an array of analytical approaches and listening strategies that can provide a basis for future investigations of speech-based contemporary art music; it also strives to recognize the categorical overlap that exists among Spoken Word and music, and to bring Spoken Word further into the view of mainstream musicological research and analysis.

The subsequent chapter on spoken-word song is the most natural starting point from which to begin a journey into spoken-word art forms that do not presently find themselves center-stage in musical scholarship, largely because this type of Spoken Word

---

<sup>24</sup> *Merz 24*, accompanying notes to the 1932 score. In Mizelle, “Kurt Schwitters’s Ursonate and Quanta and Hymn to Matter.”

retains recognizable musical practices. Spoken-word songs combine naturally produced speech with instrumental music, thus providing a familiar setting from which to investigate how the rhythms and intonation of naturally produced speech can work in musical ways and establish natural sounding speech as a musical instrument.

## Chapter 2

### 2 The Singer Speaks

Though not exclusive to popular music, spoken-word songs—the art of combining natural speech with instrumental music—are surprisingly ubiquitous in popular music genres, with far proponents exploring the form than one might initially suspect. William Shatner, David Bowie, the Shangri-Las, James Brown, Patti Smith, the Velvet Underground, Morrissey, Madonna, to name only a small few, have tried their hand at probing the borders between speech and music through spoken-word song.<sup>25</sup> While far from being the only figure in popular music to create spoken-word songs, American singer-songwriter Tom Waits is undoubtedly one of the most prolific. With more than fifty spoken-word songs recorded over five decades, his contributions to this unique genre provide a rich repertoire worthy of closer study.<sup>26</sup>

---

<sup>25</sup> Many more popular music artists have all written or performed at least one spoken-word song, including Belle and Sebastian, Black Flag, Talking Heads, Blur, Iggy Pop, John Cale, Tindersticks, Leonard Cohen, Jim Morrison, Pearl Jam, Sonic Youth, etc. A list of spoken-word songs by these and other artists is provided in Appendix A.

<sup>26</sup> While analytical scholarship on Tom Waits's music is, at present, minimal, a few notable studies that focus on exploring the words and music of this rich repertoire include Margaret E. Thomas, "Text and Temporality: Toward an Understanding of Rhythmic Irregularities in the Music of Tom Waits," *Music Theory Online* 22, no. 4 (2016); Corinne Kessel, *The Words and Music of Tom Waits* (Westport, Conn.: Praeger, 2009); Gabriel Solis, "'Workin' Hard, Hardly Workin': Hey Man, You Know Me - Tom Waits, Sound, and the Theatrics of Masculinity," *Journal of Popular Music Studies* 19, no. 1 (1974): 26–59. See also James Braxton Peterson, "The Depth of the Hole: Intertextuality and Tom Waits's 'Way Down in the Hole,'" *Criticism* 52, no. 3–4 (2010): 461–85; Stephen Patrick Tchir, "Tom Waits's Rain Dogs: Influences and Musical Genre" (M.A. thesis, University of Alberta, 2013); Christos Iraklis Tsatsoulis and Markus Hofmann, "Focusing on Maximum Entropy Classification of Lyrics by Tom Waits," in *IEEE International Advance Computing Conference (IACC)*, 2014, 664–67.

This chapter examines two spoken-word songs by Tom Waits, whose recitation style has garnered critical praise for its unique musicality. Through recording analyses of “Watch Her Disappear” from the 2002 studio album *Alice* and “The Ghosts of Saturday Night (After Hours at Napoleone’s Pizza House)” from the 1974 album *The Heart of Saturday Night*, I investigate how the rhythms and intonation of Wait’s recitations work in musical ways to establish natural sounding speech as a musical instrument.

## 2.1 Tom Waits’s Spoken-Word Songs

In his seemingly improvised (so as to sound spontaneous) spoken-word songs, American singer-songwriter Tom Waits plays words like a fiddle. “I’m a word guy,” Waits says. “Words are music, really. I mean, people ask me, ‘do you write music, or do you write words?’ But it’s all one this at its best.”<sup>27</sup>

In various interviews about his spoken-word songs, Waits acknowledges the influence of Beat generation poets such as Allen Ginsberg, Jack Kerouac, William S. Burroughs, Ken Nordine, Charles Bukowski, to name only a few, and frequently cites one album in particular as a model for much of his own spoken-word works: “there’s a fascinating album that came out in ’57 on Hanover records: *Kerouac/Allen*. It’s Jack Kerouac telling stories, with Steve Allen playing piano behind him. That album sort of sums up the whole

---

<sup>27</sup> Amanda Petrusich and Tom Waits, “Interview: Tom Waits” (*Pitchfork* November 2006), in *Tom Waits on Tom Waits: Interviews and Encounters*, ed. Paul Jr. Maher (Chicago: Chicago Review Press, 2011), 419.

thing. That's what gave me the idea to do some spoken pieces myself."<sup>28</sup> Three albums from the mid-1970s—*Nighthawks at the Diner*, *Blue Valentine*, and *Small Change*—feature several spoken-word songs that are unmistakably modeled after the Beat generation. These early examples often involve musical settings that are metrical, tonal, and in a jazz/blues style, such as “Emotional Weather Report,” “Frank’s Wild Years,” and “Romeo is Bleeding.”

After 1980, Waits’s use of spoken-word song becomes more strategic, limited to one or two per album and incorporating an eclectic mix of musical styles and genres, including waltzes (“Watch Her Disappear,” “Lucky Day Overture”), rhumbas (“Crossroads,” “Telephone Call from Istanbul”), tangos or habaneras (“Red Shoes by the Drugstore”), polkas (“Kommienzuspadt”), hymns (“That’s the Way”), vintage rock ‘n’ roll (“Get Lost”), and modern industrial (“Hell Broke Luce”). More experimental songs, such as “What’s He Building in There,” “Children’s Story,” “The Ocean Doesn’t Want Me,” “Army Ants,” and “Circus” create something akin to soundscapes that support text recitation.

In his spoken-word songs, Waits’s texts and recitations are as diverse as his musical style. Some of his spoken-word songs use rhyming texts (“The Ocean Doesn’t Want Me,” “Pasties and a G-String,” and “Hang on St. Christopher”) while others are more clearly prose (“Watch Her Disappear,” “Frank’s Wild Years,” “Big Joe Phantom 309,” “Shore Leave,” “Crossroads”). Some of the songs in which Waits combines speech with music

---

<sup>28</sup> Todd Everett and Tom Waits, “Tom Waits: In Close Touch With the Streets” (*Los Angeles Free Press* October 1975), in *Tom Waits on Tom Waits: Interviews and Encounters*, ed. Paul Jr. Maher (Chicago: Chicago Review Press, 2011), 46.

might best be thought of as *speech song*, with controlled rhythms and the occasional passage of identifiable pitch sequences; others are best described as *rhythmic speech*, comparable in many ways to rap; while others still are truly spoken-word, employing natural and connected speech. In every case, Waits builds an intricate text-music relationship that distinguishes his spoken-word songs from his Beat predecessors. Such distinction is not lost on critics and fans of Waits's music, nor is it lost on Tom Waits himself who, when speaking about the influence of the Beats on his own spoken-word songs says, "I'm considered a songwriter so [doing spoken-word songs] is something I have to watch... I don't call them poetry, 'cause there are too many poets I admire; but they are in an oral tradition. I call them 'metropolitan double-talk'."<sup>29</sup> Critics too, have written about the difference between beat generation exploration of spoken-word song and Waits's unique contributions to the development of the form. Robert Ward, for example, writes that "Waits has an advantage over his Beat predecessors, because he is first a songwriter. He makes the words and music fit... Waits is not merely cashing in on nostalgia—he is improving on the form."<sup>30</sup>

So how, exactly, does Waits "make the words and music fit"? If one studies these songs closely, they reflect a surprisingly intricate degree of interaction between speech and music. In this chapter, that is the question. In close examinations of selected spoken-word

---

<sup>29</sup> Jeff Burger and Tom Waits, "Waits: The Beats Go On..." (*Melody Maker* June 1975), in *Tom Waits on Tom Waits: Interviews and Encounters*, ed. Paul Jr. Maher (Chicago: Chicago Review Press, 2011), 41.

<sup>30</sup> Robert Ward and Tom Waits, "Play It Again, Tom" (*Syracuse Times* July 1976), in *Tom Waits on Tom Waits: Interviews and Encounters*, ed. Paul Jr. Maher (Chicago: Chicago Review Press, 2011), 55.

songs, I explore just how Waits fits speech and music together, paying particular attention to the rhythmic and melodic elements of Waits's speech in relation to the groove structure and harmonic material of the music with which his speech is combined. Ultimately, the objective of this chapter is to demonstrate how musical Waits's particular way of speaking is in his spoken-word songs. These songs are not simply stories told over an independent and unrelated musical accompaniment; rather, they are integral parts of a musical whole, equally musical and equally forms of music making. When asked, in the late summer of 1985, which, of the countless instruments that" Waits plays, he considers to be his primary instrument, he answers that "vocabulary is [his] main instrument."<sup>31</sup> Taking him at his word, this chapter explores how Waits plays his words in the most musically satisfying ways.

## 2.2 Recording Analysis

Because Waits's spoken-word songs exist, first and foremost, as recordings, the following analyses came about through close listening, transcription, and the use of a variety of recording analysis software. The latter has become possible in the last decade through the work of several empirically-minded musicologists who developed sophisticated methods for analysis of recorded music—tools that can assist the analyst to identify, isolate, and measure sounding events as received through recordings of those events. and we also have sophisticated theoretical models for interpreting the various

---

<sup>31</sup> Pete Silvertown and Tom Waits, "The Sultan of Sleaze" (*You Magazine*, August 1985), in *Tom Waits on Tom Waits: Interviews and Encounters*, ed. Paul Jr. Maher (Chicago: Chicago Review Press, 2011), 154.

sonic parameters that can be measured.<sup>32</sup> In an effort to prioritize the performative nature of spoken word, I adopt analytical tools and methodologies developed in empirical musicology and music theory for the study of music performance. Specifically, I examine recorded spoken-word songs using recording analysis software *Sonic Visualiser* to generate a spectrogram and/or waveform of the sounds produced in a spoken-word performance.<sup>33</sup> A spectrogram is a visual representation of sound. It displays the amplitude of the frequency components of a sound source over time. According to musicologist Nicholas Cook, spectrograms “are most useful for homing in on the details of performance—the un-notated nuances that are responsible for so much of music’s meaning... When they are integrated into the working environment for studying recordings, as in *Sonic Visualiser*, they help to transform listening into analytical interpretation.”<sup>34</sup>

In principle, a spectrogram illustrates all aspects of a sound source. As such spectrograms can be used to extract and measure various kinds of information: variations in intensity, which can be used to indicate relative loudness; fundamental frequencies ( $f_0$ ), which can

---

<sup>32</sup> Nicholas Cook, “Methods for Analysing Recordings,” in *The Cambridge Companion to Recorded Music*, ed. Nicholas Cook et al. (Cambridge: Cambridge University Press, 2009), 221–45.

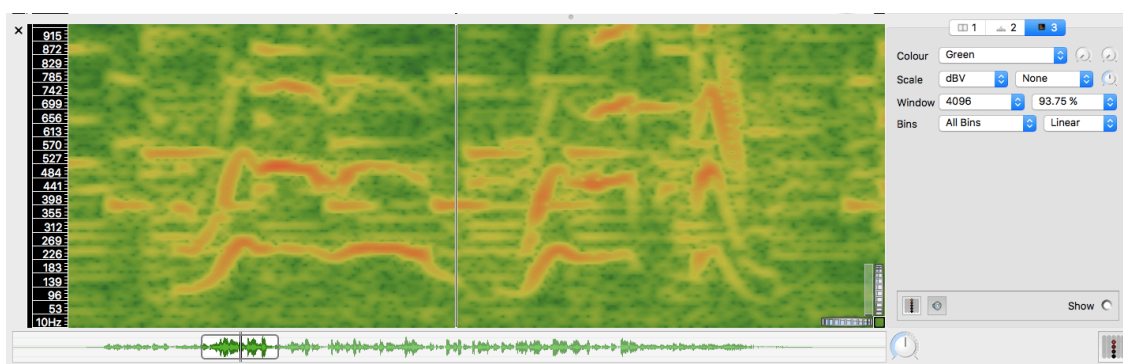
<sup>33</sup> *Sonic Visualiser* is a freeware distributed under the GNU General Public License (v2 or later) and developed at the Centre for Digital Music at Queen Mary, University of London. See Chris Cannam, Christian Landone, and Mark Sandler, “*Sonic Visualiser: An Open Source Application for Viewing, Analysing, and Annotating Music Audio Files*,” in *Proceedings of the ACM Multimedia International Conference*, 2010. One precedent for such an analytical approach to popular music analysis is Steven Rings, “A Foreign Sound to Your Ear: Bob Dylan Performs ‘It’s Alright, Ma (I’m Only Bleeding),”” *Music Theory Online* 19, no. 4 (2013).

<sup>34</sup> Cook, “Methods for Analysing Recordings,” 226.



be used to generate pitch and therefore relative intonation; syllable onsets, which can be used to generate inter-onset intervals and therefore relative timing or speech rhythm; etc. When reading a spectrogram, the horizontal axis shows time, the vertical axis shows frequency (indicated by a scale in Hertz on the left), while differences in amplitude are shown by color and/or shading (Figure 2.1).<sup>35</sup>

**Figure 2.1 Log-frequency spectrograms generated by Sonic Visualiser for an excerpt from “Dave Brubeck” (1959) by Jack Kerouac and Steve Allen.**

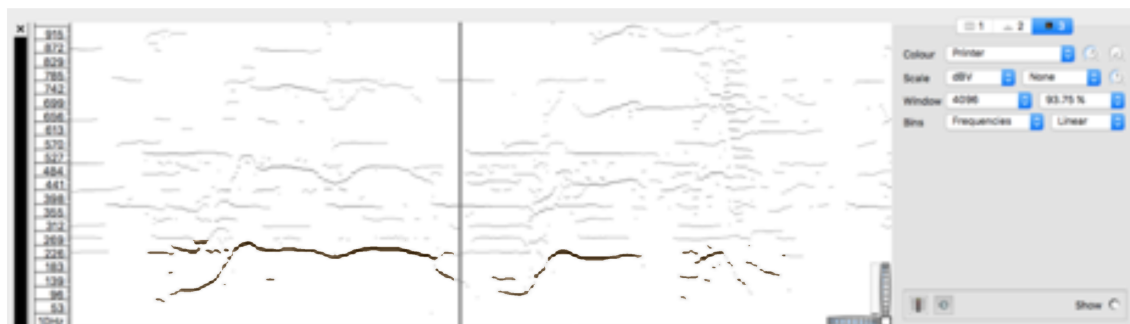


In using Sonic Visualiser to examine recordings where speech and music are occurring together, it can at times be challenging to isolate speech events from musical instruments in a normal log-frequency spectrogram. An alternative spectrogram representation that proves useful in differentiating speech from other instruments is a peak-frequency spectrogram, which isolates the central frequencies of each visible formant in a signal (Figure 2.2). Both log and peak frequency spectrograms will be used throughout this chapter.

---

<sup>35</sup> While occasionally figures will include spectrogram or waveform images for the reader to consider, at times and wherever possible, I will represent the information gleaned through the recording analysis software, in musical notation or in simplified timeline grids.

**Figure 2.2 Peak frequency spectrogram generated by Sonic Visualiser for an excerpt from “Dave Brubeck” (1959) by Jack Kerouac and Steve Allen.**



Another advantage to analysis software such as Sonic Visualiser is the ability to generate precise timing and frequency measurements through the use of various Vamp plugins. A Vamp plugin is a binary module that can be loaded by a host application and fed audio data. Vamp plugins do not generate audio output, but rather some sort of symbolic information calculated from the audio data. Typical things that a Vamp plugin might calculate include the precise location of note or event onset times, visualizable representations of the audio such as spectrograms, or curve data such as amplitude or fundamental frequency.<sup>36</sup> For the purposes of my investigation of the rhythmic interactions of speech and music, the Vamp plugins most useful to me are those that calculate event onsets (for which I use a tempo and beat tracker alongside a note-onset detector).<sup>37</sup> For the purposes of my investigation of the melodic interactions of speech

<sup>36</sup> For more information about the benefits of Vamp plugins, see <https://www.vamp-plugins.org/rationale.html> (accessed September 23, 2019).

<sup>37</sup> C Duxbury et al., “Complex Domain Onset Detection for Musical Signals,” *Proceedings of the 6th Conference on Digital Audio Effects*, 2003. Also Matthew E.P. Davies and Mark D. Plumbley, “Context-Dependent Beat Tracking of Musical Audio,” *IEEE Transactions on Audio, Speech and Language Processing* 15, no. 3 (2007).

and music, the Vamp plugins most useful to me are those that calculate fundamental frequencies (for which I use a probabilistic YIN plugin for pitch and note tracking).<sup>38</sup>

Vamp plugins are an essential part of recording analysis, in no small part because of their ability to calculate precise and complex measurements, but despite their precision, no plugin can ever fully replace the measurements determined through close listening conducted by an actual human being. Thus, in addition to using the Vamp plugins mentioned above, the following analyses also reflect my own listening experience, as all measurements are verified through comparison with my own manual data extractions.

While Sonic Visualiser and the various Vamp plugins designed for it are extremely useful for recording analysis, at times it was also necessary to use additional recording software Audacity and Adobe Audition which, unlike Sonic Visualiser, are designed as editing software rather than analysis software. These programs were used to cut tracks into their constituent phrases in order to compare them and to carefully isolate Waits's vocal track from the other instrumental sounds with minimal impact on the vocal signal.

---

<sup>38</sup> The name pYIN is a probabilistic (thus the 'p' in the name) variant of the YIN algorithm for measuring fundamental frequencies in recorded speech. YIN, in this case, does not stand as an acronym for anything but rather refers to the 'yin' of yin and yang, and alludes to the interplay between autocorrelation and cancelation that it involves. For more detailed explanation of YIN and pYIN, see Matthias Mauch and Simon Dixon, "PYIN: A Fundamental Frequency Estimator Using Probabilistic Threshold Distributions," in *IEEE International Conference on Acoustics, Speech and Signal Processing*, 2014. Also Matthias Mauch et al., "Computer-Aided Melody Note Transcription Using the Tony Software: Accuracy and Efficiency," *Proceedings of the First International Conference on Technologies for Music Notation and Representation*, 2015.

Isolating Waits's vocals was necessary in particular for tracking and analyzing the intonation of Waits's speech, as the plugins designed for pitch tracking and fundamental frequency(f0) estimation were primarily designed for monophonic recordings. All data collected from modified recordings was compared back to the original recording for verification, and all computer-generated data was compared with my own manual transcriptions.

## 2.3 Speaking in/of Time

In both music and speech, the patterning of events gives rise to hierarchically organized rhythms that are perceived as metrical.<sup>39</sup> Meter can be said to dynamically modulate attention in time in such a way that processing efficiency is influenced. This possibly takes place in a similar way in music and speech, as shown by impacts of musical meter on the phonological processing of speech.<sup>40</sup> With regard to their principles of organization, speech rhythm and musical rhythm are comparable arrangements of contrasting elements: the former is grounded on patterns of stress, while the latter is based on recurring durations.

Despite a long-standing interest in the overlapping neurological and cognitive processes involved in perceiving speech and musical rhythm, the vast majority of speech-music

---

<sup>39</sup> In speech, this hierarchical organization is called prosody; prosodic structures share significant similarities with musical meters, grouping, and phrase-rhythmic structures. Caroline Palmer and Sean Hutchins, "What Is Musical Prosody?," in *Psychology of Learning and Motivation: Advances in Research and Theory*, ed. B.H. Ross (Amsterdam: Elsevier Press, 2006), 245–278.

<sup>40</sup> Nia Cason and Daniele Schön, "Rhythmic Priming Enhances the Phonological Processing of Speech," *Neuropsychologia* 50, no. 11 (2012): 2652–58.

rhythm studies rely on stimuli that present speech and music separately. On the musical side, studies have explored how listeners perceive temporal regularity in music performances.<sup>41</sup> On the speech side, studies have explored how temporal regularity influences a listener's perception of rhythm in spoken-word.<sup>42</sup> Yet, little research explores what happens when speech and music occur simultaneously, as is the case in spoken-word song. Moreover, the growing literature exploring the rhythmic relationship between speech and music has a tendency to rely on stimuli that are either entirely devoid of (or at best, limited in) artistic expression. In perception studies that do explore how musical rhythm influences our perception of speech rhythm, musical stimuli is regularly metronome clicks and speech stimuli is one-off sentences recorded expressly for the sake of experiment.<sup>43</sup>

While such studies are immensely valuable in their own ways, they leave unexplored the large and impressive realm of speech acts strategically designed for artistic expression. Neuroscientist Aniruddh Patel recognized this oversight in his 2008 book, *Music, Language, and the Brain*, suggesting that a new direction for research in the processing of speech rhythm as it relates to musical rhythm might consider how the perception of verbally improvised speech changes when accompanied by metrical music. Such a study could shed light on the “temporal relations between verbal and musical accent points as a

---

<sup>41</sup> Edward W. Large and Caroline Palmer, “Perceiving Temporal Regularity in Music,” *Cognitive Science* 26, no. 1 (2002): 1–37.

<sup>42</sup> Hugo Quené and Robert F. Port, “Effects of Timing Regularity and Metrical Expectancy on Spoken-Word Perception,” *Phonetica* 62, no. 1 (2005).

<sup>43</sup> Nia Cason, Corine Astésano, and Daniele Schön, “Bridging Music and Speech Rhythm: Rhythmic Priming and Audio-Motor Training Affect Speech Perception,” *Acta Psychologica* 155 (2015): 43–50.

piece unfolds in time.”<sup>44</sup> The following analytical examination endeavors to demonstrate the fruitfulness of Tom Waits’s unique style of spoken-word song for investigating how speech and music can influence each other when combined artistically.

Of all Tom Waits’s spoken-word songs, one of my personal favorites is his 2002 recording of “Watch Her Disappear” from the studio album *Alice*. The album is a selection of songs composed over a decade earlier for an opera of the same name, created in collaboration with director Robert Wilson and designer Paul Schmidt, and premiered in December of 1992 by the Thalia Theatre in Hamburg Germany.<sup>45</sup> The opera (and indeed, the album as well, though less explicitly) is about Lewis Carroll (b. Charles Lutwidge Dodgson), author of *Alice in Wonderland* (1865) and *Through the Looking Glass* (1871), his obsession with the young Alice Liddell, and his confusion between dream and reality. In “Watch Her Disappear,” Carroll observes Alice from afar, watching her change from day clothes to evening dress, and go out to a dance (Figure 2.3).

The song brings to the fore Carroll’s confusion between dream and reality, as the first line explains it is not a dream, but rather a dream of a dream. The text moves in and out of present and past tense, further reinforcing the dreamlike state, as dreams so often can seem to be at once then, and now, and never.

---

<sup>44</sup> Anirudh Patel, *Music, Language, and the Brain* (New York: Oxford University Press, 2008), 158.

<sup>45</sup> Performance information can be found at <http://www.robertwilson.com/alice>. Accessed September 28, 2019.

**Figure 2.3 Text and musical sections for "Watch Her Disappear" by Tom Waits.**

Text	Musical Section <sup>46</sup>
Last night I dreamed that I was dreaming of you And from a window across the lawn I watched you undress Wearing your sunset of purple tightly woven around your hair That rose in strangled ebony curls Moving in a yellow bedroom light The air is wet with sound	mm.1-16
The faraway yelping of a wounded dog And the ground is drinking a slow faucet leak Your house is so soft and fading as it soaks the black summer heat A light goes on and the door opens And a yellow cat runs out on the stream of hall light and into the yard	mm.17-32
A wooden cherry scent is faintly breathing the air I hear your champagne laugh You wear two lavender orchids One in your hair and one on your hip A string of yellow carnival lights comes on with the dusk Circling the lake with a slowly dipping halo And I hear a banjo tango	mm.33-48
And you dance into the shadow of a black poplar tree And I watched you as you disappeared I watched you as you disappeared I watched you as you disappeared I watched you as you disappeared	mm.49-64

The words and music of “Watch Her Disappear” are so closely intertwined that it is difficult to imagine one without the other, and the relationship between speech and music often reflects and reinforces aspects of the song’s overarching narrative. Perhaps the most familiar technique used to connect words and music is text/word painting. When Waits recites “wearing your sunset of purple tightly woven around your hair,” for example, the pump organ melody emulates the poetic metaphor by tightly weaving semitones above

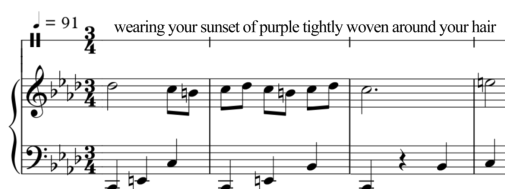
---

<sup>46</sup> Measure numbers refer to my piano transcription of the song, written in 3/4 time, and treating the first six unaccompanied melodic notes as an extended anacrusis to the first measure of the song. Each musical section is delineated by the repeated chord progression that underlies the song. Measures 1–16 represent the first cycle through the progression, and every subsequent section is either a reiteration of the same progression, or a slightly modified variation of it.

and below the chord-tone C (Figure 2.4a).<sup>47</sup> Similarly, when Waits recites “a faraway yelping of a wounded dog” the musical melody conjures up the very sound that Waits describes (Figure 2.4b). The pump organ melody leaps up a major 7<sup>th</sup>, landing on an accented non-chord tone that is held for nearly a full measure before resolving. The gesture, to my ear at least, imitates the sound of a dog yelping or howling “awhooooo!!”

**Figure 2.4 Two instances of word painting in "Watch Her Disappear" by Tom Waits.**

a) piano reduction and approximately placed words, 0:22–0:27



b) piano reduction and approximately placed words, 0:38–0:43



Text-painting is one device that draws connections between words and music, but it is certainly not the only technique that Waits employs, nor is it a technique that works directly to create connections between the act of speaking as a musical voice and the musical sounds. And while Waits undoubtedly does speak, rather than sing, the text of “Watch Her Disappear,” the rhythms of his speech operate in surprisingly musical ways. The relationship between speech and musical rhythms similarly reflect and reinforce aspects of the song’s narrative, albeit in considerably more nuanced ways than the moments of explicit word painting.

<sup>47</sup> Access to the studio recording of “Watch Her Disappear” is available through The Official Tom Waits YouTube Channel at the following link: <https://www.youtube.com/watch?v=MzXefc2fZ3Y>.



In all, the text of “Watch Her Disappear” comprises 257 syllables, of which over 100 have the potential to bear stress according to the norms and conventions of the English language and English prosody.<sup>48</sup> Of those that do bear stress, nearly three-quarters relate in some salient way to a beat or beat subdivision of the song’s triple meter groove. There are varying degrees of relatedness between stress syllables and groove beats, however, and before delving too far into a rhythmic analysis of the song, it seems prudent to explain the rationale for the various kinds of stress-beat relationships that occur.<sup>49</sup>

Much of the literature that explores expressive timing, micro-rhythm, or microtiming often focuses on jazz music, and many of the concepts and findings from that literature are useful to this study. The first key concept is the perceptual threshold for onset alignment. Early experiments in music perception found that discrepancies in onset timing of 20 milliseconds or less are unlikely to be perceived when they occur in musical or “pseudo-musical” settings. Building upon this, later studies, particularly those by Matthew Butterfield, found that in the context of jazz rhythm sections, “most ordinary

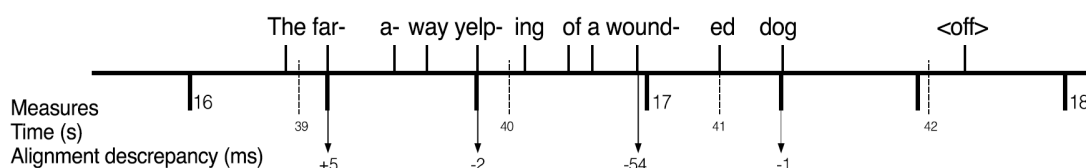
---

<sup>48</sup> For the interested reader, a vast amount of research on English speech rhythm and metrical stress exists, and perhaps the most easily accessible explanations of stress bearing parameters for English syllables can be found in Morris Halle, “The Stress of English Words 1968-1998,” *Linguistic Inquiry* 29, no. 4 (1998): 539–68.

<sup>49</sup> The notion of groove beats, rather than metrical beats, is adopted here primarily because the subject of inquiry is a recording of a song. In the absence of a notated score, the notion of meter is subject to the listener’s experience and the decision as to what, in the repeated rhythmic patterns that constitute the musical setting of “Watch Her Disappear,” constitutes a beat is determined in large part by how the patterns laid down by the bass make me want to move. See Allan Moore, *Rock: The Primary Text: Developing a Musicology of Rock*, 2nd ed. (Burlington, VT: Ashgate, 2001), 34. For more on groove from a music-theoretical perspective, see Mark Butler, *Unlocking the Groove: Rhythm, Meter, and Musical Design in Electronic Dance Music* (Bloomington: Indiana University Press, 2006), 339–40. Also Charles Keil and Stephen Feld, *Music Grooves: Essays and Dialogues* (Chicago: University of Chicago Press, 1994).

listeners, irrespective of music training or stylistic preference, are unable to discern a discrepancy of 30 milliseconds or less between bass and drums in the jazz rhythm section with any consistency.”<sup>50</sup> In Fernando Benadon’s research on expressive timing in jazz soloists, however, the perceptual threshold for onset alignments seems to increase substantially, as soloists using expressive timing tend to produce asynchronies in a range of 50–80 milliseconds in relation to the underlying beat in order to emphasize and contour their improvisations, and “timing deviations are evident and intentional-sounding only at this magnitude and above.”<sup>51</sup> Many observations made in studies of expressive timing in jazz parallel the rhythmic activity of Waits’s speech in his spoken-word songs. Take, for example, the relationship between stressed syllable onsets and groove beats in two utterances from Waits’s recording of “Watch Her Disappear” (Figure 2.5).<sup>52</sup>

**Figure 2.5 Onset time-point graph and alignment discrepancies in line 7 from “Watch Her Disappear” by Tom Waits (0:37–0:43).**



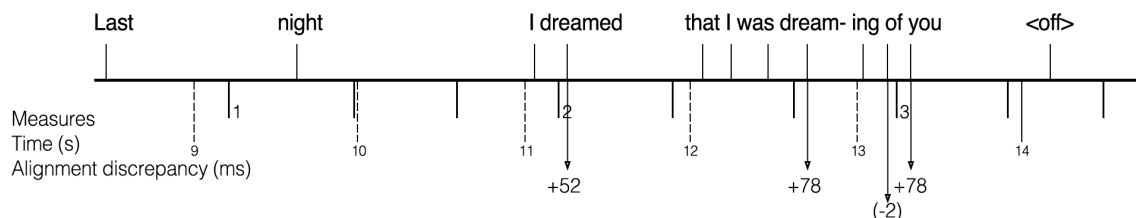
<sup>50</sup> Matthew W. Butterfield, “Participatory Discrepancies and the Perception of Beats in Jazz,” *Music Perception: An Interdisciplinary Journal* 27, no. 3 (2010): 165.

<sup>51</sup> Fernando Benadon, “Time Warps in Early Jazz,” *Music Theory Spectrum* 31, no. 1 (2009): 6. Indeed, in many of Benadon’s examples, temporal deviations between groove beats and events in melodic solos can be as large as 330 milliseconds (see for instance, Example 5 from Benadon’s study).

<sup>52</sup> Measurements are made with Sonic Visualiser through Vamp plugin calculations and verified by my own manual derivations.

Listening through both utterances, all stressed syllables sound aligned with beat onsets in the musical groove. Upon closer examination of the precise microtiming, however, onset deviations between syllables and beats range from 0 to 54 milliseconds. Much like the jazz solos studied by Benadon, Waits’s speech reflects a higher threshold for perceptual alignment between the solo instrument (in this case the words) and the musical groove upon which the solo takes place. At times, however, the deviations between beat and stressed syllable onsets reach magnitudes well beyond 80 milliseconds, and in these cases the concept of participatory discrepancies becomes useful. Take the opening utterance of the song (Figure 2.6). While the first two syllables are audibly offset from the groove beats, several stressed syllables are closely related to the musical rhythms. The syllable “dreamed,” occurring 52 milliseconds after the downbeat of m. 2, ostensibly aligns with the musical groove, while the syllables “dream-” and “you” occur 78 milliseconds after beats 3 of m. 2 and the downbeat of m. 3, respectively.<sup>53</sup>

**Figure 2.6 Onset time-point graph and alignment discrepancies in line 1 from “Watch Her Disappear” by Tom Waits (0:08–0:15).**



The onset difference between syllable and beat in these instances is unlikely to be perceived as salient rhythmic durations, but rather as participatory discrepancies, in the sense that they represent the kind of “differences that make a difference,” as Bateson

<sup>53</sup> This duration is roughly equal to a 32<sup>nd</sup> note in relation to the song’s overall tempo of 91 beats per minute.

describes the phenomenon.<sup>54</sup> In these two specific cases, they are still a kind of alignment, but the kind which Butterfield and others call “behind the beat.”<sup>55</sup>

For the syllable “dream,” one could argue that the reason it can be heard as a participatory discrepancy is because there is no other syllable in closer proximity to the beat at that moment; but the same cannot be said for the syllable “you.” One might be inclined to suggest that the syllable “of”—which nearly synchronized with the downbeat of m. 3—is aligned with the beat as that such an alignment should trump any relationship between that beat and the subsequent syllable “you.” Such an argument would, however, be to privilege measurement over experience. Perceiving beat and syllable alignments in spoken-word songs involves more than just temporal proximity of event onsets—it also involves the prominence relations of those events.

A syllable that does not bear stress within an utterance is unlikely to be perceived as aligned with a groove beat, regardless of its proximity to that beat because, prosodically speaking, it cannot act as the beginning of a rhythmic unit (except in cases where mispronunciation—*emphases* on the wrong syllables—is deliberately exploited). In this way, despite the 78 milliseconds difference between beat and syllable onset, “you” is quite easily perceived as aligned with the downbeat of m. 3 through a “behind the beat” participatory discrepancy.

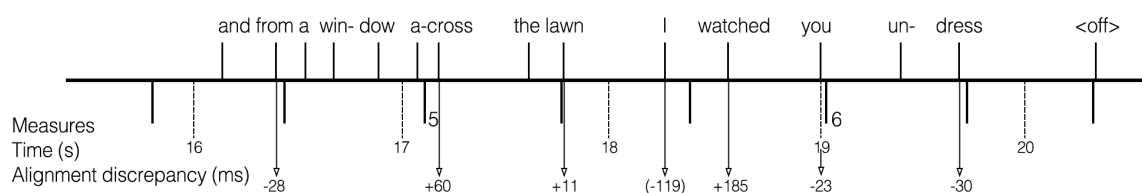
---

<sup>54</sup> Gregory Bateson, *Steps to an Ecology of Mind: Collected Essays in Anthropology, Psychiatry, Evolution, and Epistemology* (New York: Ballantine Books, 1972), 16.

<sup>55</sup> Butterfield, “Participatory Discrepancies and the Perception of Beats in Jazz,” 160.

Another interesting example of participatory discrepancies can be heard in the second line of “Watch Her Disappear,” wherein alignment differences between groove beat and syllable onsets are as large as 185 milliseconds—nearly a triplet eighth-note in relation to the quarter note pulse (on average, 659 milliseconds long) moving at 91 beats per minute (Figure 2.7).

**Figure 2.7 Onset time-point graph and alignment discrepancies in line 2 from “Watch Her Disappear” by Tom Waits (0:15–0:21).**



In this line, a number of syllables (“from,” “lawn,” “you,” and “-dress”) are well within the 50 milliseconds threshold for perceptual alignment with groove beats. and the syllable “-cross”, occurring 60 milliseconds after the downbeat of m. 5, involves a behind-the-beat participatory discrepancy. But the relationship between groove beats and the syllables “I” and “watched” is somewhat more complicated to explain. On one hand, it is not impossible to hear the syllable “watched” as an 185 milliseconds behind-the-beat participatory discrepancy in relation to the third beat of m. 5, despite both the magnitude of the discrepancy and the fact that the syllable “I” is closer in proximity to said beat.

On the other hand, it is also possible to hear neither “I” nor “watched” as aligned with groove beats at all, but rather as a triplet rhythm over two beats in the groove—what

Harald Krebs would call a “metrical grouping dissonance” of the G3/2 type.<sup>56</sup> This ambiguity of the rhythmic interactions between speech rhythm and groove is influenced by the prosodic malleability of the utterance itself: both syllables “watched” and “you” have the potential to bear prosodic stress.

English is not known for any mutability of word accent, as the relative emphasis of syllables within a word is not governed by a rhythmic rule but rather is a property of the word. But phrasal stress patterns are, at times, variable, so long as they adhere to the rules of English prosody. These rules are best expressed by metrical stress theory, which posits that “the phonetic and phonological differences between stress and ordinary features can best be accounted for if one abandons the assumption that stress is a feature. Instead, the theory represents stress as a hierarchically organized rhythmic structure.”<sup>57</sup> In treating stress this way, the rules to which stress must adhere are as follows:

- 1) Stress is hierarchical, with stress languages having multiple degrees of stress, and these degrees can appear within the phonology of a language rather than being a result of phonetic rules.<sup>58</sup>

---

<sup>56</sup> Harald Krebs, *Fantasy Pieces: Metrical Dissonance in the Music of Robert Schumann* (New York: Oxford University Press, 1999). While the inter-onset intervals of the three syllables in question are not exactly isochronous, their time-spans are close enough to be heard as perceptually equivalent and expressively timed.

<sup>57</sup> Bruce Hayes, *Metrical Stress Theory: Principles and Case Studies* (Chicago: University of Chicago Press, 1995), 26.

<sup>58</sup> The phonological and prosodic segments of English speech, from smallest to largest, that contribute to the rhythmic structure of an utterance include the syllable, metrical foot, phonological word, phonological phrase, and intonational phrase. These segments serve as *domains*, or *layers*, for stress in a metrical theory of English speech rhythm.

- 2) Stress is cumulative, such that each word or phrase culminates in a single strongest syllable bearing main stress.
- 3) Stress is rhythmically distributed, such that syllables bearing equal levels of stress tend to occur at roughly equidistant intervals from each other, falling into patterns of stress alternation.
- 4) Stress lacks assimilation, such that a stressed syllable does not induce stress on the immediately preceding or subsequent syllable.

Metrical phonologists favour a grid structure to represent stress patterns, where stress is indicated by an X at the various levels of prosodic hierarchical structure. English, being a quantity-sensitive language, disavows stress on light syllables, such that in the case of the phrase “I watched you undress,” four of the five syllables can potentially bear stress: “I,” “watched,” “you,” and “-dress.”<sup>59</sup> The normative stress pattern for the utterance would assign second level prominence to the syllable “watched” and third level prominence to the syllable “-dress” (Figure 2.8). This pattern is normative because in English phrase stress is directional, with right-most syllables favored. In other words, the syllable

---

<sup>59</sup> Quantity-sensitivity refers to syllable weight, and what elements make up a particular syllable. Syllables, formally defined as “a phonological constituent composed of zero or more consonants, followed by a vowel, and ending with a shorter string of zero or more consonants,” can be heavy or light. John Goldsmith, *Autosegmental and Metrical Phonology* (Oxford: B. Blackwell, 1990), 108. A heavy syllable contains either a long vowel or a coda consonant. The English words “you” and “watched” exemplify the two types of heavy syllable. A syllable shorter than a heavy syllable is called a light syllable. For example, if C=consonant and V=vowel, then CV syllables are light in all contexts; CVC syllables are heavy in non-final contexts, but light finally; CVV and CVVC syllables are heavy in all contexts. See also Hayes, *Metrical Stress Theory: Principles and Case Studies*, 56–57.

“watched” bears greater stress because the syllable “-dress” bears the greatest stress and, in doing so, disavows stress on the syllable “you.”

**Figure 2.8 Normative metrical stress grid for line 2 from “Watch Her Disappear” by Tom Waits.**

phrase			X				X			X				
word			X		X		X		<b>X</b>	X				
syllable	X	X	X	X	X	X	X	X	X	X				
	and	from	a	win-	dow	a-	cross	the	lawn	I	watched	you	un-	dress

There is, however, a circumstance in which the greatest stress of the utterance would fall on “you” rather than “-dress,” should the speaker wish to emphasize *who* they are watching over *what* they are watching someone do. Stressing the syllable “you” over the syllable “-dress” without any other change to the stress pattern would cause a stress clash between syllables “watched” and “you,” which violates the rule that stress is rhythmically distributed (commonly referred to as the Rhythm Rule). To resolve this, stress shift takes place, transferring the stress of “watched” leftwards. While the syllable “I” has the potential to bear stress independently, in the context of the full utterance the stress immediately preceding “I” prevents it from bearing second level stress, so the stress of “watched” shifts into the stress of “lawn” (Figure 2.9).

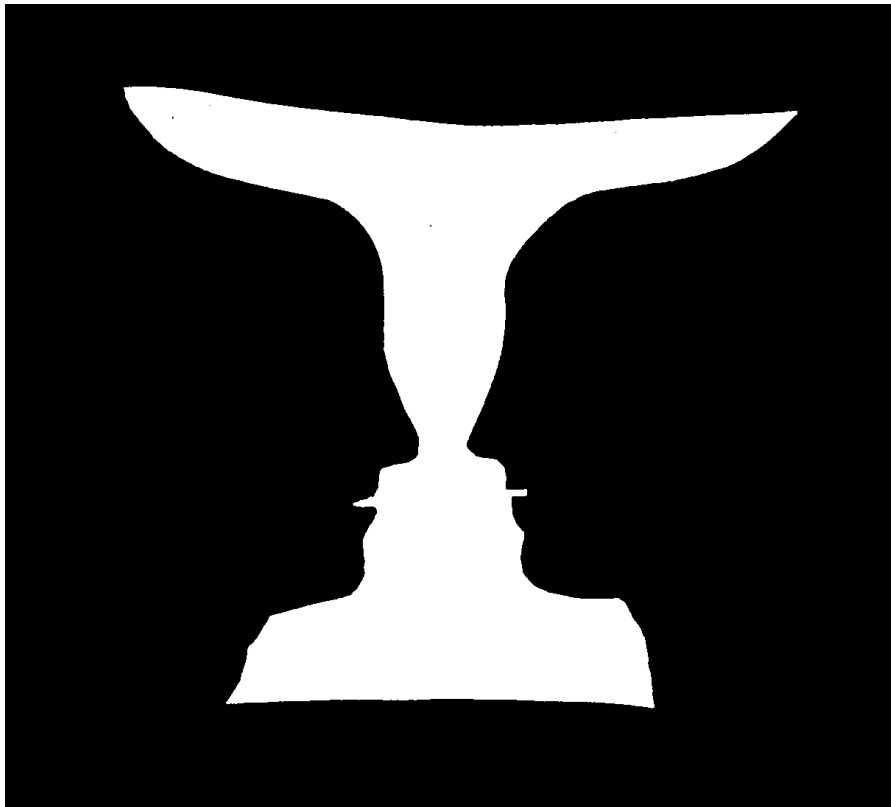
**Figure 2.9 Alternate metrical stress grid with stress shifts for line 2 from “Watch Her Disappear” by Tom Waits.**

phrase			X				X			←	←			
word			X		X		X		←	<b>X</b>	←			
syllable	X	X	X	X	X	X	X	X	X	X	X			
	and	from	a	win-	dow	a-	cross	the	lawn	I	watched	you	un-	dress



The ambiguity between hearing the passage as expressively timed or hearing the passage as a metrical grouping dissonance strongly reflects my own hearing of the phrase. While I gravitate towards hearing the syllable “watched” as aligned with the groove, I can easily hear a triplet over two beats if I listen for it. The aural experience of a phrase such as this one is reminiscent of visual experience of the familiar Rubin’s Vase, a set of ambiguous or multistable forms developed by Danish psychologist Edgar Rubin in the early twentieth century (Figure 2.10).

**Figure 2.10. “Rubin’s Vase,” afbildn 3 by Edgar Rubin.<sup>60</sup>**



---

<sup>60</sup> Edgar Rubin, *Synsoplevede Figurer: Studier i Psykologisk Analyse* (København og Kristiania: Gyldendal, Nordisk Forlag, 1915), 30. Available in digital format through the

The illusion presents the viewer with a mental choice of two equally valid interpretations: that of two black faces with a white background or a white vase with a black background. While one interpretation will often strike the viewer first, given even a small amount of time the viewer invariably recognizes the other and will alternate in quick succession between the two. Analogously, the passage from “Watch Her Disappear” involves a cognitive dissonance that is present when tracking both speech rhythms and musical rhythms simultaneously. The listener confronts a mental choice of two equally valid interpretations and, while one or the other may initially resonate with the listener, given some time and closer attention the listener will invariably recognize the other and perhaps enjoy the tension between the two.

This kind of rhythmic relationship between speech and music only increases as the song unfolds. In the second verse, yet another rhythmic device serves to link the words and the music in “Watch Her Disappear,” namely, rhythmic imitation. In the phrase, “you wear two lavender orchids, one in your hair and one on your hip,” not only are there prominent alignments between stressed syllables, but there is also a relationship between the rhythmic profiles of the speech and the pump organ melody (Figure 2.11). Both speech and melody employ triplet figures on the first and third beats of m. 35, as though they are playing and speaking together.

---

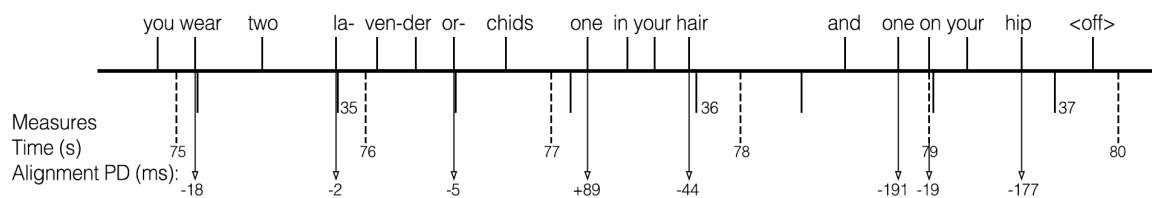
[Center for Research Libraries](#). Accessed September 25, 2019. A similar reference to Rubin’s Vase appears in Diana Deutsch, “The Perceptual Organization of Streams of Sound,” in *Musical Illusions and Phantom Words* (New York: Oxford University Press, 2019). DOI: 10.1093/oso/9780190206833.001.0001. Accessed September 29, 2019.

**Figure 2.11 Comparison of speech rhythm and melodic rhythm for lines 14-15 from “Watch Her Disappear” by Tom Waits (1:15–1:20).**



The aural sense of rhythmic connection continues into the subsequent measure but, in the same way as the previous example, involves some metrical (and cognitive) dissonance. An examination of the phrase from an onset time-point perspective reveals a close alignment of stressed syllables with the musical groove in the first half of the phrase, while the second half is slightly displaced (Figure 2.12).

**Figure 2.12 Onset time-point graph and alignment discrepancies in lines 14–15 from “Watch Her Disappear” by Tom Waits (1:15–1:20).**



The stressed syllables “wear,” “lav-,” and “or-” are all perceptually aligned with the groove beats, while the first iteration of the syllable “one,” and the subsequent stressed syllable “hair” both involve a somewhat more substantial degree of participatory discrepancies. The final portion of the utterance, however, involves significantly larger onset differences between the speech stresses and groove beats, somewhere between a 16<sup>th</sup> note (165 milliseconds) and a triplet 8<sup>th</sup> note (220 milliseconds). The mental choice

presented to the listener here involves two possibilities: to hear the large onset deviations between stressed syllables and beats as “on top of the beat” participatory discrepancies; or to hear the large onset deviations as a kind of metrical displacement dissonance of the kind Krebs would call D3–1, wherein the triplet figure formed by the syllables “one on your” shifts back by one triplet 8<sup>th</sup> note.<sup>61</sup>

Both these two examples serve to demonstrate some of the rhythmic ways in which Waits “makes the words and music fit.” More often than one might expect from a cursory listen to the song, prominent syllables in the text audibly align with groove beats. However, the perception of such alignments, especially in moments where close examination reveals surprising onset deviations, is in large part a product of the differences between what one listens for in speech as compared to music. In speech, prominent events are determined first and foremost by *where* they are in relation to other events around them, while in music prominences (particularly, groove beats) are determined first and foremost by *when* they occur in relation to events pre- and proceeding. Thus, the desire to hear prominent syllables as aligned with groove beats involves tracking two different streams of expectation simultaneously. The two parallel streams of expectation are coordinated, structured, and perceived in surprisingly similar ways, such that the brain concurrently anticipates where the next stress will be and when the next beat will take place.<sup>62</sup> In a

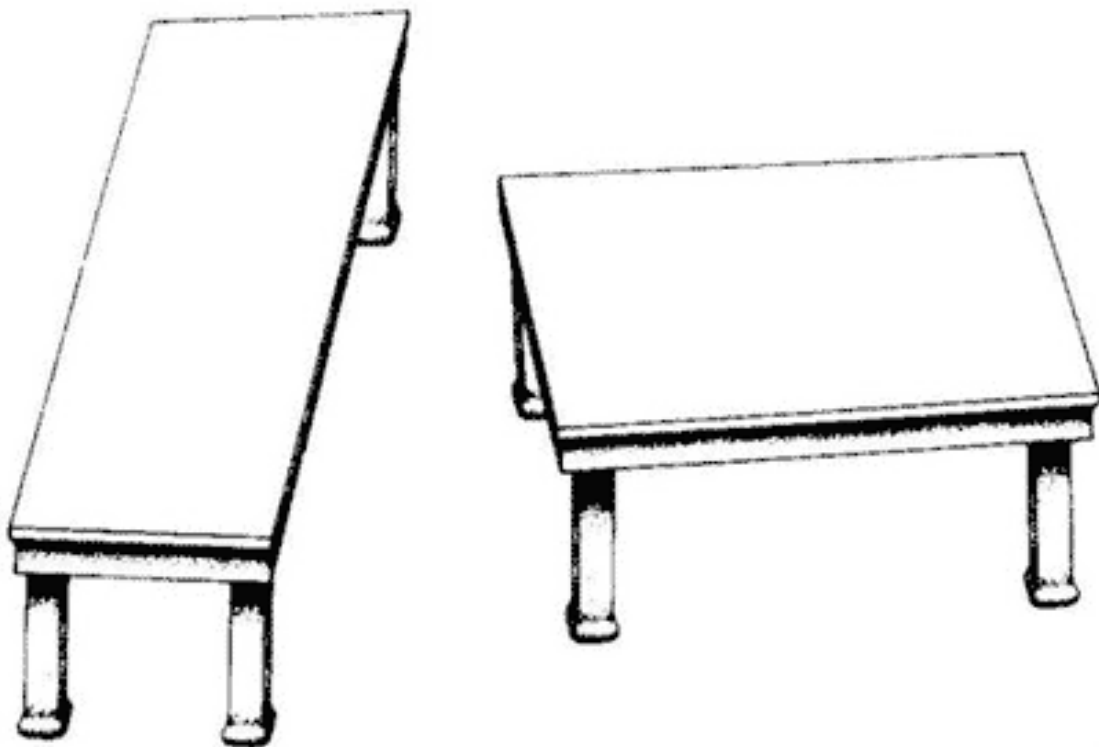
---

<sup>61</sup> Of course, in all instances presented there is the possibility of hearing no relationship between the rhythms of the speech and the musical groove, but to listen to spoken-word songs in this way is not unlike looking at Rubin’s Vase as neither two black faces nor as one white vase but rather as the line that separate the black from the white, in which case all one sees is a rather poorly drawn vase.

<sup>62</sup> This hypothesis is echoed in Chantal Lemire, ‘At the “Crossroads”’: Interactions Between Speech Rhythm and Musical Rhythm in Tom Waits’s Spoken-Word Songs’

song that superimposes the natural rhythms of speech with metrically organized music, one's expectation of when a pulse will occur and where a stress will occur can quite easily conflate; furthermore the conflation of these cognitive processes may explain why such large onset deviations between events can be perceived as far less egregious than they might otherwise be in some other context. Such perceptual alignments can be thought of as aural analogs to optical illusions. Take, for example, Roger Shepard's familiar table illusion (Figure 2.13).

**Figure 2.13. *Turning the Tables* (1990) by Roger Shepard.<sup>63</sup>**




---

(MA Thesis, University of British Columbia, 2013), and supported by much of the research on the neurological intersections of speech and music explored by various scholars over the last decade.

<sup>63</sup> Roger N. Shepard, *Mind Sights: Original Visual Illusions, Ambiguities, and Other Anomalies, with a Contemporary Commentary on the Play of Mind in Perception and Art*

The table on the left appears longer and thinner than the table on the right but they are actually identical in dimensions, simply rotated 90 degrees. The two tables appear different because of the viewer's perspective towards them. The brain interprets the image on the retina according to past experience interacting with the world: how we see what we see depends significantly on context, even when we see the simplest qualities of the world. One can say the same for aural illusions, such as those that occur in Waits's spoken-word songs and those aural illusions demonstrated in "Watch Her Disappear." In the context of a metrically organized musical groove, the text recitation, despite its rhythmic irregularities, is interpreted in relation to the regularity of the music by which it is supported. The aural illusion of alignment occurs when the two different projections merge together. Their similarities enable the merger, but their differences produce a distorted impression of what actually sounds.

The ways in which Waits "makes the words and music fit" suggested thus far—from word painting to stress-beat alignments, to rhythmic imitations—are perhaps only the tip of the iceberg with respect to the text-music relationships that Waits creates. At a deeper level, Waits's particular way of speaking the text of "Watch Her Disappear" (and indeed the various ways he combines speech and music in all his spoken-word songs) reflects much of the over-arching programmatic themes that shape the narrative. In this case, the song both is and isn't a dream about a relationship that is both deeply felt and entirely fantastic.

---

(New York: W.H. Freeman & Co., 1990), 127–28. Reprinted under the fair dealing exception ([Section 29](#)) provided by the Canadian Copyright Act.

In the same way, the aural illusions and cognitive dissonances that unfold for the listener in “Watch Her Disappear” capture this very tension between dream and reality. By straddling the boundaries between speech and song, Waits ensures that the “words and music fit” without compromising the fundamental principles of speech rhythm. He may not entirely be singing, but he is undoubtedly speaking in time.

To a certain extent, the same can be said about how Waits’s speech intonation interacts with the harmonic and melodic materials of his spoken-word songs. This may be hard to imagine as yet, if the only spoken-word songs by Tom Waits with which the reader is familiar are the two thus far discussed, as in these (and in many of his spoken-word songs) Waits’s vocal timbre is so complex that perceiving any concrete pitch information can be quite difficult. Nevertheless, Waits has a number of spoken-word songs where the tonal properties of speech and music are intricately linked. The following section explores Waits’s use of pitch and vocal intonation to “make the words and music fit” in these songs.

## 2.4 Speaking in/a Tune

Speech is made to be heard as song, and this is achieved without transforming the sounds in any way... but simply by repeating a phrase several times over.<sup>64</sup>

---

<sup>64</sup> Diana Deutsch, *Phantom Words and Other Curiosities* (La Jolla: Philomel Records, 2003).

In 1995, Diana Deutsch discovered a fascinating auditory phenomenon whereby naturally flowing speech transformed into song when it was heard repeatedly.<sup>65</sup> Aptly named the speech-to-song illusion, Deutsch's discovery has since been confirmed by many empirical studies involving both musically trained subjects and non-musicians.<sup>66</sup> The many neurological implications notwithstanding, the speech-to-song illusion suggests not only that the characteristics which determine what we call speech and what we call song are closely linked, but also that the boundaries between language and music are perhaps more malleable than we tend to think.

---

<sup>65</sup> Diana Deutsch, *Musical Illusions and Paradoxes* (La Jolla: Philomel Records, 1995).

<sup>66</sup> Julia Merrill et al., "Perception of Words and Pitch Patterns in Song and Speech," *Frontiers in Psychology* 3, no. 76 (2012). Adam Tierney et al., "Speech versus Song: Multiple Pitch-Sensitive Areas Revealed by a Naturally Occurring Musical Illusion," *Cerebral Cortex* 23, no. 2 (1 February 2013): 249–54. Simone Falk, Tamara Rathcke, and Simone Dalla Bella, "When Speech Sounds like Music," *Journal of Experimental Psychology: Human Perception and Performance* 40, no. 4 (2014): 1491–1506. Christina M. Vanden Bosch Der Nederlanden, Erin E. Hannon, and Joel S. Snyder, "Everyday Musical Experience Is Sufficient to Perceive the Speech-to-Song Illusion," *Journal of Experimental Psychology: General* 144, no. 2 (2015): 43–49; and "Finding the Music of Speech: Musical Knowledge Influences Pitch Processing in Speech," *Cognition* 143 (2015): 135–40. Elizabeth H. Margulis, Rhimmon Simchy-Gross, and Justin L. Black, "Pronunciation Difficulty, Temporal Regularity, and the Speech-to-Song Illusion," *Frontiers in Psychology* 6 (2015): 1–7. Emily Graber, Rhimmon Simchy-Gross, and Elizabeth Hellmuth Margulis, "Musical and Linguistic Listening Modes in the Speech-to-Song Illusion Bias Timing Perception and Absolute Pitch Memory," *The Journal of the Acoustical Society of America* 142, no. 6 (2017): 3593–3602. Rhimmon Simchy-Gross and Elizabeth Hellmuth Margulis, "The Sound-to-Music Illusion: Repetition Can Musicalize Nonspeech Sounds," *Music & Science* 1 (2018): 1–6. Adam Tierney, Aniruddh D. Patel, and Mara Breen, "Acoustic Foundations of the Speech-to-Song Illusion," *Journal of Experimental Psychology: General* 147, no. 6 (2018): 888–904.



In 2011, Deutsch, Trevor Henthorn, and Rachael Lapidis proposed two central hypotheses about the relationship between speech and song.<sup>67</sup> First, that the neurological processes involved in detecting pitch salience is inhibited when listening to the normal flow of speech. Second, that the exact repetition of spoken phrases causes these processes to become disinhibited, enhancing the salience of pitch perception. All this shows the intrinsic musical and melodic properties of speech, setting the stage for exploring whether these melodies might be decipherable in Waits's spoken-word songs.

Tom Waits is probably best known for his unique vocal timbre, which one critic aptly describes as “soaked in a vat of bourbon, then hung out to dry in a smoke hut for a few months, only to be dragged outside and run over by a car. In other words, it may not be suited for singing at all.”<sup>68</sup> While Waits does seem to exploit the gravelly, ambiguous grain of his voice in many of his spoken-word songs, there are some in which the speech intonation is not only audible, but impressively musical. It was, in fact, precisely one of these songs that first piqued my interest in the musical properties of Spoken Word altogether. The song was “Crossroads” from the 1993 album *The Black Rider*—another studio recording of song selections from a musical play of the same name created in collaboration (once again) with director Robert Wilson.<sup>69</sup>

---

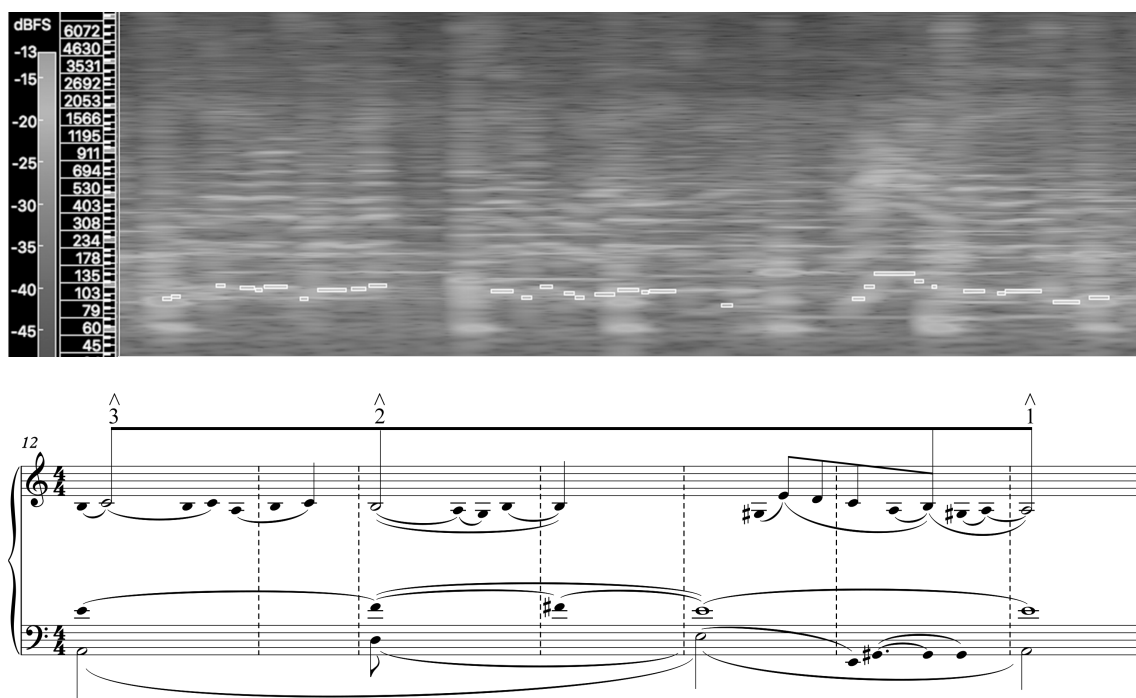
<sup>67</sup> Diana Deutsch, Trevor Henthorn, and Rachael Lapidis, “Illusory Transformation from Speech to Song,” *The Journal of the Acoustical Society of America* 129, no. 4 (2011): 2245–52.

<sup>68</sup> Gavin Wright, “Everything You Can Think of Is True: The Life and Music of Tom Waits,” in *Waits Mini Biography, Narrated by Marianne Hoff Engesland* (Grieg Audio: Audiobook, 2009).

<sup>69</sup> A complete analysis of the rhythmic interactions between speech and music in “Crossroads” is the central subject of Lemire, “At the 'Crossroads': Interactions Between

While my first attempts at analysis focused on examining the prosodic metricality of Waits’s speech in “Crossroads,” it was his melodic intonation that initially attracted me to the song. It was fascinating that, although I was perfectly cognizant that Waits was speaking the words, I could not help but sing along. Something about how he spoke seemed to allude to some latent melodic line. In other words, he seemed to be speaking “in tune.” Take, for example, the following brief passage from “Crossroads” (Figure 2.14).

**Figure 2.14 Annotated spectrogram and melodic/harmonic reduction for mm. 12–18 from “Crossroads” by Tom Waits (0:19–0:31).<sup>70</sup>**



Speech Rhythm and Musical Rhythm in Tom Waits’s Spoken-Word Songs.” Access to the studio recording of “Crossroads” by Tom Waits is available through The Official Tom Waits YouTube Channel at the following link:  
<https://www.youtube.com/watch?v=UUhGqSCS6tY>

<sup>70</sup> Measurements were determined by ear and verified through the pYIN: Notes plugin. As the plugin is designed to calculate all frequencies in the passage, occasionally discrete pitches from the instrumental voices made their way into the resultant data. Measure

Waits intones nearly every syllable on fundamental frequencies stable enough to be perceived—both by the ear and by a computational approach using a probabilistic fundamental frequency estimator—as discrete pitches. The stressed syllables in this line are consistently the most stable, and the whole phrase translates readily into musical notation, revealing that many of the frequencies upon which Waits intones are closely related to the harmonic progression driving the phrase. The emphasis on C3 in mm. 12–13 functions as the third of the underlying tonic (a minor) harmony, while the emphasis on B2 in m. 14 transforms the underlying d minor triad into a  $d^{+6}$  harmony.<sup>71</sup> Similarly, the emphasis on E3 in m. 16 doubles the root of the underlying dominant harmony while the B2 and G#2 in m. 17 function as the fifth and third of the ongoing dominant, respectively. Finally, in m. 18, the A2 serves as the root of the tonic harmony. Not only does the passage demonstrate how consistently Waits’s vocal intonation relates to the underlying harmonic progression of the song, it also demonstrates how the large-scale melodic shape of Waits’s spoken tones model familiar musical patterns, and a large-scale stepwise descent from mediant to tonic is distinctly audible over the course of the utterance.

Certain types of Waits’s spoken-word songs seem to exploit the melodic properties of speech more than others and, in particular, many of his spoken-word songs from the mid-1970s that involve jazz/blues musical settings and relatively lengthy, through-composed prose narratives are of this kind. The 1974 album *The Heart of Saturday Night* is the first

---

numbers refer to my own score transcription, in which the piece is notated in common time with a quarter note equal to approximately 465 milliseconds.

<sup>71</sup> Alternatively, one could also interpret this sonority as a B half-diminished seventh.

of Waits's studio albums to include spoken-word songs. Bones Howe, producer and engineer for the album, recalls that when the two first met, Waits specifically discussed his desire to include spoken-word on the album and in the end, two songs made the cut: "Diamonds on my Windshield," and "The Ghosts of Saturday Night (After Hours at Napoleone's Pizza House)."<sup>72</sup> The former was evidently recorded in one take, as Waits recalls, "Jim [Hughart] just started playing a modal bass line and I just started talking and Jim Gordon started playing a cool twelve-bar shuffle on brushes and we just winged it in one take..."<sup>73</sup>

Connections to mid-century American urban landscape painter Edward Hopper permeate the album's carefully crafted aesthetic. Known for painting still scenes of Americana, Hopper has been dubbed a documenter of "urban alienation," particularly through his scenes in which small groups of people occupy the same public spaces but with minimal to no interaction between them. Not just in artwork alone, *The Heart of Saturday Night* pays thoughtful homage to Hopper's glistening late night streets, diners and alleyway dwellers—and nowhere more explicitly than in the album's "footnote" track, which sums up many of the themes and characters that crop up throughout the album (cigarettes, cabs, sailors, gas station attendants). Titular connections between the album's title track and its closing spoken-word song ought not be lost on the reader, and where "The Heart of Saturday Night" offers the listener a heartfelt song, "The Ghosts of Saturday Night" offers, perhaps, some melodies' lingering ghost—ethereally and expertly hidden within

---

<sup>72</sup> Barney Hoskyns, *Low Side of the Road: A Life and Times of Tom Waits* (New York: Broadway Books, 2009), 107.

<sup>73</sup> Hoskyns, 111.

spoken word—that begs many repeated listens (and a listener’s willingness to believe in it at all) in order to take solid form.

Like many of Waits’s songs, “The Ghost of Saturday Night” involves irregular phrase rhythm, sensitive harmonic relationships, and delightful counterpoint. Over eight musical phrases that follow the same essential harmonic plan, Waits conjures an evening at an all-night diner located across the street from a Texaco station. Waits does not sing the words, but he is not simply speaking either; rather, he plays speech as a musical instrument. And, whereas in “Watch Her Disappear” the rhythmic properties of Waits’s speech seem the primary means through which words and music fit together, in the “Ghosts of Saturday Night” he exploits their melodic interplay. Studies on pitch perception in speech identify three key parameters necessary for speech events to register as melodically intoned: 1) increased stability of fundamental frequencies within syllables; 2) durational consistency between stressed syllables; and 3) increased regularity of acoustic signals across several syllables.<sup>74</sup> As with many of Waits’s spoken-word songs, not every syllable is spoken in such a way as to elicit a clear enough fundamental frequency so as to be perceived as this or that pitch; but enough syllables are spoken in this way that a melody emerges from among them. By working through each musical phrase, we can begin to piece together a larger, overarching melodic structure that ultimately connects the whole song together. The 1974 studio version comprises eight musical phrases, each following a similar harmonic progression and phrase rhythm, but with occasional phrase extensions here and there. Rendered in musical notation and taking the beat moving at approximately 60 beats

---

<sup>74</sup> Tierney et al., “Speech versus Song: Multiple Pitch-Sensitive Areas Revealed by a Naturally Occurring Musical Illusion,” 250.

per minute as the primary groove pulse, the structure of the musical phrases unfold as follows (Figure 2.15).<sup>75</sup>

**Figure 2.15 Summary of phrase structure for “The Ghosts of Saturday Night” by Tom Waits.**

A) Chord progression, phrases 1, 2, 3 and 7

B) Chord progression, phrases 4 and 5

C) Chord progression, phrase 6

D) Chord progression, phrase 8

Chord progressions for each section:

A)  $B^{\flat}add^{\flat}$   $E^{\flat 9}$   $B^{\flat}add^{\flat}$   $E^{\flat 9}$   $B^{\flat}add^{\flat}$   $A^{\flat}m7$   $D^{\flat}$   $G^{\flat}m$   $/F$   $E^{\flat 9}$

B)  $B^{\flat}add^{\flat}$   $E^{\flat 9}$   $B^{\flat}7$   $E^{\flat 7}$   $B^{\flat}7$   $E^{\flat 7}$   $B^{\flat}7$   $A^{\flat}7$   $D^{\flat}$   $G^{\flat}m$   $/F$   $E^{\flat}maj^{\flat}$

C)  $B^{\flat}add^{\flat}$   $E^{\flat 9}$   $B^{\flat}7$   $E^{\flat 7}$   $B^{\flat}7$   $E^{\flat 7}$   $B^{\flat}7$   $A^{\flat}7$   $D^{\flat}(\sharp 5)$   $G^{\flat}m$   $/F$   $E^{\flat}maj^{\flat}$   $B^{\flat}7$   $C^{\flat 9}/E^{\flat}$   $F^{\flat 9}/A$   $G^{\flat}m$   $/F$   $E^{\flat}maj^{\flat}$

D)  $B^{\flat}add^{\flat}$   $E^{\flat 9}$   $B^{\flat}7$   $E^{\flat 9}$   $B^{\flat}7$   $E^{\flat 7}$   $B^{\flat}7$   $A^{\flat}7$   $D^{\flat}$   $G^{\flat}m$   $/F$   $E^{\flat}maj^{\flat}$   $B^{\flat}add^{\flat}$   $E^{\flat 9}$   $B^{\flat}$

<sup>75</sup> Access to the studio recording of “The Ghosts Of Saturday Night (After Hours At Napoleone’s Pizza House)” is available through The Official Tom Waits YouTube Channel at the following link:

<https://www.youtube.com/watch?v=0uRX-XfpHr4>

The introductory phrase is four measures long and involves a two-beat extension of its final  $E\flat^9$  harmony. I consider this an extension of the phrase's final measure from four to six beats (rather than an additional two-beat, independent fifth measure) because it extends the standing harmony through repeated articulations of the same chordal figure at higher and higher octaves. In this way, the fifth beat of the measure does not accrue additional significance as a downbeat.

The same two-beat extension occurs on every final measure of each phrase in the song. Four of the eight phrases (the first three and the penultimate) follow the same harmonic progression as presented by the introduction, while the remaining four phrases involve various phrase extensions (summarized in Figure 2.15b through d). The fourth and fifth phrases are extended by an additional measure at their outset, repeating the alternation of  $B\flat^7$  and  $E\flat^9$  a third time (where in the four-measure version the alternation occurs only twice). The sixth phrase similarly extends by an additional measure in the first half, as well as by two additional measures in the second half: this occurs at the moment where the listener is primed to expect an extended  $E\flat^9$  harmony; but the extension is omitted and the final two measures of the progression (with an inversive variant on the  $D^7$ ) are stated one more time. The final phrase also extends by an addition of a measure alternating between  $B\flat^7$  and  $E\flat^9$  in its first half and extended again at its end to include a two measure cadential close to the song. To this musical setting Waits performs the text of "The Ghosts of Saturday Night" which, in printed versions, typically breaks down into five verses (Figure 2.16). For the most part, verses correspond to musical phrases, with the first and seventh musical phrases unaccompanied by text recitation.

**Figure 2.16 Table showing text and musical form for "The Ghosts of Saturday Night (After Hours at Napoleone's Pizza House)" by Tom Waits.**

	<b>Text Verses</b>	<b>Musical Phrases</b> <sup>76</sup>	<b>Time Stamps</b>
	<i>[Introduction]</i>	1 (mm.1-4)	0:00–0:19
1	A cab combs the snake, Tryin' to rake in that last night's fare, And a solitary sailor Who spends the facts of his life like small change on strangers...	2 (mm. 5-8)	0:19–0:37
2	Paws his inside P-coat pocket for a welcome twenty-five cents, And the last bent butt from a package of Kents, As he dreams of a waitress with Maxwell House eyes And marmalade thighs with scrambled yellow hair.	3 (mm. 9-12)	0:37–0:56
3	Her rhinestone-studded moniker says, "Irene" As she wipes the wisps of dishwater blonde from her eyes	4 (mm.13-17)	0:56–1:19
	And the Texaco beacon burns on, The steel-belted attendant with a 'Ring and Valve Special'... Cryin' "Fill'er up and check that oil" "You know it could be your distributor and it could be your coil."	5 (mm. 18-22)	1:19–1:42
4	The early mornin' final edition's on the stands, And that town cryer's cryin' there with nickels in his hands. Pigs in a blanket sixty-nine cents, Eggs - roll 'em over and a package of Kents, Adam and Eve on a log, you can sink 'em damn straight, Hash browns, hash browns, you know I can't be late.	6 (mm.23-29)	1:42–2:13
	<i>[Interlude]</i>	7 (mm.30-33)	2"13–2:31
5	And the early dawn cracks out a carpet of diamond Across a cash crop car lot filled with twilight Coupe Devilles, Leaving the town in a-keeping Of the one who is sweeping Up the ghost of Saturday night...	8 (mm. 34-41)	2:31–3:07

The second and third phrases of music correspond to the first and second verses of text, respectively. The third verse, however, spans two musical phrases, where the first two lines of text are recited over the fourth musical phrase and the remaining four lines of text are recited over the subsequent musical phrase. The final verse corresponds to the final musical phrase and involves substantial rubato over its last three measures and final line of text. Although this particular analysis does not focus on the rhythmic relationship

<sup>76</sup> Measure numbers are derived from my transcription of the recording, in which I choose to assign the quarter-note value to a duration spanning approximately one second.

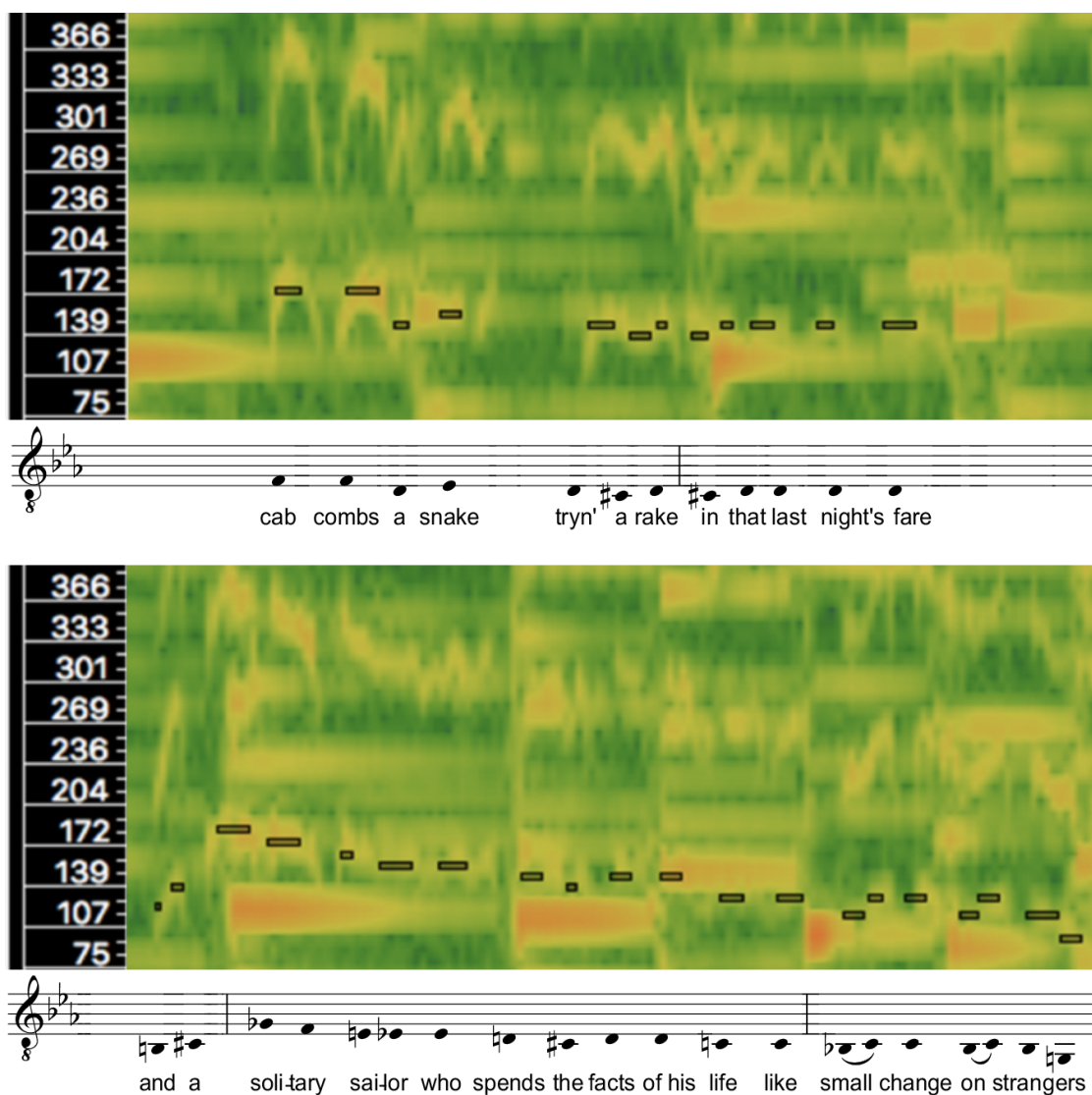


between speech and music in “The Ghosts of Saturday Night,” there is much that one could say for the musicality of Waits’s speech rhythm, as stressed syllables invariably align with correspondingly prominent groove beats and as many of the rhythmic devices demonstrated in the previous analysis are often – if not more so – employed here as well.

Equally significant to the musicality of “The Ghosts of Saturday Night,” however, is the intonation of Waits’s speech. One will hear impressive stability and consistency in the fundamental frequencies Waits employs as he delivers the text, as well as a high degree of acoustic regularity of intonation across larger phrase utterances. Consider, for example, the first spoken verse of “The Ghosts of Saturday Night” (Figure 2.17).

The annotated spectrogram shows, among other things, the stability and consistency with which Waits intones his speech, as even in a multi-track audio file, the frequencies and overtones of the spoken voice are visibly distinct. The scale to the left of the spectrogram shows the frequency range of the excerpt, and the black markers on the spectrogram identify the fundamental frequency ( $f_0$ ) extracted for each of the syllables in the utterance, rendered in musical notation below. The melodic profile of Waits’s speech often corresponds to chord tones in the underlying musical accompaniment and, if that is not enough, his vocal inflections tend to follow traditional voice-leading schema at both a chord-by-chord level and on a larger, structural level.

Figure 2.17 Annotated spectrogram for verse 1 from “The Ghosts of Saturday Night” by Tom Waits (0:19–0:37).



When the pitch rendering of the opening phrase of the song is placed in the context of the harmonic progression supporting it, we notice how closely connected the two seemingly distinct parts are (Figure 2.18).<sup>77</sup>

<sup>77</sup> The reader will undoubtedly recognize the influence of Schenkerian notational devices in this (and the several subsequent) melodic reductions, and a word should be said to explain the degree to which a Schenkerian approach is adopted. I do not intend the reader

**Figure 2.18** Reduction of opening verse from “The Ghosts of Saturday Night” by Tom Waits.

The first stressed syllable, “cab” is recited on F3, the fifth of the underlying Bb<sup>7</sup> harmony, followed by chromatic passing tones (omitted for clarity in this example) that lead to the next stressed syllable, “snake,” intoned on Eb3, the root of the subsequent Eb<sup>9</sup> harmony. Anticipating the downbeat of the next measure, the stressed syllable “rake” is intoned in D3, the third of Bb<sup>7</sup> and descends chromatically to Db3, the seventh of the Eb<sup>9</sup> harmony to which it corresponds. A similar chromatically embellished descent unfolds over the course of the next two lines, where the word “solitary” is intoned at first with an accented upper neighbor, Gb3 that immediately falls to and rests on F3 over a Bb<sup>7</sup> harmony. The word “sailor,” intoned on E3 and Eb3, acts as the fifth of the A<sup>7</sup> harmony, which resolves appropriately to the root of the D<sup>7</sup> harmony as the stressed syllable “facts” is intoned on

---

to view this, nor the various reductions that follow in this chapter, as true Schenkerian graphs, as I adopt only the representational style of a Schenkerian approach. Stemmed notes indicate vocal frequencies stable and consistent enough to elicit the perception of a musical pitch, and beamed notes indicate those pitches that one perceives as related chord-tones of the underlying harmony and serve as significant tones in an overarching melodic shape. Slurs express the relationship between chord-tones and neighboring non-chord tones, as well as to indicate moments in which the vocal inflections arpeggiate multiple stable chord-tones over time.

D3. The syllable “life” descends to C3, becoming the seventh of the underlying harmony and, on the downbeat of m. 8 resolves to B $\flat$ 2, the third of G minor. This process of vocal descent, spoken in tune to the musical setting, unfolds for nearly every verse of the song.

**Figure 2.19 Spectrogram and harmonic reduction of verse two, "The Ghosts of Saturday Night" by Tom Waits (0:36–0:56).**

The figure illustrates the vocal line and its harmonic support for the second verse of "The Ghosts of Saturday Night" by Tom Waits. It is divided into three sections:

- Top Section:** A spectrogram showing the vocal line's frequency content over time. Below it is a musical staff with a treble clef and a key signature of two flats (B-flat major/D minor). The lyrics are: "paws his in side peacoat pocket for a welcome twenty-five cents and the last bent butt".
- Middle Section:** A second spectrogram and musical staff continuing the vocal line. The lyrics are: "from a package of Kents as he dreams of a waitress with Maxwell House eyes marmalade thighs and scram bled yel low hair".
- Bottom Section:** A piano accompaniment in a grand staff (treble and bass clefs). The bass line features a series of chords: B $\flat$ 7, E $\flat$ 9, B $\flat$ 7, E $\flat$ 9, B $\flat$ 7, A7, D7, gm, 4/2, and E $\flat$ 9.

The second verse, though not identical to the first, reveals a similar melodic structure, and similar voice leading tendencies over analogous harmonic progressions, particularly in its second half (Figure 2.19). The first line of the utterance, “paws his inside peacoat pocket” prolongs an F. The word “peacoat,” begins with an accented upper neighbour G<sup>3</sup> on the stressed syllable “pea” and descends to chord-tone F<sup>3</sup> on “-coat” and remains on F<sup>3</sup> for the stressed syllable of “pocket.” With chromatic embellishments, the utterance “for a welcome twenty-five cents” continues to prolong F, now harmonized by E<sup>b</sup><sup>9</sup> and transferred down the octave to F<sup>2</sup>. The subsequent utterance, “and the last bent butt from a package of Kents as he dreams of a waitress with Maxwell House eyes,” recaptures the original F<sup>3</sup>, and closely parallels the melodic structure of the analogous measures of the previous verse. The F<sup>3</sup> of “last bent butt” functions as the fifth of B<sup>b</sup><sup>7</sup>, descends chromatically through E<sup>3</sup> (the fifth of A<sup>7</sup>) to E<sup>b</sup><sup>3</sup> (the flattened fifth of the same harmony) to D<sup>3</sup> (the root of D<sup>7</sup>) then C<sup>3</sup> (the chordal seventh) and, finally, resolves to B<sup>b</sup> (the third of g minor). Whereas the final measure of the first verse emphasized the chord tones of g minor at its close, the intonation of the second verse moves with the changing harmonies as Waits arpeggiates through the upper chord tones of an E<sup>b</sup><sup>9</sup> chord: the initial stressed syllable of “marmalade” sounds as D<sup>b</sup><sup>3</sup> (the chordal seventh), the initial stressed syllable of “scrambled” as F<sup>3</sup> (the chordal ninth), and the stressed syllable “hair” as B<sup>b</sup><sup>7</sup> (the chordal fifth).

The first two verses of “The Ghosts of Saturday Night” seem to exist right on the boundary between speech and song. Waits is not technically singing, but rather speaking in tune with the music. Such a distinction, if presently ambiguous in the mind of the

listener, becomes explicit with the third verse, which Waits splits into two parts. The third verse is literally interrupted, and the melodic profile of Waits's intonation is similarly cut short (Figure 2.20).

**Figure 2.20 Spectrogram and melodic/harmonic reduction for “The Ghosts of Saturday Night,” verse 3A (0:57–1:09).**

The figure displays three musical components for the specified time range:

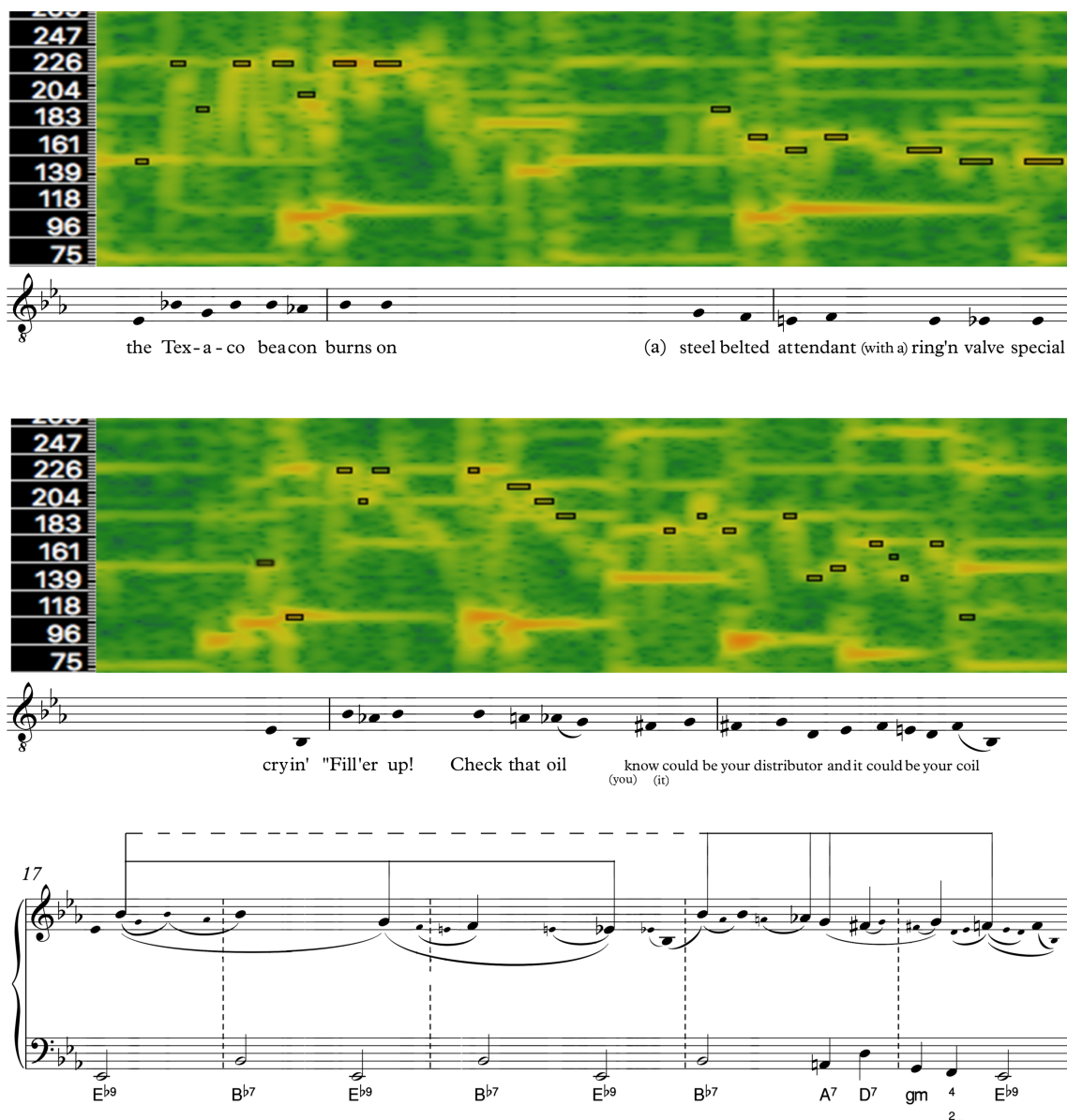
- Spectrogram (Top):** A spectrogram showing the frequency spectrum of the audio. The x-axis represents time, and the y-axis represents frequency. The spectrogram shows a clear melodic line with several distinct horizontal bands of energy, indicating the pitch of the notes.
- Melodic Line (Middle):** A single-staff musical notation in treble clef, showing the melody. The notes are: G4 (quarter), A4 (quarter), B4 (quarter), C5 (quarter), B4 (quarter), A4 (quarter), G4 (quarter), F4 (quarter), E4 (quarter), D4 (quarter), C4 (quarter). The lyrics are: her rhine - stone studded mo ni ker says I - rene.
- Piano Accompaniment (Bottom):** A two-staff musical notation in treble and bass clefs. The melody is in the treble clef, and the bass line is in the bass clef. The chords are: B<sup>b</sup>7, E<sup>b</sup>9, B<sup>b</sup>7, E<sup>b</sup>9, B<sup>b</sup>7, A<sup>7</sup>, D<sup>7</sup>, gm /F, E<sup>b</sup>9.

With chromatic embellishments, the utterance “her rhinestone studded moniker says Irene” emphasizes F3, while the utterance “as she wipes the wisps of” presents a substitution of B $\flat$ 2 for the expected E $\flat$ 3 (shown in brackets). The descent of the line continues through D3, sounded on the strong syllable of the word “dishwater” and rests on C3 with the final words “blonde from her eyes.”

The arrival on C3 at the end of the line is interpreted as the seventh of a D<sup>7</sup> harmony, left unresolved in the melody as Waits’s text recitation abruptly halts, but the harmonic progression continues on. At the outset of the third line of the third verse, which coincides with a new musical phrase, Waits not only resolves the lingering C3 abandoned from the previous line, but does so by, for the first and only real time, singing in earnest (Figure 2.21).

Whereas in the previous verses Waits’s pitch range stays within the bounds of naturally spoken speech typical for adult males (85–180 Hz), in opening utterance of this verse, “and the Texaco beacon burns on,” he reaches above that range to emphasize—indeed, to sing— B $\flat$ 3 (233 Hz). The overt singing is short-lived, however, as the very next utterance “a steel-belted attendant with a Ring n’ Valve Special” immediately reverts back to speech, spoken in tune with the music. Over the first three measures of the phrase Waits outlines an E $\flat$  major triad in arpeggiated descent from B $\flat$ 3 (the fifth of E $\flat$ <sup>9</sup>) to G3 (the third of E $\flat$ <sup>9</sup>) and, passing through F3 (the fifth of B $\flat$ <sup>7</sup>), arrives on E $\flat$ 3 (supported by E $\flat$ <sup>9</sup>).

Figure 2.21 Spectrogram and melodic/harmonic reduction for verse 3B from "Ghosts of Saturday Night" by Tom Waits (1:20–1:39).



With the subsequent utterance, “cryin’ fill’er up, check that oil,” Waits recaptures the  $B^b_3$  over a  $B^b_7$  chord, descends stepwise to the  $A^b_3$  (the seventh of the same harmony) and resolves the seventh to  $G_3$  (the seventh of  $A^7$ ). In turn, the following utterance, “you know it could be your distributor” continues to exhibit standard voice-leading



procedures: G3 resolves down by step to F#3 (the third of D<sup>7</sup>) which, acting as temporary leading-tone, resolves up to G3 (now the root of G minor). The remaining utterance of the verse, “or it could be your coil” descends by chromatically embellished step-wise motion to F3 (the ninth of E<sup>b</sup><sup>9</sup>) before resting back on B<sup>b</sup>2.

By now the reader will have begun to recognize the melodic spectre that lies beneath Waits’s speech and, as is my hope with this entire enterprise, begun to listen for that melody as it relates to its musical setting. The remaining three verses of the song continue in the same vein, with the fourth verse lingering in the upper range introduced by the previous verse and the fifth and sixth verses returning back to the vocal range of the opening phrases (Figure 2.22). Readers should, as they consider this example, keep in mind that despite the apparent similarities to a Schenkerian sketch, that is not the nature of the figure.

The ostinato harmonic progression that governs the song is, despite the ubiquity of A<sup>b</sup>, in the key of B<sup>b</sup> and, as such, lacks any structural dominants. In this way, to suggest a large-scale fundamental structure is somewhat tricky and illogical (at least from a Schenkerian perspective). Rather, the example above is meant to be a description of the melodic structure that underlies Waits’s speech in “The Ghosts of Saturday Night.” A larger melodic form can be heard over the ostinato progression—one that begins by emphasizing a stepwise descent from F3 supported by dominant harmony to B<sup>b</sup>2 supported by tonic harmony. The descending melody is twice successful at reaching the tonic goal, but on the third descent stops short on C3, the seventh of D<sup>7</sup>, which resolves only after some time and through reaching over to B<sup>b</sup>3.

**Figure 2.22 Full melodic/harmonic reduction of “Ghosts of Saturday Night,” studio recording.**

The musical score is presented in three systems, each with a grand staff (treble and bass clefs) and a key signature of two flats (B-flat major or D-flat minor). The first system covers Verses 1, 2, and 3a, starting at measure 5 and ending with a double bar line. The second system covers Verse 3b and Verse 4, starting at measure 18. The third system covers Verse 5 and an Instrumental Interlude, starting at measure 29. Vertical dashed lines connect the melodic lines in the treble and bass staves across measures, indicating harmonic relationships. The score concludes with a final double bar line.

Over verses 3b and 4, B $\flat$ 3 descends back to F3, but the F3 is somewhat unstable as it functions as the ninth of an E $\flat$ <sup>9</sup> harmony, arpeggiated by the harmonic progression at the end of verse 4. Finally, after a four-measure instrumental interlude, verse 5 reiterates the original melodic descent from F3 to B $\flat$ 2, once over alternating B $\flat$ <sup>7</sup> and E $\flat$ <sup>9</sup> harmonies and again over a harmonic progression that arpeggiates E $\flat$ <sup>9</sup> and arrives to close on B $\flat$ .

## 2.5 Speaking Live

If the reader has taken the time not just to consider in abstract the proposal of a melodic structure imbedded in Waits's speech in "The Ghosts of Saturday Night," but also to listen carefully to Waits's recording of the song, then the idea that Waits is speaking in tune should (hopefully) resonate. The question may linger, however, as to whether the melody presented for "The Ghosts of Saturday Night" is one that is truly integral to the song or one that exists only in the specific instance of this studio recording. While on tour to promote the album, *The Heart of Saturday Night*, on which the song in question was first released, Waits performed "The Ghosts of Saturday Night" live at Ebbets Field, Colorado on October 8, 1974 and his performance was recorded.<sup>78</sup>

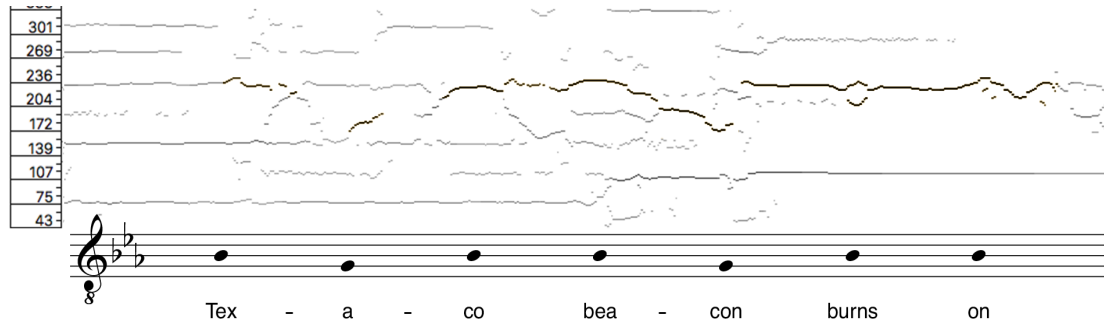
As with many live performances, not every detail of the song is the same as the studio version, but a comparison of the two reveals that far more details than one might expect are consistent, suggesting that when Waits speaks the words of "The Ghosts of Saturday Night," he is indeed conjuring up a real, albeit subtle, melody. There are some obvious moments where Waits's live performance of "The Ghosts of Saturday Night" bears close resemblance to the studio version. In both recordings, for example, Waits sings outright the utterance "Texaco beacon burns on" with the same rhythmic and pitch inflections (Figure 2.23).

---

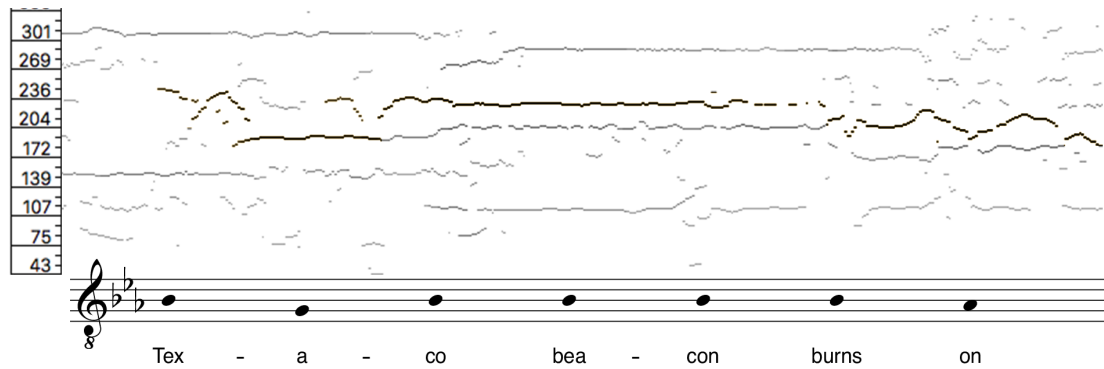
<sup>78</sup> A recording of this performance was made available to the public in a 2001 limited release. Tom Waits, *Dime Store Novels Vol. 1: Live from Ebbets Field (1974)* (NMC Music Ltd 8 00945 00822 8, 2001). Compact disc.

**Figure 2.23 Melodic similarities between line 11 from live and studio recordings of “The Ghosts of Saturday Night” by Tom Waits.**

a) “The Ghosts of Saturday Night,” studio recording released October 15, 1974.



b) “The Ghosts of Saturday Night,” live at Ebbots Field, Colorado, October 8, 1974.

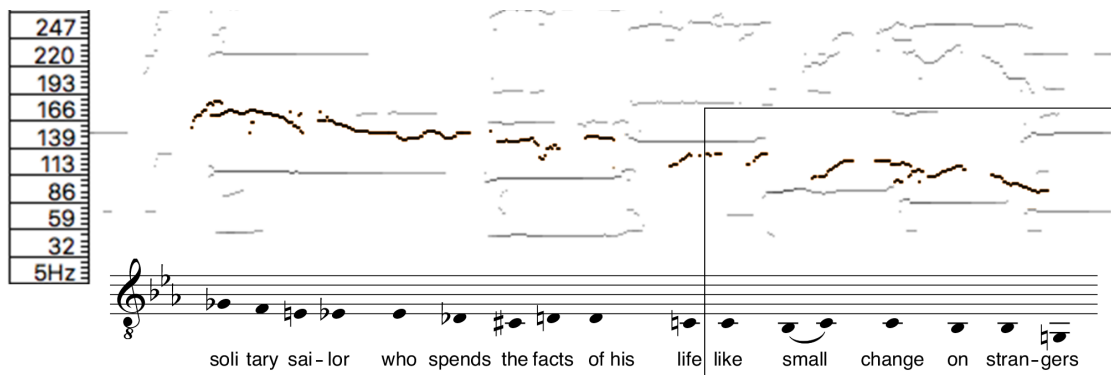


Each spectrogram highlights only the peak frequencies of Waits’s vocal inflections. Both live and in studio, Waits performs “Texaco beacon burns on” to the same melody with near-identical pitch inflections, emphasizing Bb3 and only diverging on one or two occasions. In addition to this shared moment in Waits’s two versions of the work, there are several more subtle passages that reveal an underlying rhythmic and melodic identity to Waits’s speech, particularly at end points in phrases. The first verse, for example, is a compelling example of the similarities between the studio version as compared to the live recording. Among the similarities between the way that Waits recites the first two lines of this verse is their metric placement, as both spoken phrases begin and end in similar

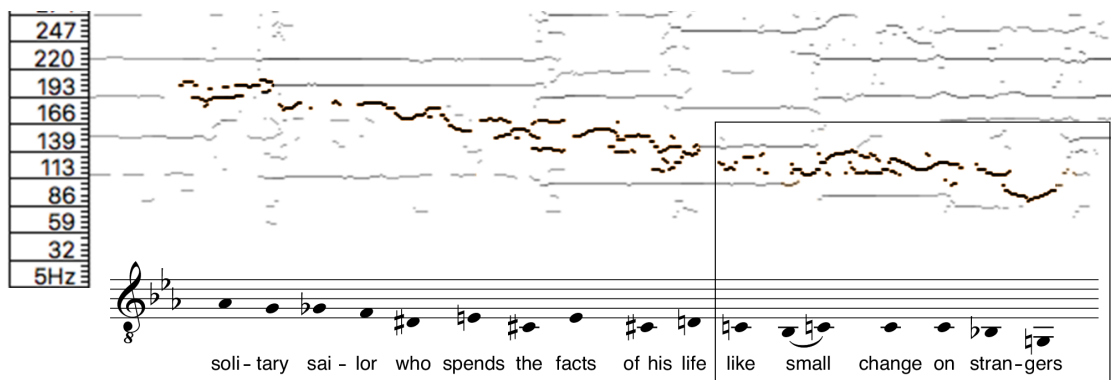
metric positions in relation to the musical setting of the narrative. Similar too are their melodic contours and pitch inflections (Figure 2.24).

**Figure 2.24 Melodic similarities between line 2 from live and studio recordings of “The Ghosts of Saturday Night” by Tom Waits.**

a) “The Ghosts of Saturday Night,” studio recording released October 15, 1974.



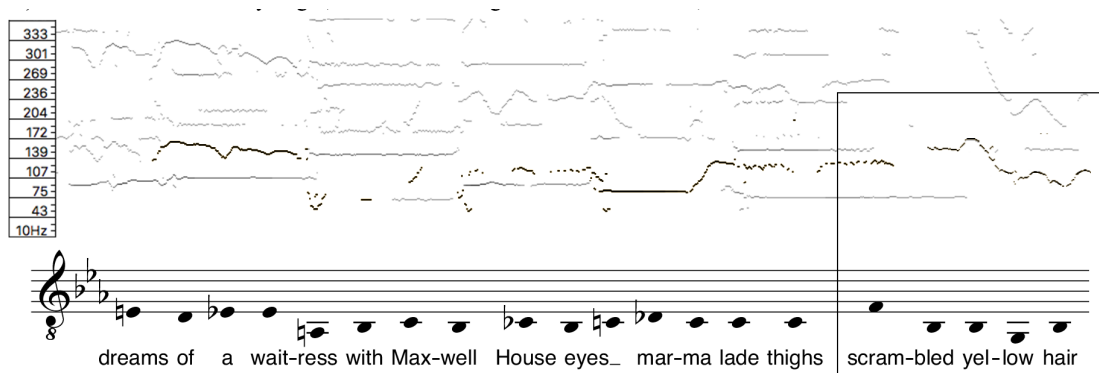
b) “The Ghosts of Saturday Night,” live at Ebbots Field, Colorado, October 8, 1974.



When Waits recites the phrase, “like small change on strangers,” the specific pitches he uses are remarkably similar. Save for the syllable “on,” which is intoned as a B $\flat$  in the studio version and as a C in the live version, every other syllable is recited on the same pitch. Similarly, the last phrase of the second verse is recited much the same way in studio as Waits recites it live (Figure 2.25).

**Figure 2.25 Melodic similarities between lines 7–8 from live and studio recordings of “The Ghosts of Saturday Night” by Tom Waits.**

a) “The Ghosts of Saturday Night,” studio recording released October 15, 1974.



b) “The Ghosts of Saturday Night,” live at Ebbots Field, Colorado, October 8, 1974.

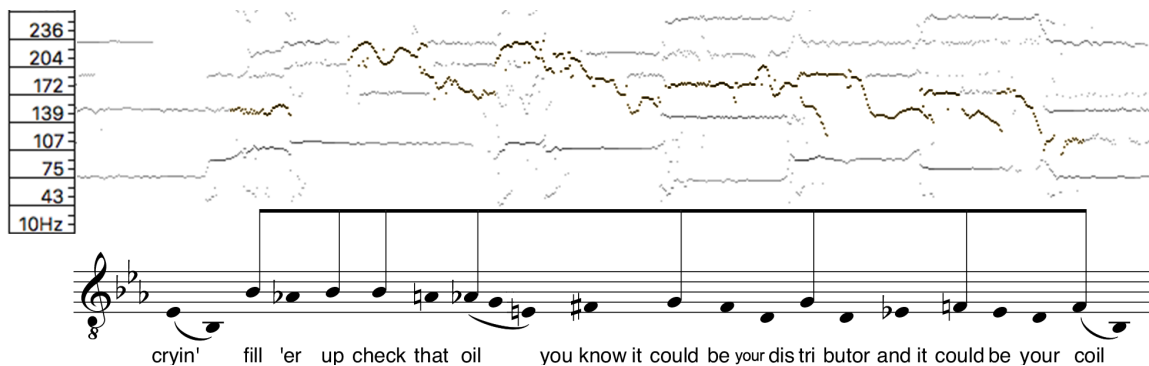


These short examples of melodic similarities between Waits’s recitation of the text in studio and in live performance might, on their own, suggest that any melodic similarities can be explained either by the fact that Waits is, at times, singing outright, or that the endings of vocal phrases naturally coalesce around the same region in his voice. But there are longer passages that would suggest otherwise—full phrases in which Waits’s recitation in live performance parallels the melodic structure of the studio recording.

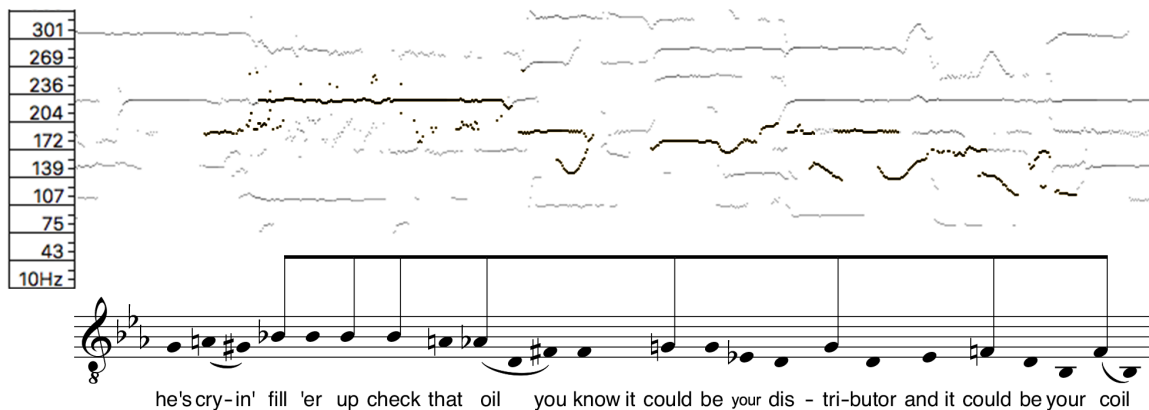
In his delivery of the line, “fill ‘er up, check that oil; you know it could be your distributor and it could be your coil,” the melodic similarities between live and studio versions are the most impressive (Figure 2.26). Although not entirely identical, the two versions express the same descending line from B $\flat$ 3 to F3, following the same chromatic descent from B $\flat$ 3 to A $\flat$ 3 over “fill ‘er up, check that oil.” The latter half of the utterance, “you know it could be your distributor and it could be your coil,” involves identical melodic contours and similar pitch inflections on prominent syllables.

**Figure 2.26** Longer excerpt comparing melodic inflections in lines 12-13 from live and studio recordings of “The Ghosts of Saturday Night” by Tom Waits.

a) “The Ghosts of Saturday Night,” studio recording released October 15, 1974.



b) “The Ghosts of Saturday Night,” live at Ebbots Field, Colorado, October 8, 1974.



In both versions, for example, the first syllable “could” is preceded by a lower neighbour tone (chromatically raised in the studio version) and recited on G3. The word “distributor” in both versions emphasizes G3 and D3, and the second syllable “could” arrives on F3. Finally, both versions employ the same leap down a fifth from F3 to B♭2 on the word “coil” to conclude the phrase.

## 2.6 Conclusions

The analyses presented in this chapter show how Waits’s speech rhythm employs familiar musical devices such as word painting, rhythmic imitation, and attention to stress and beat alignment between speech and musical meter. Similarly, the close study of Waits’s speech intonation reveals how musically informed Waits’s recitations are, as far too often to be coincidental, extensive passages of prose are not only intoned to frequencies consonant with the underlying musical harmonies but also follow familiar musical schema in their voice leading and melodic patterns. Waits is consistently musical in his manner of speech and his ample repertoire of spoken-word songs are testaments to the various ways that he treats vocabulary as an instrument, as well as the strategic ways in which he makes music out of that instrument. This is not to suggest, however, that the musical properties of Waits’s spoken-word songs are intentionally determined by the songwriter in the sense that one would imagine him meaning to hit just this note on just this syllable at just this moment in time. But there is a degree of intentionality to which I mean to draw attention, namely, that of speaking musically. Whether he is screaming into a megaphone, banging on a conundrum, beatboxing nonsense syllables, or telling a story, Tom Waits is a consummate musician. He speaks often of a belief that all things are



musical objects, so it is no surprise that he makes words, whether sung or simply spoken in time and in tune, so impressively musical.

Going beyond Waits's artistry, the analyses presented here also reveal the potential that spoken-word song as a genre has for further explorations of the relationship between speech and song. As neuroscientists and cognitive psychologists continue to explore this relationship, many seek to find and understand some explanation for the boundaries between speech and song. We know, for example that one exists in the speech-to-song illusion, as Tierney et al. write: "It is normally obvious to listeners whether a human vocalization is intended to be heard as speech or song. However, the two signals are remarkably similar acoustically. A naturally occurring boundary case between speech and song has been discovered where a spoken phrase sounds as if it were sung when isolated and repeated."<sup>79</sup> In this case the boundary is determined primarily by isolation and repetition.

Tom Waits's spoken-word songs suggest that there may be another naturally occurring boundary case between speech and song which reveals itself when speech is presented in combination with music. The influence of a musical setting on the perception of speech rhythm and intonation has significant implications for a listener's sense of rhythmic regularity, event alignment, and frequency perception in speech.

While the analyses presented here are of my own listening experience, it is probable that the aural illusions I describe are shared by other listeners. One direction for research in

---

<sup>79</sup> Tierney et al., "Speech versus Song: Multiple Pitch-Sensitive Areas Revealed by a Naturally Occurring Musical Illusion," 1.

this area is to examine how listeners perceive speech acts that are not only supported by musical settings but strategically spoken so as to be musical themselves. It would be particularly interesting to study these relations in spoken-word songs by artists who identify as musicians (such as Tom Waits) versus artists who identify as poets (such as, for example Jack Kerouac or Allen Ginsberg) to see if there is a greater flexibility and/or precision in the manner in which the speech-music interactions are handled.

Going beyond the research potential of studying spoken-word songs, my analyses raise a larger question about the extrinsic dimensions that influence how we perceive speech and song. In studying the way in which Waits's spoken-word songs can be thought of as genuinely song, rather than just as musically accompanied speech, I argue that Spoken Word not only can be thought of as musical but is *better* thought of as music. So many of the experiments designed to investigate relationship between speech and song invariably operationalize an *a priori* assumption that speech and song are somehow in their purest forms at odds with each other and thus reinforce the notion that the perceptual experience of listening to speech differs in phenomenal character from that of listening to song.

The notion that how we think about speech versus song significantly influences how we perceive speech in relation to song leads to larger implications for how we perceive Spoken Word generally in relation to music. Such a notion requires that we reconsider what assumptions underly our concept of speech and music categories, and how those assumptions might influence and potentially limit further analysis. If, at the outset, Spoken Word is considered categorically distinct from music, then our experience of Spoken Word versus music will necessarily be different. But if, at the outset, we shed such categorical predispositions, then we open our research (and ourselves) to the

possibility that a great many more experiences in our lives can be far more musical than we may have imagined.

In the next chapter, “The Poet Performs: Marc Kelly Smith and Slam Poetry,” I venture deeper into the realm of pure spoken word, investigating how a listener might hear or imagine musical qualities in performed speech in the absence of a musical setting. While the current chapter’s analytical methods emphasized the subtleties of durational timing and pitch frequency, the next chapter changes methodological gears to investigate the musical elements of slam poetry. As a genre invariably comprised of works which fall into the nebulous poetic void traditionally called free verse, the following chapter considers how theories of form developed specifically for understanding formal structure in western European art music might productively apply to slam poetry and indeed, modern poetry in general that is similarly rooted in the concept of performance.

## Chapter 3

### 3 The Poet Performs

If we can conceive of music as performance—indeed, if we tend to agree with director and drama theorist Richard Schechner that “just about anything can be studied ‘as’ performance,”—then it follows that just about anything studied as performance might also be studied for its musicality.<sup>80</sup> Of course, some types of performances are more conducive to musical study than others. Where the previous chapter was concerned with uncovering the musicality of spoken-word song, this chapter is concerned with yet another such type of spoken-word performance activity: slam poetry.

In this chapter, I explore how theories of form developed for the purposes of understanding musical form can apply to and shed light on slam poetry, a spoken-word genre that does not involve an explicit musical setting like Tom Waits’s spoken-word songs in the previous chapter. Through analyses of Marc Kelly Smith’s 2011 slam poems “Kiss It” and “I Wear My Father’s Coat,” I demonstrate how re-envisioning the emergent design of a slam performance as a musical form can ultimately communicate more about the formal structure that a textual representation of a slam poem might suggest.<sup>81</sup>

---

<sup>80</sup> Richard Schechner, *Performance Theory, Revised and Expanded Edition* (London; New York: Routledge, 2006), 38.

<sup>81</sup> Marc Kelly Smith hosted the original poetry slam at the *Get Me High Lounge* in the Bucktown neighbourhood of Chicago in 1986. In 1987, it became a weekly event called the Uptown Poetry Slam, held at the Green Mill Lounge in Chicago, Illinois. The Uptown Poetry Slam still runs weekly at the Green Mill Lounge, and Smith is still the host. Smith’s founding role in the slam poetry movement is documented in multiple sources, including Marc Kelly Smith, *Take the Mic: The Art of Performance Poetry, Slam and the Spoken Word* (Naperville: Sourcebooks MediaFusion, 2009), Cara Losier Chanoine,

### 3.1 Slam Poetry and/as Music

A “slam” is a poetry competition where poets perform original works alone or in teams before an audience that judges the work as much for its performance as for its content, style, or form. Established in 1986–87 at Chicago’s Green Mill Jazz club, the competition quickly spread across the country and, being an important movement in poetry since the early 1990s, slam has restored interest in poetry in performance. The rise of poetry slams across America picked up on the performative aspects of earlier Beat and Negritude movements and for young poets of diverse backgrounds, slam poetry has been seen as a democratizing force and is often highly politicized, drawing upon racial, economic, and gender injustices as well as current events for subject matter.<sup>82</sup>

Venerated slam poet and founder of *Project Voice*, Sarah Kay, describes slam poetry as “creative poetry that doesn’t just want to sit on paper — that something about it demands that it be heard out loud.”<sup>83</sup> Like many slam poets, Kay has several published books of poetry and regularly contributes to the large and growing body of slam poetry in print. This hardly means that Kay disavows text representations of her spoken-word poetry. Rather she is touching upon an essential aspect of Spoken Word that breaks the bounds of the page and the written medium. The sonic events of a slam performance (rhyme, alliteration, rhythm, intonation, dynamics, articulation, etc.) are fundamental

---

*Poetry, Points, and Performance: Expanding the Scope of Slam Poetry Analysis* (PhD dissertation, Indiana University of Pennsylvania, 2018), and Price-Styles, “MC Origins: Rap and Spoken Word Poetry.”

<sup>82</sup> Smith, *Take the Mic: The Art of Performance Poetry, Slam and the Spoken Word*.

<sup>83</sup> Sarah Kay, “[If I Should Have a Daughter](#),” *Ted Talks* (March 3, 2011 in Long Beach, CA.), 18:14, 2011.

elements of the work and crucial to its structural design. In this way, slam poetry has much in common with music.

There are many resemblances between slam poetry and western art music traditions. For starters, there is the tradition that a piece of music exists, at least in part, on a page—in some written form which represents the latent or dormant possibilities of its sonic manifestation. The same tradition can be found in slam poetry. The difference, however, is most readily apparent through the language used to describe the written form of a musical work versus that of a work of slam poetry. In musical discourses, the written form a work, though sometimes referred to as “the music,” also has a distinct name: *the score*.

However subtle, this linguistic distinction enables music scholars to recognize the multiplicity of contexts that make up our concept of a musical work. In poetic discourses, this distinction is less explicit. The pages of symbols that represent a work of poetry are typically called “the poem,” as though its existence on the page can be understood as the work fully and completely. While there may be poetic traditions where this is true (and I suspect far less than we might imagine), slam poetry is not one of them. In the realm of slam poetry, where the performance of a work is essential to the ontology of the work, it is imperative that the written form of the work be understood not as “the poem” but rather more as a musical score.

Slam poetry, according to comparative oral traditions scholar John Miles Foley, is a perfect example of the oral media category he calls *voiced texts*, the kind of oral poetry that “begins life as a written composition only to modulate to oral performance before a

live audience.”<sup>84</sup> Foley distinguishes between oral traditions by differentiating between the modes of composition (written or oral), performance (written or oral), and reception (written or aural) of a work. Slam poetry represents voiced texts because they are composed in a written form, performed orally, and received aurally. Any study of the written form of a voiced text should be undertaken with its oral/aural forms at the forefront of the analyst’s mind. With respect to western European art music traditions, the very same modalities are at play: musical works of this tradition are composed in a written form, performed orally, and received aurally.

Kay’s description of slam poetry, “that something about it demands that it be heard out loud,” echoes certain aspects of musicologist Nicholas Cook’s idea that there is something about music which demands it be understood not as contained within a score but beyond it, comprising all of its temporal, performative, social, and sonic facets. This is the central claim of his 2002 book, *Beyond the Score*—that to think deeply about music requires thinking differently about the nature of a musical object and only by thinking of music as performance can we start to make sense of musical scores.<sup>85</sup> Cook does not suggest that musical scores are insignificant or irrelevant altogether; rather he argues that “the score” is not synonymous with “the music.”

---

<sup>84</sup> Foley, *How to Read an Oral Poem*, 43. Not to be confused with Nicholas Cook’s notion of “voiced texts,” by which he means to critique the text-centrism of traditional music-analytical methods that ignore or marginalize performance. See Nicholas Cook, *Beyond the Score: Music as Performance* (Oxford: Oxford University Press, 2013), 17–23.

<sup>85</sup> Cook, *Beyond the Score: Music as Performance*, 1.

On its own, a score is an insufficient account of a musical work and that performance is an essential part of the musical experience. A very similar perspective towards practices in poetic circles led to the formation of slam poetry. Slam-founder Marc Kelly Smith describes slam's primary distinction within the realm of poetry in general as the merging of the art of performance with the art of writing poetry.<sup>86</sup> In other words, the poem is not synonymous with the page.

In his telling of the slam origin story, Smith writes that “the seeds of slam were sown in the womb of Get Me High Jazz Club in Chicago when (way back in 1984) yours truly upset the status quo by daring to enthusiastically embrace performance as an essential part of the poetic experience.”<sup>87</sup> Alongside the other members of the Chicago Poetry Ensemble, Smith had become disillusioned with the state of and general attitude toward poetry in America. By the 1980s, the feeling was that poetry had been firmly shackled to the page and those few lingering traditional poetry events “had diminished to sporadic, self-absorbed, non-adventures cramped uncomfortably in bookstore aisles and attended by a handful of insular followers.”<sup>88</sup>

Slam poetry emerged in response to these critiques of poetic practices: it was a reaction to poets who had lost sight of their audience; to critics whose intellectual elitism rendered poetry accessible to fewer and fewer people; to the false superiority of the poet over the audience. Poetry is a communicative art and without an audience it loses its punch. In

---

<sup>86</sup> Smith, *Take the Mic: The Art of Performance Poetry, Slam and the Spoken Word*, 5.

<sup>87</sup> Smith, 6.

<sup>88</sup> Smith, 21.



slam poetry the performer is no celebrity; the performer is the audience's servant. In fact, the performance imperative of slam poetry has considerably more to do with the audience than it does with the performer.

Audience participation is an essential element for a successful slam. Only the audience determines whether the sounds coming out of a person's mouth are a poem and whether that mouth belongs to a poet—that is to say, whether something has been communicated to them. Slam communities take performer-audience relationship very seriously, which is another aspect of slam that is shared by Cook's argument that "as enacted in the practice of performance, the musical work is an irreducibly social construct."<sup>89</sup> Cook elaborates on this point in his 2018 book, *Music as a Creative Practice*, and particularly in the book's second chapter, where he writes that "performers and listeners collaborate in creative practice," and that "to engage with music is in some sense to be part of a community, whether in the literal circumstances of social listening or in some more symbolic form."<sup>90</sup>

Not only is the listener (or audience) an essential part of the art form itself, but so too is the sense of community that develops when the role of the audience is taken seriously.

The same is true for slam poetry, as Smith explains:

Slam reaches far beyond those in the spotlight, indeed beyond the walls of any particular venue. It encompasses all the forces involved in staging a show, including ... the audience itself. More important than any individual performer or event, slam is a community of people who have

---

<sup>89</sup> Cook, *Beyond the Score: Music as Performance*, 239.

<sup>90</sup> Nicholas Cook, *Music as a Creative Practice* (Oxford: Oxford Scholarship Online, 2018), 22; 20.

discovered a dynamic way of presenting poetry aloud onstage in full public view, enabling its passion, wisdom, and beauty to be experienced with total impact.<sup>91</sup>

Despite all that slam poetry has in common with music, it is not typically thought of *as* music. Studies of slam poetry traditionally belong to the domains of literary or cultural studies, dominated by discussions of its significance as a cultural movement, or its ambiguous literary status. While scholars of slam poetry ardently contend that the work *is* its performance, few offer insights on how to integrate the sound structures realized through performance into analyses of these works. Literary scholar Walter Bernhart points out one reason why this might be:

What is missing in any written text of poems is a graphic representation of their prosodic dimension, which is largely irrelevant for a purely semantic appreciation of the poems, but essential for their oral articulation. The elements of stress, intonation, rhythm (i.e., time segmentation), timbre, sound volume and tempo all those, in common parlance cannot be identified in the written document, and yet they absolutely determine the oral performance. Even the most barren musical score indicates at least pitches and durations, not to speak of elaborate post-Debussy scores with their minute specifications of dynamics, tempo, tone colour and so on. But a poem on a page has none of all this.<sup>92</sup>

Bernhart brings to light one of the problems that arises in the study of Spoken Word from a literary perspective, namely, that the tools for analytical inquiry developed in fields like literary criticism and literary studies are primarily intended for written texts, and as such,

---

<sup>91</sup> Smith, *Take the Mic: The Art of Performance Poetry, Slam and the Spoken Word*, 25.

<sup>92</sup> Walter Bernhart, "Rhythmical Ambivalence of Poetry Performance: The Case of Elizabethan Verse," in *Word and Music Studies: Essays on Performativity and on Surveying the Field*, ed. Walter Bernhart and Michael Halliwell (Amsterdam: Rodopi, 2011), 89.

are insufficient for the study of performance poetry such as the slam. Although Bernhart is not specifically thinking of slam poetry when he covets the nuances of musical scores, his suggestion that voiced texts require a more musically informed analytical perspective certainly motivates the kind of analytical approaches towards which scholars of slam poetry are increasingly gravitating.

### 3.2 The Ethnopoetic Approach

The dominant analytical approach to slam poetry at present is an ethnopoetic one, a mode of analysis that seeks to improve the accuracy of a slam “score” and enable faithful reperformances of a slam event. The ethnopoetic approach, proposed by literary scholars Dennis Tedlock and Dell Hymes in the late 1980s and further developed by John Miles Foley in the early 2000s, is essentially a way of analyzing oral poetry through the construction of an annotated score by developing a notational system that can account for the paralinguistic and non-linguistic elements of the slam event.<sup>93</sup> According to Foley, ethnopoetics has “two actions on its agenda: to accurately *represent* oral poetry on its own terms and by doing so to foster the reader’s more faithful *reperformance*.”<sup>94</sup> These two actions represent a balancing act for the ethnopoetic analyst, as increasingly accurate precision in transcription runs the risk of becoming overwhelming to the reader but an imprecise transcription will risk the fidelity of reperformance.

---

<sup>93</sup> Dennis Tedlock, *The Spoken Word and the Work of Interpretation*, *The Journal of American Folklore* (Philadelphia: University of Pennsylvania Press, 1985); Dell Hymes, *Narrative Inequality: Toward an Understanding of Voice* (London: Taylor & Francis, 1996).

<sup>94</sup> Foley, *How to Read an Oral Poem*, 96.

In her recent work on slam poetry, scholar Cara Losier Chanoine makes a strong case for the ethnopoetic approach as particularly well-suited to slam poetry analysis, in that it reveals the ways in which text and performance function symbiotically within the slam poetry genre. She advocates for a perspective towards slam poetry that affords both text and performance equal places in the constitution of the work and develops an ethnopoetic legend that is surprisingly communicative, specific enough to elicit a reperformance of the work in one's mind.<sup>95</sup> One of Chanoine's primary foci is comparative analysis, examining what she terms performative *mouvance* between multiple versions—both text and performance—of a single slam poem, “Like” by Mike McGee. Another approach related to ethnopoetics is Julia Novak's 2011 work *Live Poetry: An Integrated Approach to Poetry in Performance*.<sup>96</sup> Novak adopts principles from musical notation to graphically portray pitch, rhythmic, and dynamic processes which shape a live poetry performance. Novak's adaptation of musical notation for the purposes of live poetry is a decisive step towards developing a methodology for examining literary works that exist first and foremost as performances.

With all the advances and innovative refinements to ethnopoetics, however, a number of caveats remain. First, in pursuing the development of a notational system that captures the paralinguistic complexities of live events, the entire enterprise of ethnopoetics can

---

<sup>95</sup> Chanoine, *Poetry, Points, and Performance: Expanding the Scope of Slam Poetry Analysis*.

<sup>96</sup> Julia Novak, *Live Poetry: An Integrated Approach to Poetry in Performance* (Amsterdam: Rodopi, 2011). Novak employs a proto-notation in her analysis of Jackie Hagan's 2007 performance of “Coffee and Tea.” Her analysis identifies recurrent rhythmic patterns in Hagan's poetry, as well as recurrent melodic contours that work together to create a most musical performance.

inadvertently suggest that the nuance and detail of performance poetry will submit to analysis only after it is bound to the page and captured in written form. In other words, the ethnopoetic approach can appear to implicitly accept and perpetuate a kind of text-centrism that slam poetry intentionally eschews. Second, the more faithfully one represents the myriad paralinguistic facets of oral poetry, the greater the risk of information overload and in turn, the greater the experiential difference between attending a live performance event and reperforming that event through an ethnopoetic transcription. Of course, this is not entirely a bad thing: the difference shines a giant floodlight on the vast amount of non-linguistic communication that goes into performance. But it does seem to mean that the more accurately a performance is represented in transcription, the less capable the ethnopoetic transcription will be at capturing any sense of temporality.

Ethnopoetic transcriptions of a slam performance (which would typically be under 3 minutes and 10 seconds long, as per the rules of slam competitions) can be several pages long and take considerably longer to read and understand than the performance upon which they are based. And because temporality is a fundamental aspect of form, the final caveat to the ethnopoetic approach is its inability to produce any analytical insights with regard to the formal properties of slam poetry. From a formal perspective, slam poetry invariably falls within what literary scholars call *free verse*. Unlike couplets, sonnets, limericks, etc., free verse follows no preconceived, regular pattern of meter, rhyme, or

other traditional poetic techniques.<sup>97</sup> When, however, one defines free verse by what it lacks, then form in such poetry becomes little more than a set of absences and identifying form becomes a process of checking for rhyme and finding no pattern; checking for meter and discovering no regularity. Ultimately, the result suggests, erroneously, that free verse poetry is without any sense of formal structure. If our method for investigating the formal structure of free verse poetry focused less on what traditional techniques are not present and more on how the words, phrases, rhythms, and motives function within such poetry, then we may discover unique formal structures analogous to familiar musical forms.

### 3.3 Reforming Formal Functions

One of the most well-known treatments of musical form to date is William Caplin's theory of formal functions.<sup>98</sup> A formal function describes the specific role played by a particular passage in the formal organization of a musical work. The theory treats form as a hierarchical, temporally situated process in which three general function types—initiating, medial, and concluding—govern the unfolding of events. The idea that a particular passage, or “unit,” has a functional role in establishing the form of a work and that its function is largely determined by its temporal situation is an attractive starting point for an investigation of form in slam poetry for a number of reasons. First, such an approach would enable the analyst to treat the text of a slam poem not so much as a

---

<sup>97</sup> For definitions of couplets, sonnets, limericks, and other poetic forms, see Clare Cavanagh et al., eds., *The Princeton Encyclopedia of Poetry and Poetics*, 4th ed. (Princeton, N.J.: Princeton University Press, 2012). Accessed November 30, 2020.

<sup>98</sup> William E. Caplin, *Classical Form: A Theory of Formal Functions for the Instrumental Music of Haydn, Mozart, and Beethoven* (Oxford: Oxford University Press, 1998).

transcription but as a score—both of a performance and with the potential for performance. Second, the approach would allow the analyst to move away from the traditional approach to poetic forms which treat line breaks, meter, and rhyme as the exclusive and essential formal parameter of poetry.<sup>99</sup> Finally, the approach can reveal unique formal structures in poetry that otherwise would fall into the nebulous formal void that is presently free verse. While the overarching benefits certainly merit an attempt at investigating the formal structure of slam poetry according to the principles of formal functions, the application of Caplin's theory of musical form to a repertoire so far outside the theory's original purview is not without significant challenges. A central challenge is the degree to which concepts of tonality not only inform but very often determine what constitutes a particular formal function. One reason for this is the specificity of subject matter in the formation of Caplin's theory.

The full title of Caplin's 1998 book, *Classical Form: A Theory of Formal Functions for the Instrumental Music of Haydn, Mozart, and Beethoven*—in which he first presents his theory in full to the musical community—explicitly states the theory's highly specialized context. Indeed, Caplin takes pains to be specific about the scope of his theory and by doing so recognizes that functionality can, and often does change depending on context.<sup>100</sup> Because formal functions developed with such a specific and highly contextualized repertoire in mind, any application of them outside the context of that

---

<sup>99</sup> I have yet to find in my research any literary theory that does not either prioritize these parameters or rely on them exclusively.

<sup>100</sup> A frying pan on the stove is a kitchen item; in the hands of a villain it becomes a weapon; and in the hands of an artist, the mind reels at the possibilities.

repertoire will require some measure of adaptation. In the case at hand, this involves teasing out the contextual specificity of Caplin's functional definitions from the essential functional roles each concept plays in establishing the form of a work.

So how does one take the "Classical music" out of *Classical Form* without altering or obscuring the essential principles of formal functions? Jason Yust asks a very similar question in his recent 2018 book, *Organized Time: Rhythm, Tonality, and Form*.<sup>101</sup> Yust recognizes an inherent potential in Caplin's theory of formal functions to "point the way to a more flexible theory of form based on processes and structuring principles rather than fixed schemata."<sup>102</sup> Though for reasons (and repertoire) entirely dissimilar from the present goals, Yust addresses the same challenge that I face in applying formal functions to slam poetry. Yust takes a decisive and innovative step towards a generalized theory of formal functions, capable of application outside the specialized language of the high Viennese school, by disentangling tonal structure from formal structure. In doing so, he arrives at a small but malleable set of form-structural criteria. Yust's approach is not so much a critique of Caplin's theory but an evolution of it, wherein the formal functions Caplin proposes can have powerful applicability in contexts and repertoires hitherto unimaginable.<sup>103</sup>

---

<sup>101</sup> Jason Yust, *Organized Time: Rhythm, Tonality, and Form* (Oxford: Oxford University Press, 2018).

<sup>102</sup> Yust, 59.

<sup>103</sup> This may be an appropriate place to mention a few others whose take on musical form could have easily been adopted here as well, particularly Antares Boyle, "Formation and Process in Repetitive Post Tonal Music" (PhD Dissertation, University of British Columbia, 2018). Also Patricia Howland, "Formal Structures in Post-Tonal Music," *Music Theory Spectrum* 37, no. 1 (2015): 71–97.



Yust's form-structural criteria are based on four key devices: repetition, fragmentation, caesura, and contrast. He uses these devices to re-form Caplin's various formal functions in such a way that relies solely on formal structure, rather than on an integration of formal and tonal principles. A presentation phrase, for example, according to Caplin's theory, means "an initiating intrathematic function consisting of a unit (usually a basic idea) and its repetition, supported by a prolongation of tonic harmony."<sup>104</sup> While in this definition Caplin invokes the principle of tonal prolongation, Yust might argue that this represents an alignment of formal and tonal features. The driving formal feature of a presentation phrase is repetition. Similarly, Caplin's definition of a continuation phrase is "a medial intrathematic function that destabilizes the prevailing formal context by means of fragmentation, harmonic acceleration, faster surface rhythm, and harmonic sequence," invoking tonal features once again by indicating the implications for harmonic motion during this functional unit.<sup>105</sup> By disentangling the tonal structure from this definition Yust argues that the essential formal criteria for a continuation phrase is fragmentation, which can play out in a variety of ways including (but not limited to) sequential repetition or what Yust more generally terms *motion*.<sup>106</sup>

Finally, Caplin's theory defines a cadential phrase as "a concluding intrathematic function that produces the requisite conditions for thematic closure. It is supported exclusively by one or more cadential progressions," which in turn he defines as "a

---

<sup>104</sup> Caplin, *Classical Form: A Theory of Formal Functions for the Instrumental Music of Haydn, Mozart, and Beethoven*, 256.

<sup>105</sup> Caplin, 254.

<sup>106</sup> Yust, *Organized Time: Rhythm, Tonality, and Form*, 61–63.

progression that confirms a tonality by bringing about its fundamental harmonic functions.” Yust responds yet again by disentangling the formal and tonal aspects of this definition of cadential phrase by arguing that musical cadences, and by extension, cadential progressions and phrases are not solely harmonic phenomena. The essential formal criteria for a cadence is “discontinuity: a resting point in the melody or a distinct change in texture.”<sup>107</sup>

Yust adopts the term *caesura* to refer to such occurrences, in effort to differentiate between the more traditional concept of cadence, which involves a coordination between essential tonal and formal criteria. In short, Yust’s reduced set of form-structural criteria enables Caplin’s theory of formal functions to apply in contexts where the common practice concepts of tonality do not explicitly play a driving role. Compare, for example, the relationship between a musical “sentence” and the formal structure of a limerick (Figure 3.1).

Measures 1–2 of Mozart’s Piano Sonata in C major, K.330, fulfill the function of a basic idea and are repeated in mm. 3–4. Together, mm. 1–4 form the presentation phrase of a sentence (a basic idea followed by a repetition of that basic idea). Measures 5–6 reduce the unit lengths from two measures long to one measure long, employing a technique of fragmentation, particularly common in a continuation phrase.

---

<sup>107</sup> Yust, 62.

**Figure 3.1 Comparison of musical and poetic sentences.**

a) Mozart Piano Sonata in C Major, K.330, *i*, mm.1–8.

Presentation

basic idea                      basic idea %

C: I<sub>PED</sub> (V<sup>7</sup>) I (V<sup>7</sup>) I

Continuation

frag.                      frag.                      cadential

(IV) I  $V^8_6$   $\begin{matrix} 7 \\ 5 \\ 4 \\ 3 \end{matrix}$  I<sup>#2-3</sup>

IAC

b) Limerick from St. John’s, Nova Scotia, circa 1880.

Presentation

basic idea                      basic idea %

There was a young rustic named Mallory      who drew but a very small salary

Continuation

frag.                      frag.                      cadential

when he went to a show      his purse made him go      to a seat in the uppermost gallery

“CAD”

Measures 7–8 then employ a standard harmonic formula that Caplin terms a “cadential progression.” The limerick employs an analogous structure: the first line of poetry is easily understood as primary material and, as such can be thought of as a basic idea.

The repetition of its rhythmic and rhyme structure in the second poetic line suggests that the first two lines of the poem form a presentation phrase. The third and fourth lines of poetry are noticeably reduced in length, comprising roughly half as many syllables as the first and second lines of poetry, thus emulating the same kind of fragmentation seen in mm. 5–6 of the Mozart example. The limerick concludes with a final line that reflects the rhythms of, as well as rhymes with, the poem’s opening material.

The question arises, naturally, as to whether it is reasonable to call this last line “cadential” in the same way that mm. 7–8 of the Mozart excerpt are “cadential.” Caplin defines tonic function as “the central *harmony* of a *key*, the one to which all others relate, and derive their meaning.”<sup>108</sup> These harmonic functions represent the elements of a formulaic schema based on two essential features: its elements are directionally charged, and relationally determined. Rather than defining the terms and concepts of formal functions according to the specific, ordered harmonic functions fundamental to tonality—tonic, pre-dominant, dominant, (and back to tonic), in that order—these specific elements and their particular order can be thought of as the specific formulaic schema that characterizes Classical music’s unique way of functioning in and through time.

---

<sup>108</sup> Caplin, *Classical Form: A Theory of Formal Functions for the Instrumental Music of Haydn, Mozart, and Beethoven*, 257.

Although the specific notion of tonality is not pertinent to the structure of a limerick, the essential features of a directionally charged and relationally determined schema are most certainly present in the form of rhythm and rhyme. Just as the presentation phrase in Mozart's Piano Sonata in C major, K.330 emphasizes the tonic harmony, the limerick's "presentation" phrase emphasizes the end-rhyme between the words "Mallory" and "salary." The return of this rhyme at the end of the poem is no less functionally formulaic than is the return of the tonic (and descending third gesture in the piano right-hand) at the close of the piano sonata's main theme.

In the following two analyses of slam poems by Marc Kelly Smith (sections 3.4 and 3.5), several formal functions are used to demonstrate the relationship between the formal structure of these works and the familiar formal structures often found in common practice music. For the reader's ease, the following table summarizes the functions and formal concepts discussed in section 3.4, providing revised definitions for each which extricate reference to specific tonal devices and emphasize essential formal criteria (Figure 3.2).

**Figure 3.2 Table of Generalized Definitions for Formal Functions.**

Basic Idea (b.i.)	An initiating function consisting of a relatively short idea that usually contains several internal motivic elements constituting the primary material of a theme.
Contrasting Idea (c.i.)	A concluding function consisting of a relatively short unit that follows and contrasts with (i.e., is not a repetition of) a basic idea.
Presentation	An initiating intrathematic function consisting of a unit (usually a basic idea) and its repetition.
Continuation	A medial intrathematic function that destabilizes the prevailing formal context by means of fragmentation, acceleration, faster surface rhythm, and sequential activity.
Cadential	A concluding intrathematic function that produces the requisite conditions for thematic closure, articulated by motion towards caesura.
Cadential Progression	A formulaic schema that leads towards caesura—a moment in time marking the structural end of a theme or theme-like unit
Cadence/Caesura	The goal of a cadential unit, marked by discontinuity: a resting point or distinct change in the sounding activity.
Sentence	A simple theme consisting of a presentation phrase and a continuation (or continuation =>cadential) phrase.
Sequential Progression	A progression that projects a consistent pattern, classified according to the governing invariable element which links each unit within the progression (eg. descending 5 <sup>th</sup> sequence in tonal musi; alliterative sequence in poetic music).
Abandoned Cadence	The failure to realize an implied or anticipated caesura by altering the material which previously signaled resolution.
Fragmentation	A reduction in the length of units in relation to the prevailing grouping structure. Fragmented units do not necessarily contain melodic-motivic material derived from the preceding units.
Small Ternary	A tripartite theme consisting of an exposition (A), contrasting middle (B), and recapitulation (A').
A Section (small ternary exposition)	An initiating intrathematic function consisting of a complete thematic unit ending with a major caesura. The first unit of the small ternary form.
Contrasting Middle	A medial intrathematic function that loosens the prevailing formal organization through fragmentation, motion, and/or sequential activity, and closes with contrasting caesura material. The second unit of the small ternary form.
A' Section (small ternary recapitulation)	A concluding intrathematic function that represents a return (often adjusted and altered) of an earlier exposition. The third unit of the small ternary form.

While most definitions should be clear, the concept of “cadence” requires additional discussion, as in this instance it is difficult to retain the term without invoking tonal

features. Yust's concept of caesura, which does not comprise any tonal requisites, functions in the same way as cadence does for Caplin. While slam poetry rarely involves such an overt and familiar rhyme scheme as the limerick example presented above, slam poems nonetheless do invoke formulaic patterns that guide the listener and function to articulate formal events. Often, those formulae operate according to what John Miles Foley would call "immanent art," in the sense that they are defined by and through the unfolding of a specific piece within a specific tradition and may not always transfer between works in identical fashion.<sup>109</sup> Despite the contextual emergence of a particular formula—its particular rhetorical content—its functional role can often be interpreted as analogous to one or another of the various formal functions identified in Caplin's theory, particularly when the invocations of tonality are teased out from the essential formal criteria that define each function.

### 3.4 Formal Structure of a Slam Poem

The limerick discussed in the previous section offers easy comparison to a musical sentence in no small part because of its metric regularity and codified rhyme scheme. In slam poetry, however, meter and rhyme are almost never used so systematically. Nevertheless, the form-functional concepts borrowed and adapted from music can tell us a lot about the structure of a slam performance. Smith's slam poem "Kiss It" is portrayed

---

<sup>109</sup> Foley, *How to Read an Oral Poem*, 120.

(by the poet) in text-form as beginning with two six-line verses, separated by a two-line “refrain”-like unit (Figure 3.3).<sup>110</sup>

**Figure 3.3 Opening two verses of Smith's "Kiss It," print format.**

<i>a</i>	If you need to kiss it,	<i>a</i>	If you need to leave it,
<i>b</i>	Kiss it.	<i>b</i>	Leave it.
<i>a</i>	If you need to kick it,	<i>a</i>	If you need to love it,
<i>b</i>	Kick it.	<i>b</i>	Love it.
<i>a</i>	If you need to scream it,	<i>a</i>	If you need to hold it,
<i>b</i>	Scream it.	<i>b</i>	Hold it.
<i>c</i>	But kiss it, kick it, scream it	<i>c</i>	But leave it, love it, hold it
<i>d</i>	Now.	<i>d</i>	Now.

A traditional, literary interpretation of the poetic structure, assigning each line alphabetic symbols to denote similarities and changes, would suggest that the first verse and refrain of “Kiss It” elicits an *ababab cd* structure (as does the second verse-refrain). When Smith performs the poem, however, his performance deviates from the text-version and, as a result, elicits a very different structure than the one suggested by the printed text. The emergent formal structure of the first verse in performance can be treated as a slam-poetic analog to a musical sentence (Figure 3.4).

**Figure 3.4 Form-functional analysis of verse 1 from “Kiss It” by Marc Kelly Smith.**

Presentation b.i.	b.i.	Continuation fragmentation		cadential		
If you need to kiss it kiss it	If you need to kick it kick it	If you need to scream it	scream it	If you need to scream it	scream it	but kiss it kick it scream it now
			(audience)		(audience)	Caesura

<sup>110</sup> The performance to which this analysis refers is the [2011 National Poetry Slam in Boston, MA](#). Accessed September 18, 2020.



Smith's performance of the first two lines replicates a presentation phrase: a basic idea ("If you need to kiss it, kiss it") is followed by an ever-so slightly modified repetition of that basic idea ("If you need to kick it, kick it"). Looking only at the printed text, the third line appears to be an additional repetition of the basic idea, thus extending the presentation phrase. But Smith's performance of this line is drastically different from way he delivers the previous two lines; not only does he insist on audience participation for the completion of the line, he does so twice, resulting in fragmentation of material derived from the basic idea. Smith recites the beginning of the line ("If you need to scream it...") then waits for the audience to finish the line for him. Seemingly dissatisfied with the audience's "scream," he repeats his fragment again, eliciting a more emphatic response from the audience. The effect of this fragmentation and repetition is a sense of continuation rather than initiating function. Fragmentation accelerates in the subsequent line, as does Smith's syllabic delivery. The line reiterates the last two syllables of the previous three lines, and Smith's performance of them motions toward the concluding word "now."

According to the musically informed concept of cadential progressions, this line cannot technically constitute a cadential idea. But the notion of a cadential progression can be generalized for application outside the domain of tonal music; that is, rather than defining a cadential progression according to the specific, ordered harmonic functions fundamental to tonality—tonic, pre-dominant, dominant, (and tonic if authentic), in that order—the specific elements of the progression and their particular order can be thought of as a formulaic schema that leads towards caesura—a moment in time marking the structural end of a theme or theme-like unit. Smith's timbral accentuation of this end

phrase and his durational elongation of the word “now” are both formulaic cadences in speech; they indicate to the listener that the sentential unit is complete. The first time we hear Smith recite this cadential idea, these phonological features imbue the line with a cadential quality; as the performance goes on, the words of the line itself become the formulaic schema that effects (or implies) a significant cadential gesture. Delivered at a much higher degree of performative intensity, Smith’s performance of the second verse of text is much closer to the printed version (Figure 3.5).

**Figure 3.5 Form-functional analysis of verse 2 from “Kiss It” by Marc Kelly Smith.**

Presentation			Cadential
b.i.	b.i.	b.i.	
if you need to leave it leave it	if you need to love it love it	if you need to hold it hold it	but leave it love it hold it now

*Caesura*

This time, Smith performs the third line in the same fashion as the first two lines; in other words, the basic idea is repeated three times, producing an extended presentation phrase.<sup>111</sup> What was the cadential idea of the first “sentence” here constitutes the entirety of the second phrase, resulting in a notably compressed cadential unit.<sup>112</sup> Whereas the first time that Smith delivered this line, our indications of cadential function were limited

<sup>111</sup> This poetic form of sentence is analogous to what Stephen Rodgers calls a ‘manic’ sentence with three presentations, exemplified in Schubert’s setting of the first verse of “Halt!” from *Schöne Müllerin*, mm.12-57. Manic sentential structure is also used in “Am Feierabend,” mm. 8–24. See Stephen Rodgers, “Sentences with Words: Text and Theme-Type in *Die schöne Müllerin*,” *Music Theory Spectrum* 36/1 (2014): 58–85.

<sup>112</sup> A similarly extended presentation comprising three iterations of a basic idea occurs in the first subordinate theme of Mozart’s Piano Sonata in D, K. 576, i, mm. 28–41. See Caplin, *Classical Form: A Theory of Formal Functions for the Instrumental Music of Haydn, Mozart, and Beethoven*, 98, example 8.1. A similarly compressed cadential phrase occurs in the first theme of Haydn’s String Quartet in F Minor, Op. 55, no. 2, iv, mm. 1–6.

to characteristics of vocal cadence and timing, now the unit itself has garnered a certain measure of cadential behavior. With the conclusion of the second sentence, the word-unit “But [verb] it, [verb] it, [verb] it, now” is established as a formulaic schema with cadential function. As Smith’s poem continues, the formal structure of the text—both as printed and as performed—becomes increasingly complex. Smith begins the third passage of “Kiss It” in the same way as the previous verses, but quickly veers off from the pattern he had established earlier, both on the page and on the stage (Figure 3.6).

**Figure 3.6 Final verses of “Kiss It” by Marc Kelly Smith, print format.**

<i>a</i>	If you need to squeeze it,	<i>c</i>	Then be it, tell it, spill it,
<i>b</i>	Squeeze it.	<i>c =&gt; i</i>	Squeeze it out of each
<i>a</i>	If you need to spill it,		instantaneous moment.
<i>b</i>	Spill it.	<i>j</i>	Make the juice, the jive, the jazz,
<i>a =&gt; e</i>	If you need to tell the world		the jism,
<i>f</i>	You’ve got more to you	<i>k</i>	The mysticism that ism you!
<i>g</i>	Than the world has as of yet	<i>l</i>	Grab at the moon!
<i>h</i>	Allowed you to be,	<i>m</i>	And hold the stars hot inside your head.
		<i>n</i>	'Cause now is all there ever was
		<i>o</i>	And all there ever will be.
		<i>c</i>	So kiss it, kick it, scream it
		<i>d</i>	Now!

What begins like a third repetition of the *ab* pattern—“If you need...”—becomes a far lengthier thought spanning the next several lines of text. In comparison with the brevity of line that had come before, this new passage takes on a run-on sentence quality. In the midst of this occurs a return to the refrain couplet, “Then be it, tell it, spill it”; but the couplet is not complete. Instead, it initiates another lengthy diversion that continues until the final two lines of the work.

Despite the particular spacing, the text as shown in Example 6 is far more continuously formed than the notion of “verses” delineated by the separation of text into groups of lines. When thinking musically about slam poetry, however, it is possible to consider the

formal structure in a different, and arguably more productive way. The now-familiar “presentation” phrase is prattled off quickly, at nearly twice the speed as the analogous lines had hitherto been delivered (Figure 3.7).

**Figure 3.7 Form-functional analysis of verse 3 from “Kiss It” by Marc Kelly Smith.**

Presentation b.i.	b.i.	Continuation fragmentation					cadential	
If you need to squeeze it squeeze it	if you need to spill it spill it	if you need	to tell the world	you’ve got more to you	than the world	has as of yet	allowed you to be	then be it tell it

*abandoned cadence  
minor caesura*

A third, incomplete statement of the basic idea then gives way to a series of fragmented utterances. In Smith’s performance, the fragmentation is achieved by temporal and timbral means: Smith places brief pauses between each fragment, while the intonation of his voice for each fragment is directed towards a similar frequency. These fragments exhibit a notable prosodic pattern in which a three, then four, then five syllable pattern is repeated twice, and followed by what has come to serve as cadential material: “Then be it, tell it...”.<sup>113</sup> This indicates a progression towards a cadence that has up to now been accomplished by the word “now.” In this passage, however, Smith pauses after the imperative, “tell it,” and instead of completing the cadential gesture with “squeeze it, now,” he abandons the formulaic ending he had established by turning the third imperative “squeeze it” into a new beginning.<sup>114</sup>

<sup>113</sup> Note that in print-format, this line reads “then be it, tell it, spill it;” in Smith’s [2011 National Slam performance](#), however, he omits “spill it” (refer to 00:45–00:49).

<sup>114</sup> A similar musical example might be found in the opening of Chopin’s Mazurka in g minor, Op. 67, no. 2, mm. 1–16. Measures 7–8 set up a half-cadential progression such

Using Caplin's terminology, the passage is far more "loose-knit" than the previous two phrases. Interpreting Smith's performance in this way renders the print-based delineation between the last line of verse 3 and the first line of verse 4 basically inconsequential with regard to formal structure, since the two lines participate as functional units of the same continuation phrase. Similarly, the performance-based fragmentation of "If you need / to tell the world / you've got more to you / than the world / has as of yet / allowed you to be" substantially differs from the 'lines' indicated in the written text.

If the various phrases of the performative structure for "Kiss It" discussed thus far resemble the kinds of thematic formal types and inter-thematic functions we often come across in Classical music of the high Viennese school, then on a larger level, the overall form of the performance also resembles a common formal structure of Classical music, namely, the "balanced sectional binary."

Binary form is, as one might expect, comprised of two parts and each part is typically repeated.<sup>115</sup> The first part of small binary usually consists of a tight knit, conventionally constructed theme that is often repeated. The second part often begins with material that is derived from or possibly identical to the opening material of the first part and, in the case of a balanced binary form, digresses from that material for a time (often both harmonically and thematically) and concludes by bringing back the closing material of the A section. These features are evident in Smith's slam poem (Figure 3.8).

---

that a half-cadence is expected in m. 9; instead, however, the cadential dominant is inverted and the cadential arrival abandoned for a new beginning.

<sup>115</sup> A good example for comparison is Bach's Polonaise in G minor from the Anna Magdalena Notebook, BWV Anh 119.

**Figure 3.8 Full binary form analysis for "Kiss It" by Marc Kelly Smith.**

A SECTION	Presentation b.i.		Continuation fragmentation				cadential					
	If you need to kiss it kiss it it	If you need to kick it kick it	If you need to scream it to scream it	scream it	If you need to scream it	scream it	but kiss it kick it scream it now  <i>caesura</i>					
B SECTION	Presentation b.i.		b.i.		b.i.		Continuation cadential					
	If you need to leave it leave it		if you need to love it love it		if you need to hold it hold it		but leave it love it hold it now  <i>caesura</i>					
B SECTION	Presentation b.i.		Continuation fragmentation				cadential					
	If you need to squeeze it squeeze it	if you need to spill it spill it	if you need d	to tell the world	you've got more to you	than the world	has as of yet	allowed you to be	then be it tell it  <i>abandoned caesura</i>			
	Compound Basic Idea b.i.		Continuation alliterative sequence		rhyme		allit.sequence					
squeeze it out of each		instantaneous moment		let the juice	the jive	the jazz	the jism	the mysticis m	that ism you	grab at the moon	and hol d	the star s your hot head <i>open cadence</i>
Extended Cadential Progression					A <sup>1</sup> (closing)							
'cause now is all there ever was					and all there ever will be		so kiss it kick it scream it now  <i>caesura</i>					

Together, the first two sentences comprise a closed A section. Despite the small variation in their sentential construction, these phrases can be thought of as a modified repeat rather than the traditional literal repeat of the A section that is standard in binary form. The B section begins with a presentation phrase that closely parallels the opening material of the A section, but soon digresses into new and expanded material that reaches an open cadence before concluding with material taken directly from the closing cadential progression of the A section.

Smith's "Kiss It" is a straightforward example of how formal functions can not only be productively applied to study the emergent form of a slam performance, but that they are also far *more* productive analytically than a line/verse based textual analysis. While the notions of "verse" and "line" were efficient for my analysis of the formal structure of 'the first two verses,' correspondences between the spatial layout of a slam poem in print-format and the formal features of a slam performance should never be assumed, nor assumed to be of consequence. Any analysis of a musical work that took the staves of its score to be indicative of formal, musical boundaries would at once be suspect (and rightly so). Of course, it is not impossible for the spatial layout of a score to mark formal-structural boundaries of the music—and indeed, it is not uncommon to see scores in which the spatial boundaries of a staff or a page coincide with a formal boundary of a theme or section of music—but the idea that a spatial layout might be the initial point of investigation with regard to form is an absurd premise.

With regard to formal structure, text-based poetic analysis is ill-suited to be the primary analytical tool used to investigate the formal features of slam poetry. Music-theoretical approaches like that of form-functionality—although perhaps still not perfectly applicable or generalizable—can draw attention to formal structures emergent in a slam performance that are otherwise obscured by analytical approaches such as ethnopoetics.

### 3.5 Speaking of/in Sonata Form

Smith's 2011 performance of "I Wear My Father's Coat" is yet another, somewhat more complex example. On one occasion where Smith was performing this piece, he noted that "I Wear My Father's Coat" perhaps is already his legacy, saying that "this is probably the

only poem that will ever be remembered after [he's] dead.”<sup>116</sup> In print versions, the work's text is often presented as a series of several verses of varying line lengths that do not rhyme or project any consistently metrical poetic pattern.

From a poetic perspective of form or structure, “I Wear My Father's Coat” falls into the expansive category of poetry deemed free verse, and to talk about large-scale formal structure beyond free verse from a literary perspective in which poetic forms are primarily distinguished by meter and rhyme would be relatively unproductive. From a musical perspective—particularly one that adopts the language and formal strategies of Caplin's functionally driven theory—a much different interpretation of the formal design emerges. The first two phrases of “I Wear My Father's Coat” reflect a similar opening strategy employed in “Kiss It,” namely that of varied repetition of a main idea (Figure 3.9).

**Figure 3.9 Main Theme of “I Wear My Father's Coat” by Marc Kelly Smith.**

Compound Basic Idea/Antecedent b.i.		Cadential c.i.		
I'm wearing my father's coat	he has died	I didn't like him but I wear the coat		
<i>minor caesura</i>		<i>cadential caesura</i>		
Compound Basic Idea/Antecedent b.i.		Cadential c.i.		(post- cadential)
I'm wearing the coat of my father	who is dead	I didn't like him but I wear the coat	just the same	
<i>minor caesura</i>		<i>cadential caesura</i>		

The basic idea, “I'm wearing my father's coat” is followed by a contrasting idea, “he has died.” Together, these two units form either an antecedent phrase or a compound basic

<sup>116</sup> Intro to “I Wear My Father's Coat”

[https://www.youtube.com/watch?v=\\_fqtNDkuPcw](https://www.youtube.com/watch?v=_fqtNDkuPcw). Accessed September 20, 2020.



idea (CBI), depending on whether 1) the passage involves a cadential formula *or* 2) whether the distinction between the two phrase types requires coordination between content and form. On the one hand, given that cadential formulae are necessarily immanently determined and that we have little information with only two short units of material with which to work, the phrase would be difficult to construe as an antecedent. Without definitive cadential closure, the logical interpretation of the phrase is that it functions as a compound basic idea. On other hand, Yust's separation of tonal and formal structure suggests that the essential difference between CBI and antecedent phrases is not technically a formal distinction.<sup>117</sup> Rather, the distinction refers to the presence or absence of formal and tonal coordination. For the repertoire at hand, where coordination with tonal structure (in the classical music sense) does not factor in, one finds a distinction without difference: any interthematic initiating phrase is either a presentation or an antecedent, defined respectively by repetition or contrast.

The subsequent phrase, "I didn't like him but I wear the coat," however, might rightly be called a cadential phrase. To a certain extent, I consider the phrase cadential because of a process of elimination of other end-phrase alternatives. Caplin identifies three main types of concluding interthematic phrases: cadential phrases, consequent phrases, and continuation phrases (of which there are various types, depending on the technique employed). The phrase is not a consequent, as the basic idea of the first phrase is not the opening material of this phrase; and the phrase does not employ any of the typical techniques one expects of a continuation (fragmentation, or breakdown in the size of

---

<sup>117</sup> Yust, *Organized Time: Rhythm, Tonality, and Form*, 62.

melodic units; liquidation or the removal of “characteristic” figures; sequential repetition(s); and accelerated surface rhythm or activity). Thus, this leaves the cadential phrase type. At the same time, the cadential function of the phrase is reinforced by Smith’s performance, as he employs familiar strategies for linguistic cadence. His intonation descends, his vocal amplitude decreases, and he slows down his speech as he reaches the end of this phrase. As if to confirm the structure (and the significance of this theme as the *main* theme for the work), Smith reiterates it a second time with slight modifications: the possessive contraction of “father’s coat” is reversed to “coat of my father” and the tense of the contrasting idea is changed from passive to active.<sup>118</sup>

With these two utterances, the main theme of “I Wear My Father’s Coat” is firmly established and new, anecdotal material serves to transition from main theme’s observational remove to the introspective musings of the forthcoming subordinate theme (Figure 3.10).

**Figure 3.10 Transition of “I Wear My Father’s Coat” by Marc Kelly Smith.**

Antecedent		Continuation				
b.i.	c.i.	fragmentation			cadential	
a younger man stopping on the street	asks where’d you get a coat like that?	I answer:	it was my father’s	who’s now gone	passed away	the younger man shuts up
<i>open cadence</i>				<i>closed cadence</i>		

The first phrase of the transitional material presents a literal, linguistic question, which should easily be interpreted as analogous to an antecedent phrase in music (a phrase type

<sup>118</sup> The inclusion of “just the same” in the second version of the theme is in brackets because Smith only sometimes performs it. Whether he does or doesn’t changes little about the structure of the theme, as the unit functions essentially as closing material following the actual cadence established by “I wear the coat.”

that is often associated with question-like material, particularly when ending with a half-cadence). Yet the analogy is perhaps not so easy to prove in theory, in no small part because it draws attention to the problem of applying the musical concept of a half-cadence to repertoire that lacks harmonic syntax. The problem is clear in Caplin's definition of half-cadence, as "a cadential arrival articulated by the final dominant of a half-cadential progression." The issue is a cyclicity of definition, where the concept being defined (a half cadence) is referenced in the definition itself as the half-cadential progression (or passage, in the case of slam poetry). What is a half-cadential progression, then? In terms original to Caplin's theory, a half-cadential progression is "a cadential progression whose complete form brings, in order, the harmonic functions of tonic (usually in first inversion), pre-dominant, and dominant (triad in root position)."<sup>119</sup> Such a definition is so steeped in common practice traditions that to tease out the specific tonal context is near-impossible. While the notion of a half-cadence might not fully apply, the functional properties of a half-cadential event can be teased out. The idea that the passage reaches a cadence that is open — implying another phrase will follow that will lead to a more convincing arrival — is, I think, a reasonable interpretation here.

The first phrase of the transitional passage of "I Wear My Father's Coat", the question Smith poses articulates an open cadence, as it does not sound final, and it implies that another phrase will follow, and that the subsequent phrase will answer the question he poses both semantically and through an arrival articulated by a final timbral descent. Indeed, this is precisely what Smith does: he answers the question, not in the form of

---

<sup>119</sup> Caplin, *Classical Form: A Theory of Formal Functions for the Instrumental Music of Haydn, Mozart, and Beethoven*, 255.

consequential material but as a continuation of the anecdote, fragmenting the loss of his father and establishing new cadential material through the emphatic end-unit, “the younger man shuts up.” Altogether the passage fulfils much of the functional roles of a transition, if the contextualized term “key” is (temporarily) replaced with “material”: “an interthematic function that destabilizes the home ‘material’ and loosens the formal organization in order for a subordinate ‘material’ to be established and eventually confirmed.”<sup>120</sup>

Drawing on the rhetorical expectations of a transition, which Hepokoski and Darcy (2008) rigorously explore, this passage in Smith’s performance can be understood all the more as transitional.<sup>121</sup> Transitions, according to Hepokoski and Darcy, are frequently concerned with “rhetorical energy-gain” and often involve “a marked change of musical topic, rhythmic motion, and/or figure,” which is precisely what Smith creates in performance.<sup>122</sup> He speeds up during the antecedent phrase in every performance I have found of the work, and employs a new vocal timbre to state the younger man’s question.

The final timbral descent of his delivery at the end of the passage, together with a deceleration in his rhythmic delivery and a reduction in dynamics all elicit a sense of cadence, but not a cadence that we have heard before. In this way, one could argue that the transition passage is modulatory, closing on a clear cadence, but not the same cadence

---

<sup>120</sup> Caplin, 258.

<sup>121</sup> James Hepokoski and Warren Darcy, *Elements of Sonata Theory: Norms, Types, and Deformations in the Late-Eighteenth Century Sonata* (Oxford: Oxford University Press, 2008), 93–116.

<sup>122</sup> Hepokoski and Darcy, 94–95.

as established by the main theme (nor the same cadence that will ultimately close the subordinate theme... for now, the best we can say is that it *is* some kind of closed cadence, and that whatever kind of cadence it is may only be determinable in retrospect. Where the emotional state of the main theme of “I Wear My Father’s Coat” is one of observational remove, the subordinate theme is undoubtedly contrasting, as the next passage delves into a poignant description of the father, now gone, passed away (Figure 3.11).

**Figure 3.11 Subordinate theme from “I Wear My Father’s Coat” by Marc Kelly Smith.**

CBI b.i.	c.i.	Sequential progression model			ECP		
It's not that I'm trying now to be proud of my father	I didn't like him he was a narrow man	There was	more of everything he should have done	more of what he should have tried to understand	the coat fit him well	it fits me now	I didn't love him but I wear the coat
							<i>closed cadence major caesura</i>

Typical of subordinate themes, the passage here is much looser knit than the main theme. The basic idea, “it’s not that I’m trying now to be proud of my father” is considerably longer than that of the main theme, if we think in syllable or word count. Durationally, on the other hand, where the basic idea of the main theme spans 3.2 seconds of time, the basic idea I propose for the subordinate theme similarly spans 3.4 seconds of time. And, despite the few words that make up the contrasting idea of the subordinate theme, “I didn’t like him / he was a narrow man,” the unit also spans just over three seconds of time. In the same way as the main theme, no clear cadential content, rhetoric, or function emerges at the close of these two units, and the full phrase thus presents a compound

basic idea. The continuation phrase that follows begins with a sequential repetition in the two statements of what his father should have done and should have tried to do. These, in turn, lead into a series of four units that, taken together, form a nice analog for an expanded cadential progression, in the same, essential sense as Caplin finds common in subordinate theme cadences in Haydn, Mozart, Beethoven, and Schubert.<sup>123</sup>

While one could argue that the close identity between the cadential caesuras at the end of subordinate and main themes undermines, rather than reinforces, the notion of sonata form for this piece, such an argument would overlook the close identity between main theme and subordinate theme cadences in common practice tonal music. Indeed, while we deem the contrasting nature of those cadences essential to sonata form itself, most of the features of those cadences are, in fact, the same. We require the same underlying harmonic progressions, the same voice-leading, and the same kinds of rhetorical strategies for closure—the only crucial difference is that of key. When considering the function of a transition, the term “key” was only temporarily replaced with the word “material.” Smith’s cadence at the end of the subordinate theme in “I Wear My Father’s Coat” suggests that a better analog to musical keys is poetic states: of mind, emotion, being, etc.

Whereas the main theme cadences involve the phrase “I didn’t like him,” here Smith replaces the word “like” with “love”—a subtle but significant change in relation to his emotional state as he reflects on his father. The word “love” injects this particular

---

<sup>123</sup> William E. Caplin, “The ‘Expanded Cadential Progression’: A Category for the Analysis of Classical Form,” *Journal of Musicological Research* 7, no. 2–3 (1987): 215–57.

cadential moment with heightened feeling and creates a distinct contrast between the close of the two expositional themes while simultaneously retaining the immanently designed cadential formula that ultimately governs the formal structure of work.

Before summarizing the form of the exposition of “I Wear My Father’s Coat,” I will revisit the cadential caesura at the end of the transition. On one hand, the rhetorical, syntactic, and (even perhaps) semantic features of the phrase “the younger man shuts up” all suggest a sense of closure that might be akin to an authentic cadence, the content of this phrase has not been set up in such a way necessary for authentic formal closure.<sup>124</sup>

For me, this particular poetic cadence resonates in a similar way as do tonicized half-cadences, in the sense that it has some elements of authentic closure, yet still involves some degree of unresolved tension, an open-ness for new material to follow. In the print version of the piece, “I Wear My Father’s Coat,” the passages I have discussed thus far are often presented as seven verses ranging from two to four lines long, and line lengths ranging from two words to nine words.<sup>125</sup> It would be difficult to attribute some poetic form other than “free verse,” but if one thinks about musical forms and the functions that create them, one can begin to imagine not twenty-two lines of variable lengths grouped in seven irregularly sized verses, but rather as three parts comprising a main theme that expresses a state of observational remove or ambivalence, followed by a transition that

---

<sup>124</sup> It has not, as Foley would say, established enough “word-power” to achieve the status as a final specialized formula.

<sup>125</sup> Described here as published by <https://www.rattle.com/my-fathers-coat-by-marc-kelly-smith/> (accessed September 20, 2020) and as published in David Starky and Richard Guzman, eds., *Smokestacks and Skyscrapers: An Anthology of Chicago Writing* (Chicago: Wild Onion Books, 1999). Also in Marc Kelly Smith, *Crowdpleaser* (Chicago: College Press, 1996).

anecdotally opens up space for, finally, a subordinate theme that turns inward to a state of emotional reflection. In other words, what we have thus far is an exposition.

The following section of “I Wear My Father’s Coat” fits Caplin’s model for development as a formal unit, in that it “stands between the exposition and the recapitulation,” but also (albeit abstractly as always) fits as a formal function, so long as his definition is sufficiently generalized. Of the development he writes: “As a formal function, the development generates the greatest degree of tonal and phrase structural instability in the movement and thus motivates a restoration of stability (to be accomplished by the recapitulation).”<sup>126</sup>

While a great deal of Caplin’s theory emphasizes the specifics of the development’s tonal organization, which for his purposes is appropriate, he also places significance on the phrase-structural organization and form-functional aspects of development sections. He notes:

The phrase-structural technique most characteristic of a development involves the establishment of a relatively large model, which is repeated and sequenced one or more times. Subsequent fragmentation leads to an open cadence (or dominant arrival) of either the home key or a development key, after which a standing on the dominant typically appears.<sup>127</sup>

Following Erwin Ratz, Caplin calls this developmental process (and the resultant unit it creates) a *core*, and further notes that cores are most often preceded by *pre-cores*, a

---

<sup>126</sup> Caplin, *Classical Form: A Theory of Formal Functions for the Instrumental Music of Haydn, Mozart, and Beethoven*, 139.

<sup>127</sup> Caplin, 141.



thematic unit (which can be complete or incomplete) that essentially sets up the stage for core processes to unfold.<sup>128</sup> The subsequent passage in Smith’s “I Wear My Father’s Coat” begins with an incomplete pre-core unit comprising “most of us show off to one and other fashions of who we are” (Figure 3.12).

**Figure 3.12 Development from “I Wear My Father’s Coat” by Marc Kelly Smith.**

Pre-Core CBI		Pseudo-core		
b.i.	c.i.	Sequential progression		=> continuation
most of us show off to one and other	fashions of who we are	sometimes buttoned up to the neck	sometimes overpriced	sometimes surprising even ourselves in garments we would have never dreamed of wearing

*open cadence*

There are two ideas presented in this phrase: first, the idea of showing off and second, the sort of thing about ourselves that we show. The second idea is not a reiteration of the first in prosodic structure, syntactic structure, or semantically. While there are similarities to the melodic contours of Smith’s delivery of these lines across multiple performances, the phrase concludes with material that has not previously been established as holding cadential significance. It is reasonable to think of this as a compound basic idea, which functions as a pre-core—a set up for three examples (stated in sequence) of the kinds of fashions we wear and in which we show off.

<sup>128</sup> Caplin’s own account of the historical lineage of of formal functions traces the theory from Arnold Schoenberg, *Fundamentals in Music Composition*, ed. Gerald Strang and Leonard Stein (London: Faber & Faber, 1967) through to Erwin Ratz, *Einführung in Die Musikalische Formenlehre: Über Formprizipien in Den Inventionen Und Fugen J.S. Bach Und Ihre Bedeutung Für Die Kompositionstechnik Beethovens*, 3rd ed. (Vienna: Universal, 1973). He attributes the concepts of and terminology for developmental formal functions to Ratz in Caplin, *Classical Form: A Theory of Formal Functions for the Instrumental Music of Haydn, Mozart, and Beethoven*, 152.

When we speak of sequences in common practice music, we tend to expect a certain degree of consistency that might not readily be apparent in the language Smith uses here. Indeed, Caplin incorporates this expectation of pattern uniformity in his definition of sequences, sequential progressions, sequential repetitions and the model-sequence technique. But there is a difference (albiet a nuanced one) between a sequential progression and a sequential repetition (which is considered synonymous with the model-sequence technique). And the key word that differentiates the two is “projects.” A sequential progression “projects a consistent intervallic pattern among the individual voices of the harmonies. It is classified in terms of the intervallic motion of its constituent roots...”<sup>129</sup> That is, a sequential progression “projects a consistent intervallic pattern,” but do not actually *have* consistent intervallic patterns. In his analysis of the development from Haydn’s Symphony No. 102 in B-flat, ii, for example, Caplin notes that the “continuation (upbeat to m. 39) contains a sequential progression (but no model-sequence technique) as the music modulates back to the home key.”<sup>130</sup> Similarly in an analysis of the theme from Beethoven’s Andante in F, WoO 57, mm. 1–8, he differentiates between a mm.1–2 in which the basic idea is “repeated sequentially” and “a relatively rapid progression of harmonies that are largely sequential until the pre-dominant at m. 3.”<sup>131</sup>

---

<sup>129</sup> Caplin, *Classical Form: A Theory of Formal Functions for the Instrumental Music of Haydn, Mozart, and Beethoven*, 257.

<sup>130</sup> Caplin, 157. Score provided on p. 156.

<sup>131</sup> William E. Caplin, “Teaching Classical Form: Strict Categories vs. Flexible Analyses,” *Dutch Journal of Music Theory* 18, no. 3 (2013): 134. Score provided on p. 133.

The distinction between sequential progressions and sequential repetitions (involving the model-sequence technique) are crucial for his inter-thematic functional distinction between cores and “pseudo-cores.” Whereas a genuine core requires a “well-articulated process of model, sequence, and fragmentation,” a pseudo-core “features a prominent sequential organization of the harmonies, but they are not used to support the extensive model-sequence technique.”<sup>132</sup> In “I Wear My Father’s Coat,” the examples of the kinds of fashions we wear (and in which we show off) each begin with “sometimes,” projecting a kind of sequential patterning. In this way, the passage might be thought of as the kind of sequential progression that does not involve a model-sequence technique and, as such, the whole unit functions as a pseudo-core.<sup>133</sup> The third “sometimes” statement breaks off into a lengthier musing, and in doing so “destabilizes the prevailing phrase-structural, rhythmic, and harmonic context,” transforming the sequential progression into a continuation involving phrase extension and increased rhythmic activity.<sup>134</sup>

The passage in Smith’s slam poem I call the development concludes with a caesura moment at the end of the utterance “garments we would have never dreamed of wearing.” The poignancy of Smith’s pause at this moment garners enough weight to the caesura to merit a cadential function, but not one that is closed to the same degree as previous

---

<sup>132</sup> Caplin, *Classical Form: A Theory of Formal Functions for the Instrumental Music of Haydn, Mozart, and Beethoven*, 55.

<sup>133</sup> I think if the prosodic pattern of each example were identical as well, then there would be a case to argue that the passage involved sequential repetition. In addition to the “sometimes” pattern, there is another bonding element of these phases, in their inherent affective and semantic connections.

<sup>134</sup> Caplin, *Classical Form: A Theory of Formal Functions for the Instrumental Music of Haydn, Mozart, and Beethoven*, 40.

cadential events. In this case we might consider the end of the development to function as an *open cadential caesura*, preparing the way for recapitulatory material. What follows this event is indeed a return to the opening line of the work, “I wear my father’s coat” but beyond that, the final lines of the piece deviate from expositional material to the extent that one might question whether the concept of recapitulation applies at all. And if the concept of recapitulation does not apply, then to what degree is it appropriate to consider the work as a whole to be in sonata form?<sup>135</sup>

Caplin’s theory of formal functions defines recapitulation as “a large-scale concluding function that brings back, usually modified, an earlier exposition. It resolves tonal conflicts by adjusting all material into the home key.”<sup>136</sup> While the emphasis on tonal resolution problematizes the concept’s efficacy in context where common practice tonality plays a less central (or no) role in a work’s structure, we can (as we have done throughout) distill the essential formal criteria from the tonal conventions of the repertoire to which Caplin specifically refers. The issues remaining then, are whether the material of the final passage constitutes a sufficient “bringing back” of the earlier exposition and whether, should the passage not suffice to bring back the requisite expositional material, the concept of sonata form can still apply. While Caplin does address a number of circumstances in which a recapitulation may involve substantial

---

<sup>135</sup> To the same degree, I imagine, that one considers Haydn’s String Quartet no. 48 in B Minor, Op. 64 No. 2, 1<sup>st</sup> mvt, is in sonata form, despite having a recapitulation that is entirely different from its exposition. The interested reader will find a form-functional analysis of this work in Caplin, 175–76.

<sup>136</sup> Caplin, 256.

modifications to expositional material, the situation where a recapitulation is truncated (a term used by Hepokoski and Darcy) does not seem to be one that Caplin recognizes.<sup>137</sup>

If the key purpose of a recapitulation is resolution through the return of expositional material and particularly that material which occurred initially in contrasting key areas, and if such a return is a necessary requirement of sonata form (which a tripartite perspective seems to suggest), then despite all I have proposed thus far, “I Wear My Father’s Coat” cannot, by its conclusion, be a sonata. An alternative take on sonata form, however, is that it is a bipartite form defined not by what happens in the second part but crucially according to what occurs in the first. I turn once more to the recent work of Jason Yust for explanation. In Yust’s investigation of sonata form, he writes: “sonata form as it is presently understood existed within a continuum of formal types, including those that lack a main theme recapitulation or a fully-fledged development section. What all these types share is an exposition with certain predictable design features.”<sup>138</sup>

To put it plainly, what makes a work sonata form is not how it concludes but rather how it sets itself up. The form involves at least two themes (though it is quite possible to involve more) that are contrasting in at least one way (though it is quite common that the

---

<sup>137</sup> Sonata theorists James Hepokoski and Warren Darcy explain, “a genuinely truncated recapitulation... would be one that begins with P<sup>1</sup> but is cut off before the rotation has been completed, perhaps even before part 2 (S/C) has been entered.” Hepokoski and Darcy, *Elements of Sonata Theory: Norms, Types, and Deformations in the Late-Eighteenth Century Sonata*, 232. I do not know whether Caplin would accept truncated recapitulations, specifically those in which the subordinate theme is the material omitted. I only mean to say that I am not aware of any reference to or discussion of such situations in his writing on sonata forms.

<sup>138</sup> Yust, *Organized Time: Rhythm, Tonality, and Form*, 269.

themes contrast in more than one modality). Such a set up naturally lends itself to a conventional subsequent narrative strategy (or to be perfectly clear, dialectic) in which those contrasting themes are examined, compared, and developed, then ultimately brought into harmony with each other. But this is far from the only narrative according to which such a set up must proceed. Sometimes the themes do not lead to any development; sometimes their development can beget a new theme entirely; sometimes one or another theme exerts greater dominance than another and achieves resolution through subjugation; and sometimes a theme can be too complex to ever fully find a path to resolution.

It is the last of these “sometimes” that takes place in Smith’s poem. The final lines do indeed recapitulate expositional material, but only in part. The main theme returns in the motivic fragment, “I wear my father’s coat.” The transition returns by way of elusory reference to the meeting of strangers in the utterance “this is the way we make each other’s acquaintance.” Subtle similarities link this material to the expositions’ transition, particularly in the prosodic structure (Figure 3.13).

**Figure 3.13 Comparison of transition material from “I Wear My Father’s Coat” by Marc Kelly Smith.**

a) Transition material from exposition

A    youn - ger man    stop - ping me on    the street    has asked

b) Transition material from recapitulation

X								X		
X				X				X		
X	X			X			X	X		X

And it seems to me that this is the way the most of us

Both phrases share prosodic structures so similar in design that, rhythmically speaking, what is on one hand a new continuation is, on the other, a clear recollection of the expositional transition. With only slight rhythmic differences, both share essentially the same speech-rhythmic structures, represented here both as metrical stress grids and as proximate musical notation.<sup>139</sup> Abridged main theme and modified transition fuse together to form the phrase “I wear my father’s coat and it seems to me that this is the way most of us make each other’s acquaintance” (Figure 3.14).

**Figure 3.14 Recapitulation (with MT/trans. Fused) from “I Wear My Father’s Coat” by Marc Kelly Smith.**

Main Theme/transition	
b.i.	continuation
I wear my father's coat	and it seems to me that this is the way that most of us make each other's acquaintance
<i>minor caesura</i>	

The final utterance of this passage, “make each other’s acquaintance,” is interesting from a cadential perspective. The utterance shares qualities with other moments I have called

---

<sup>139</sup> The musical notation is not intended to denote the precise timing of Smith’s performance, but rather the underlying rhythmic structure of both phrases. While Smith’s performance does indeed correspond to these rhythmic structures, they are intended only representationally here.

open, but it does not properly evoke a sense of cadence. Rather it functions like what Yust would call a minor caesura, a term he invokes at moments analogous to Caplin's concept of dominant arrival, in which a theme-like unit articulates local closure but in a non-cadential way.

This phrase does not cadence, but rather arrives at a moment of opening. Such an effect evokes an expectation of new beginning, or more to come—and specifically, the return of subordinate theme material. The subordinate theme however, which centered upon the now passed away father, does not return. Indeed, how can it? What remains (as it goes in life and so too in the piece) is the coat and the choice to take it or leave it (Figure 3.15).

**Figure 3.15 Truncated recapitulation from “I Wear My Father’s Coat” by Marc Kelly Smith.**

Main Theme/transition b.i.	continuation	cadential
I wear my father's coat	and it seems to me that this is the way that most of us make each other's acquaintance:	in coats we have taken to be our own.
	<i>minor caesura</i>	<i>cadential caesura</i>

A recapitulation of this kind, where the explicit parallelism to expositional material is minimal at best and subordinate theme material is absent entirely occurs somewhat infrequently in the classical music repertoire that concern most theorists of musical form. But it is, nonetheless, recapitulatory, in the sense that it brings back what material it can and achieves some semblance of closure. Closure through resolution may be our desire, but closure through resigned acceptance is what we get.



### 3.6 Conclusions

In this chapter I have tried to demonstrate the ways in which theories of musical form and form functionality can shed light on the structural design of slam poetry, focusing on the works of slam founder Marc Kelly Smith. I have endeavored to draw parallels between the poetic devices that shape, expand, and guide the trajectory of poetic utterances and the functional devices that govern the shaping of musical forms. I chose to focus on slam poetry not because it is the only genre of modern poetry that responds productively to a music-analytical lens, but because 1) it is a type of spoken-word poetry with which I am most familiar and 2) it was a slam performance which sparked in me the idea that spoken-word poetry employs analogous formal functions to those expressed in musical forms.<sup>140</sup>

While my motivation for this experiment was driven by personal experience and epiphany, I am far from the first or only person to suggest that modern poetry is a musical art form. In 1942, poet T.S. Eliot presented a lecture on this very hypothesis, titled “The Music of Poetry,” in which he argues that “a ‘musical poem’ is a poem which has a musical pattern of sound and a musical pattern of the secondary meanings of the words which compose it, and that these two patterns are indissoluble and one.”<sup>141</sup> He refers to the “vexed question of formal pattern and free verse,” noting that while “it is sometimes assumed that modern poetry has done away with forms,” it is a mistake to think of free

---

<sup>140</sup> The particular performance that inspired this project was by Ottawa-based slam poet PruFrock Shadowrunner, 2015 Canadian Individual Poetry Slam Winner.

<sup>141</sup> T.S. Eliot, *The Music of Poetry: The Third W.P. Ker Memorial Lectures Delivered in the University of Glasgow 24th February 1942* (Folcroft, PA: The Folcroft Press, 1969), 19.

verse as a liberation from form; rather, it is a “revolt against dead form and a preparation for new form or the renewal of the old.”<sup>142</sup> Ultimately, Eliot believes that “the properties in which music concerns the poet most nearly are the sense of rhythm and the sense of structure” and concludes his lecture with the following passage:

The use of recurrent themes is as natural to poetry as to music. There are possibilities for verse which bear some analogy to the development of a theme; possibilities of transitions in a poem comparable to those in a movement of a symphony or a quartet; possibilities of contrapuntal arrangement of subject matter...More than this I cannot say, but must leave the matter to those who have had a musical education.<sup>143</sup>

In part, the application of a musical theory of formal functions is an attempt to demonstrate how spoken-word poetry makes manifest the possibilities Eliot imagines; but it is also an attempt to explore the reaches of a functional theory of musical form. Caplin’s theory is predicated upon privileging formal function over formal type, an approach that not only distinguishes his work from other scholars of musical form, but also (and more importantly) renders his theory far more malleable to repertoires beyond the classical realm. Yust’s theory of formal structure operationalizes that malleability by disentangling context-specific aspects of tonal structure from the criteria that define formal structure.

In the opening phrase of “Kiss It,” for example, I am able to identify a sentential analog not by type but by function: this poetic phrase involves a set of constituent functions that are recognizable as initiating, medial, and concluding. Similarly, in “I Wear My Father’s

---

<sup>142</sup> Eliot, 25–26.

<sup>143</sup> Eliot, 28.

Coat,” I am able to identify a subordinate theme not by its perfect resemblance to a familiar example in the classical music literature, but rather by its expression—on multiple hierarchical levels—of subordinate function. The criteria Caplin uses to identify formal functions involve multiple parameters that vary in terms of how easily they apply outside the context of the Classical repertoire, depending on how explicitly they rely on concepts of harmony and tonality. But in general, the privileging of function over type in Caplin’s theory renders it an incredibly versatile approach to form. Even in the absence of harmony and tonality, prolongational progressions engender a sense of formal initiation, sequential ones express medial functions, and cadential progressions create formal closure. While not every analytical interpretation I make in this chapter is equally convincing, my hope is that the idea can inspire in my reader a more musical way of listening to Spoken Word and indeed, of experiencing language in general. As Wittgenstein puts it, “understanding a sentence is much more akin to understanding a theme in music than one might think.”<sup>144</sup>

In the previous chapter on spoken-word songs, and in this chapter on slam poetry, the kinds of Spoken Word at play involve texts that are themselves composed of meaningful words organized according to English grammatical and syntactical systems. But the general definition of Spoken Word I present at the outset of this dissertation does not apply exclusively to such texts. Where my study of spoken-word songs endeavored to examine how Spoken Word can be musical when presented in combination with an unequivocal musical setting, and my study of slam poetry endeavored to investigate how

---

<sup>144</sup> Ludwig Wittgenstein, *Philosophical Investigations*, 2nd ed., trans. by G.E.M. Anscombe (Oxford: Basil Blackwell Ltd, 1958), frag. 527.

Spoken Word alone can emulate musical ideas and form, the following chapter investigates yet another kind of spoken word—the kind that is neither supported by an explicit musical setting nor involves words with any lexical correlates: sound poetry.

## Chapter 4

### 4 The Artist Composes

Within the vast medium of Spoken Word is yet another fascinating genre of sound poetry, written not through the composition of words imbued with semantic meaning according to syntactic laws but rather composed with phonemes expressly devoid of semantic or linguistic association. Sound poetry bridges literary and musical composition such that the phonetic aspects of speech are far more important than semantic and syntactic values. Very much a twentieth-century phenomenon, sound poetry emerged from the pioneering work of Futurist and Dadaist artists operating in the early years of the century. Founding artist of the Dadaist movement in Switzerland, Hugo Ball, called these works “verse without words,” and himself composed and performed a number of sound poems at Dadaist gatherings and events during the 1910s and 20s.<sup>145</sup>

Among the Dadaist crowds, Ball, Tristan Tzara and Raoul Hausmann led the development of sound poetry. Sound poems for solo performer (such as *K' perioum* (1918) by Raoul Hausmann) and simultaneous poems for poet ensembles (such as *Karawane* (1916) by Hugo Ball and *L'amiral cherche une maison à louer* by Richard Huelsenbeck, Marcel Janko, Tristan Tzara, and Toshiro Sawa) were staples of any Dada affair. In Futurist circles, F.T. Marinetti's program of “parole in libertà” is detailed in the

---

<sup>145</sup> Debbie Lewer, “Hugo Ball, Iconoclasm and the Origins of Dada in Zurich,” *Oxford Art Journal*, 2009, 29. Also Tobias Wilke, “Da-Da: ‘Articulatory Gestures’ and the Emergence of Sound Poetry,” *Modern Language Notes* 128, no. 3 (2013): 641.

“Technical Manifesto of Futurist Literature” from 1912, which similarly advocated for a quite radical negation of language.<sup>146</sup>

A particularly well-known example of sound poetry is Kurt Schwitters’s *Die Ursonate* (1922–1932). Translated as “the sonata in primordial sounds,” *Die Ursonate* provoked audiences who expected traditional prose, yet also continues to encourage listeners to make connections between unexpected sounds and, in doing so, create their own meaning when faced with the absence of semantic content. Composed using a structure of a classical sonata consisting of four movements and presented in text format as a score (complete with performance specifications, correct pronunciation of the letters and brief prescriptions regarding tempo, pitch, dynamics, and emotional content), *Die Ursonate* has long fascinated poets and musicians alike. The work, which makes explicit the bridge between music and literature both in title and in structural design, is an extraordinary example of the intricacies and entanglements that bind together music and spoken word.

This chapter examines how the intentional exploitation of the limen of speech and musical sound in Schwitters’s *Die Ursonate* (1922–1932) affords creative approaches to analysis. In this chapter I provide a brief introduction to Kurt Schwitters as an artist, as well as a general overview of the four movements of *Die Ursonate* (which Schwitters named *Introduction and Rondo*, *Largo*, *Scherzo*, and *Presto*, respectively), referring where appropriate to existing analyses and/or studies of the work. I then focus on the

---

<sup>146</sup> Filippo Tommaso Marinetti, *Selected Poems and Related Prose*, trans. Elizabeth R. Napier and Barbara R. Studholme (New Haven and London: Yale University Press, 2002), 77–80. For more on the “Technical Manifesto” and Marinetti’s language model, consider Marjorie Perloff, *The Futurist Moment: Avant-Garde, Avant Guerre, and the Language of Rupture* (Chicago: Chicago University Press, 2003), 56.

third movement, *scherzo*, and present an analysis that posits a deeper musical underpinning to the phonetic motives that Schwitters uses and their compositional design.

## 4.1 Kurt Schwitters

Although closely affiliated with the Dada movements in Switzerland, Germany, and the Netherlands, artist Kurt Schwitters (1887–1948) was not, strictly speaking, a card-carrying Dadaist himself. He belonged to—or rather, was the sole member of—the contemporaneous one-man movement he called *Merz* (the term is a chance fragment of ‘Kommerzbank’ [bank of commerce]), which held strongly to the principle of effacing the boundaries between the arts, and to the principle of material equality and metamorphosis through combinations. Schwitters explains:

For the sake of thrift, I took whatever I found, for we were an impoverished country. You can also shout with trash, and that is what I did, gluing and nailing it together. I called it *Merz*: it was my prayer for the victorious end to the war, for once more peace emerged victorious again. Everything was wrecked anyway, and new things had to be made from the fragments. That is *Merz*.<sup>147</sup>

Operating as an artist within a *Merz* philosophy means creating through choice, combination, distribution, and metamorphosis of materials. The foremost procedure employed in Schwitters’s works—and the driving aesthetic property of his work that aligns with the Dada movement—is the art of collage or, put more generally, assemblage.

---

<sup>147</sup> Heinz Rasch and Bodo Rasch, *Gefesselter Blick: 25 kurze Monografien und Beiträge über neue Werbegestaltung* (Stuttgart: Wissenschaftlicher Verlag Dr. Zaugg, 1930), reprinted in Kurt Schwitters, *Das literarische Werk*, edited by Friedhelm Lach (Munich: Deutsche Taschenbuch Verlag, 2005), Vol. 5: 335–336. Quoted from Megan R. Luke, *Kurt Schwitters: Space, Image, Exile* (Chicago: Chicago University Press, 2013), 5.

But where the Dadaists used collage to reflect the chaos and discord they saw in their society, often hyperbolically for shock effect, Schwitters thought of collage in a fundamentally different way. This distinction from the Dadaist approach runs deep in Schwitters's philosophy, as he explains, "whereas Dadaism merely poses antitheses, *Merz* reconciles antitheses by assigning relative value to every element in the work of art. Pure *Merz* is pure art; pure Dada is non-art; in both cases deliberately so."<sup>148</sup>

The Dadaists considered collage a profoundly modern technique because of its destructive aspects, while Schwitters considered it a profoundly versatile artistic method because of its constructive elements. Schwitters's collages acknowledge rupture and fragmentation as a given, but not as an artistic end: He believed in synthesis and unity through the joining of unrelated parts and, in this way, his artistic works—though begun under Dada influence in late 1918—shifted away from the Dadaist aesthetic in 1919 and 1920 as he developed his mature (and unique) artistic style.

Depending on the scholarly circles one keeps, Schwitters's artistic legacy and influence on modern art will differ. In artistic circles, Schwitters's particular style and innovative approach to collage works and high-relief works are among his most significant contributions. Schwitters trained as a painter and, while he continued to paint throughout his lifetime, it was what he created when he was *not* painting that had lasting influence on the artistic world. Gwendolyn Webster writes:

The language of *Merz* now finds common acceptance and today there is scarcely an artist working with materials other than paint who does not

---

<sup>148</sup> Kurt Schwitters, "Banalitäten," 1923, quoted in Friedhelm Lach, *Der Merz Künstler Kurt Schwitters* (Cologne: Verlag M. Dumont, 1971), 48.



refer to Schwitters in some way. In his bold and wide-ranging experiments he can be seen as the grandfather of Pop Art, Happenings, Concept Art, Fluxus, multi-media art, and post-modernism.<sup>149</sup>

In both literary and musical circles, however, Schwitters's influence on Happenings and Fluxus are considered most significant, particularly in relation to his innovative take on sound poetry. Although literary critics would likely cite Schwitters's 1919 poem, "An Anna Blume," as his most important literary work, music scholars quite naturally gravitate towards his epic sound poem, *Die Ursonate* (sometimes referred to as *lautsonaten* or *Die Sonate im Urlauten*).

## 4.2 *Die Ursonate* (1922–1932)

Composed over the decade from 1922–1932, *Die Ursonate* boasts not only an original, formal design but also an oft-told origin story. On a visit to the city of Prague, Kurt and Helma Schwitters attended an evening performance put on by several notable Dadaists of the day including artist, engineer, and phonetic poet Raoul Hausmann (1886–1971), who performed his now famous phonetic poem, *fmsbw*. The following day, the Schwitters and the Hausmanns went on a nature walk near the little town of Lobovice (in modern day Czech Republic) during which Schwitters expressed his fascination with Hausmann's piece by repeating the first line over and over—to Hausmann's great annoyance.<sup>150</sup> Over

---

<sup>149</sup> Gwendolyn Webster, "Artchive Online," *artchive.com* 2011. Accessed 25 July 2019.

<sup>150</sup> Nearly every source will tell this story; Hausmann himself recalls the events (though the accuracy of the whole account is at times questionable) in "29 Jahre Freundschaft mit Kurt Schwitters," unpublished manuscript, quoted in Werner Schmalenbach, *Kurt Schwitters* (Cologne: Dumont Schauberg, 1967), 222–23.

the next ten years, Schwitters expanded Hausmann's first line into the *Ursonate*, a 35 minute piece structured, as the name suggests, in a multi-movement sonata form.

The last edition of *Merz*, the avant-garde multimedia journal that appeared between 1923 and 1932 under Schwitters' editorial direction, was a complete performance score for *Die Ursonate*, with typography by Jan Tschichold. Alongside a detailed preface and substantial performance notes accompanying the score, Schwitters and Tschichold take pains to provide important annotations directly in the score, using a grid structure, line boundaries of differing thicknesses and styles, and an alpha-numeric labeling system to indicate theme units and larger formal sections.<sup>151</sup>

#### 4.2.1 Introduction and Rondo, 1<sup>st</sup> Movement

Following a more or less traditional multi-movement sonata form, the *Ursonate* begins with a sonata-rondo movement—complete with introductory and coda framing functions—that is both the longest movement of the work and the most sonically diverse, using twenty-seven phonemes (all of which occur in the movement's first theme). The amount of thematic material packed into the movement is remarkable. In the movement's introduction, Schwitters presents four distinct themes which collectively function as the expositional refrain of the movement (Figure 4.1).<sup>152</sup>

---

<sup>151</sup> For the interested reader, a complete and annotated edition of *Merz* journal was recently published in Kurt Schwitters, *Die Reihe Merz 1923–1932*, ed. Ursula Kocher and Isabel Schulz, *Die Reihe Merz 1923–1932* (Berlin/Boston: Walter de Gruyter GmbH, in cooperation with the Sprengel Museum Hannover, 2019).

<sup>152</sup> Facsimile of the 1932 print edition of *Die Ursonate* in *Merz 24* is available in Dary John Mizelle, *Kurt Schwitters's Ursonate and Quanta and Hymn to Matter* (PhD

**Figure 4.1 Annotated score of *Die Ursonate*, 1st movement, with colors indicating motivic segments of themes 1–4.**

theme 1:	1
Fümms bö wö tää zää Uu, pögiff, Kwii Ee.	
theme 2:	2
Dedesnn nn rrrrr, li Ee, mpiff tillff too, tillll, Jüü Kaa?	
theme 3:	3
Rinnzekete bee bee nnz krr müü? ziiuu ennze, ziiuu rinnzkrrmüü	
rakete bee bee,	3a
theme 4:	4
Rrumppff tillff toooo?	

The first of these four themes draws from Hausmann’s poem, “fmsbw.” It contains three motives, distinguished in this reproduction by commas.<sup>153</sup> Thus taking the four themes together, there are twelve distinct motives that make up the first movement alone. Some of those motives in their own right contain even smaller elements that transform into new thematic material (such as the “bee bee” in theme 3, which is central to the second movement of the sonata, the “ziiuu,” which transforms into motive 10 in the third movement, and the “Rrumppff” from theme 4, which occurs in modified form in the third movement).

#### 4.2.2 Largo, 2<sup>nd</sup> Movement

The abundance of thematic material in the first movement offers ample fodder for thematic interconnections between movements and the cyclic nature of the *Ursonate* is

---

Dissertation, University of California, 1977), 71–88. This is the first of many examples reproduced from Schwitters’s 1932 print edition of the full *Ursonate* in *Merz 24*.

<sup>153</sup> This is a simplification of Schwitters’s 1932 print edition, where commas, line breaks, and indentations indicate motivic distinctions.

immediately apparent in the second movement, *largo*. The simple and languid themes of this movement have already sounded in the introduction to the first movement.<sup>154</sup> In stark contrast to the first movement, the *largo* involves the smallest sonic profile of the work, using only twelve discrete phonemes. Schwitters complements this most minimal sonic profile (and most straightforward formal structure) with the most supplementary performance instructions. On the score, he provides both temporal and pitch specifications for the performer: “rhythmically performed. exactly 4/4 time. each following line is spoken successively a quarter tone lower, so the piece should begin proportionately.”<sup>155</sup> In the preface to the score, Schwitters also suggests that the performer may want to write out in musical notation the rhythms for the movement and provides an example of his idiosyncratic notation for the first four lines of the movement (Figure 4.2).<sup>156</sup>

**Figure 4.2 Example of Kurt Schwitters’s musical notations for lines 1–4 of *Die Ursonate*, 2nd movement.**

$\frac{4}{4}$	$\frac{Oo}{1}$	$\frac{bee\ bee}{4}$	$\frac{bee\ bee}{4}$	$\frac{bee}{8}$	$\frac{3}{8}$	$\frac{zee\ zee}{4}$	$\frac{zee\ zee}{4}$	$\frac{zee}{8}$	$\frac{3}{8}$	...
---------------	----------------	----------------------	----------------------	-----------------	---------------	----------------------	----------------------	-----------------	---------------	-----

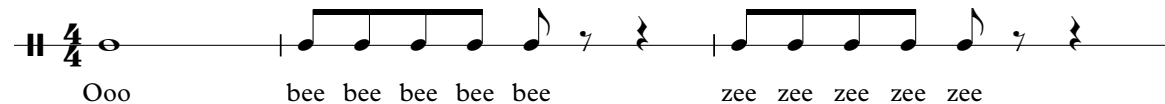
<sup>154</sup> The introduction presents both themes from the second movement (numbered 6 and 7), alongside one other theme that Schwitters counts as theme 5, which only returns once towards the end of the final movement.

<sup>155</sup> Schwitters, “erklärungen zu meiner ursonate,” reprinted in Mizelle, *Kurt Schwitters’s Ursonate and Quanta and Hymn to Matter*, 71. Original: »gleichmäßig vorzutragen. takt genau 4/4. jede folgende reihe ist um einen folgenden viertel ton tiefer zu sprechen, also maß entsprechend begonnen warden«. Translation mine. Lack of capitalization, Schwitters.

<sup>156</sup> Schwitters, “erklärungen zu meiner ursonate,” reprinted in Mizelle, 72.

Schwitters's notational system is somewhat odd but becomes a charming, if not mixed-metaphorical system for representing syllabic duration as rational proportions once properly deciphered (Figure 4.3).

**Figure 4.3 Musical interpretation of Schwitters's rhythmic notation for *Die Ursonate*, 2nd movement.**



Boxes delineate measures in 4/4 time and contain fractions that represent durations per syllable. The first measure, then, contains one syllable over the whole measure. The second measure contains two syllables, “*bee bee*,” over a fourth part of the measure, the same again, then one syllable, “*bee*,” over an 8<sup>th</sup> of the measure, followed by no syllables over the remaining 3/8ths of the measure.

#### 4.2.3 Scherzo and Trio, 3<sup>rd</sup> Movement

The third movement of the *Ursonate* is the *scherzo-trio* movement, and shares with the previous movement a cogent formal structure and quite charmingly emulates the formal design typical of a classical scherzo movement. The movement presents a new main thematic motive, III, which combines in various ways with three new numbered themes (8, 9 and 10) as well as motivic elements that, while phonetically reminiscent of themes 3a and 4 from the first movement, are remarkably distinctive in nearly every other way. Schwitters also provides three performance instructions that are important for understanding how the movement should unfold: first, one should perform the themes in distinctively different ways; second, that the *scherzo* should be lively; and finally, that the

*trio* should be very slow with lengthened phonemes. How these directives play out in relation to each other varies widely between performances of the movement and will be a subject of much discussion to come later in this chapter.

#### 4.2.4 Sonata (presto) with Cadenza, 4<sup>th</sup> Movement

The final movement not only presents seven new themes but also brings back the four themes from the first movement, interwoven with references to thematic material from the second and third movements as well. Schwitters's notes on this movement suggest that he was quite proud of the coda, which incorporates a final theme numbered 18.

Schwitters writes:

I call attention to the intentional retrograde of the alphabet to a. One expects that and anticipates the a with suspense. But twice it ends painfully on the bee. The bee sounds painful in this context. The calming resolution follows in the third alphabet to a. But now the alphabet follows to the conclusion a fourth and final time and ends very painfully at *beeee?*<sup>157</sup>

Indeed, this passage is an elegant example of how even the most abstract linguistic materials can emulate familiar (classical) musical formulae (Figure 4.4). What Schwitters describes—the retrograde alphabet that stops short of its final goal not once but twice before finally resolving—should be familiar to music theorists as cadential evasion and the “one-more-time” technique.<sup>158</sup>

---

<sup>157</sup> Schwitters, “erklärungen zu meiner ursonate,” reprinted in Mizelle, 71–72.

<sup>158</sup> Beethoven, for example, employs these very same devices in his Piano Sonata in A-flat, Op. 26, *iv*, mm. 148–154.

**Figure 4.4 Kurt Schwitters, Die Ursonate, 4th movement, coda.**

finale:	
Zärr üpsiilon iks ( <i>animated</i> )	18
Wee fau Uu	
Tee äss ärr kuu	
Pee Oo änn ämm	
Ell kaa Ii haa	
Gee äff Ee dee zee beee?	
Zärr üpsiilon iks ( <i>more animated</i> )	18
Wee fau Uu	
Tee äss ärr kuu	
Pee Oo änn ämm	
Ell kaa Ii haa	
Gee äff Ee dee zee beee?	
Zärr üpsiilon iks ( <i>simply</i> )	18a
Wee fau Uu	
Tee äss ärr kuu	
Pee Oo änn ämm	
Ell kaa Ii haa	
Gee äff Ee dee zee beee Aaaaa	
Zärr üpsiilon iks ( <i>very animated</i> )	18
Wee fau Uu	
Tee äss ärr kuu	
Pee Oo änn ämm	
Ell kaa Ii haa	
Gee äff Ee dee zee beee? ( <i>dolefully</i> )	

By presenting the German alphabet backwards, like a counting down towards *a*, Schwitters emulates the descending musical scale. The omission of the final syllable “Aaaa,” acts as a kind of abandoned cadence, motivating the repetition of the cadential progression and, as a result, extending cadential function one (or two or three) more times.

Numerous passages like this one permeate the *Ursonate*, suggesting that Schwitters was not only familiar with the techniques of musical development, but strove intentionally to replicate them through speech sounds as a compositional medium. Such passages no doubt play a central role in motivating musical analyses of the work—at least, parts of it.

As is so often the case with multi-movement sonata forms, scholarly inquiry tends to focus primarily on the longer, outer movements of *Die Ursonate*, with a sentence or two (at best) mentioning the inner two movements.<sup>159</sup> Leigh Van Handel focuses exclusively on the first movement in her form-functional examination of *Die Ursonate*, as does Jack Ox's visual analysis of the work, while Sabine Lichtenstein's critique of *Die Ursonate's* aesthetic merits, and Peter Dayan's beautifully sensitive analysis of the work's alphabetic foundation, both focus primarily on the last movement of the work.<sup>160</sup>

Despite scholarly focus on the outer movements, Schwitters was often focused on the *scherzo* movement. In a letter to Hans Richter on the subject of making a movie of the *Ursonate*, Schwitters wrote, "you know the sonata is printed. But it is not all good. Only ½ of it is very good."<sup>161</sup> While Schwitters does not go on to specify *which* half of the sonata he considers the good half, the *scherzo* appeared in print before all other movements and more times than any other; and although Schwitters was never able to make a complete recording of *Die Ursonate* as he had hoped, in the two opportunities he did have to record segments of the work, he chose both times to record the *scherzo*. The

---

<sup>159</sup> A notable exception to this is Mizelle, *Kurt Schwitters's Ursonate and Quanta and Hymn to Matter*, which presents a complete phonemic analysis of each movement of *Die Ursonate*.

<sup>160</sup> Ox, "Creating a Visual Translation of Kurt Schwitters's *Ursonate*"; Peter Dayan, "The Music of Dada Follows Kurt Schwitters to Holland (Bypassing Berlin) and Ends in the *Ursonate*," in *The Music of Dada: A Lesson in Intermediality for Our Times* (New York: Routledge, 2019), 129–56; Lichtenstein, "Listening to Kurt Schwitters's *Ursonate*"; Leigh Vanhandel, "The Classic(AI) Dada Work: Kurt Schwitters's *Ursonate*," presented at the West Coast Conference of Music Theory and Analysis (WCCMTA), April 26-28, 2017, Victoria, BC. Handout available at <https://msu.edu/~lvh/NewFiles/Schwittershandout.pdf>.

<sup>161</sup> Quoted in Schmalenbach, *Kurt Schwitters*, 372, fn. 54a.



following analysis posits the *scherzo* to be unique among the various parts that make up *Die Ursonate* and treats the movement as not only a “genuine scherzo,” but as a genuine collage in Schwitters’s own style.

## 4.3 Die Scherzo (und trio)

### 4.3.1 Musical Form

The form of the inner movements from *Die Ursonate* are relatively straightforward compared to the work’s first and final movements. The third movement, *scherzo and trio*, is written in a traditional da capo trio form, with the *scherzo* forming the large A section and da capo return. The *scherzo* comprises six distinct motives that Schwitters labels as III (lanke trr gll), 8 (pe pe pe pe pe /Ooka ooka ooka ooka), 9 (pii pii pii pii pii /Züüka züüka züüka züüka), 4 (Rrmp Rrnf/Rrmpff tilff too), 3 (Ziiuu lenn trll?), and 10 (Lümpff tümpff trll).<sup>162</sup> While some of these motives are unique to the third movement of *Die Ursonate*, others form connections with phonetic material from earlier movements. Motive 4 harkens back to a motive from the first movement and is presented in the *Scherzo* in two different forms: 4a, “Rrmp Rrnf,” and 4b, “Rrmpff tilff too.” Similarly, “ziiuu lenn trll” is phonetically related to “ziiuu ennze rinnzkrümü,” theme 3 in the first movement. The *scherzo* follows a standard small ternary (or rounded binary)

---

<sup>162</sup> The reader should take note of the distinction between thematic motives III (lanke trr gll) and 3 (ziiuu lenn trll?), as they may be a source of confusion later in my discussion of the work. Schwitters uses Arabic numerals to label themes and motives throughout *Die Ursonate* save for the first motive of the scherzo, which he gives the roman numeral III. He offers no explanation for this momentary difference.

structure, made easily identifiable by the motivic labels Schwitters provides in the 1932 edition of the work (Figure 4.5).<sup>163</sup>

---

<sup>163</sup> A reproduction of the 1932 print edition of *Die Ursonate* in *Merz* No. 24 is available in Mizelle, *Kurt Schwitters's Ursonate and Quanta and Hymn to Matter*, 71–88. The score for the scherzo can be found in Mizelle, 81. The complete text of Kurt Schwitters's *Ursonate* may also be found (and downloaded) at: <http://www.abaforum.es/users/2010/110.htm#so>.

**Figure 4.5 Annotated score of the scherzo from *Die Ursonate*, 3<sup>rd</sup> movement, indicating the formal structure.**

<b>dritter teil :</b>	<b>Formal Structure</b>
<b>scherzo</b> ( <i>die themen sind charakteristisch verschieden vorzutragen</i> )	
Lanke trr gll ( <i>munter</i> )	(M) III <i>Antecedent</i>
Pe pe pe pe pe	8
Ooka ooka ooka ooka	
Lanke trr gll	III <i>Consequent</i>
Pii pii pii pii pii	9
Züüka züüka züüka züüka	
Lanke trr gll	III <i>Cadential</i>
Rrmmp	4
Rrnnf	
Lanke trr gll	III <i>Digression</i>
Ziiuu lenn trll?	3
Lümpff tümpff trll	10
Lanke trr gll	III
Rrumpff tilff too	4
Lanke trr gll	III
Ziiuu lenn trll?	3
Lümpff tümpff trll	10
Lanke trr gll	III <i>A' (reprise)</i>
Pe pe pe pe pe	8 <i>Antecedent</i>
Ooka ooka ooka ooka	
Lanke trr gll	III <i>Consequent</i>
Pii pii pii pii pii	9
Züüka züüka züüka züüka	
Lanke trr gll	III <i>Cadential</i>
Rrmmp	4
Rrnff	
Lanke trr gll	<i>(codetta)</i>

The opening A section comprises motives III (lanke trr gll), 8 (pe pe pe pe pe /Ooka ooka ooka ooka), and 9 (pii pii pii pii pii /Züüka züüka züüka züüka), which together form a period-like phrase structure. The A section concludes with motives III and 4a fused together to create a cadential-like gesture. While motive III continues to permeate the work in the subsequent section, new motives enter into the movement that contrast from

the tight-knit structure of the opening theme. In the B section's digression (which itself establishes an internal aba structure), motive III is juxtaposed by motives 4b, "Rrmpff tilff too" and 10, "Lümpff tümpff trll," which provide new material to the passage and, in the case of motive 10, to *Die Ursonate* as a whole. The closing A<sup>i</sup> section is a literal repeat of the opening theme with an additional codetta featuring a final statement of motive III. The small ternary *scherzo* is itself nested within a larger quasi compound ternary form that includes a *trio* which borrows all motivic elements from theme 3 of the first movement and themes 6 and 7 from the second movement (Figure 4.6).

**Figure 4.6 Annotated print format showing formal structure of the trio from *Die Ursonate*, 3<sup>rd</sup> movement.**

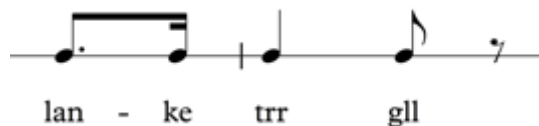
Scherzo	A Section (small ternary)	
trio:		Contrasting Middle
Ziiu iuuu ziiuu aaau ziiuu iuuu ziiuu Aaa	3    (3+7)	a
Ziiu iuuu ziiuu aaau ziiuu iuuu ziiuu Ooo	3    (3+6)	a'
Ziiu iuuu ziiuu aaau ziiuu iuuu	3	a (retransition)
da capo Scherzo	A Section (small ternary)	

Given the derivation of the *trio* material from the first and second movements, the following analysis pertains only to the *scherzo* portion of this movement, as the central focus of analysis is on determining where the motives that make up the *scherzo* movement come from.

### 4.3.2 Motivic Fragments

Motive III, “lanke trr gll”—the only motive in the entire sonata not given an Arabic numeral— occurs at the start of each line in the Scherzo, sounding a total of twenty times over the course of the movement. Ernst Schwitters’s 1954 recording is surprisingly consistent in his rhythmic delivery of this motive, both in terms of syllabic emphasis and durational values (Figure 4.7).<sup>164</sup>

**Figure 4.7 Notated rhythm of motivic fragment III.**



The ubiquity of the motive over the course of the *Scherzo*, together with its rhythmic (and melodic) profile, recalls the equally incessant motivic cell that Beethoven employs almost exclusively in the scherzo movement of his Piano Sonata No. 29, Op. 106. Indeed, where Schwitters writes of “lanke trr gll” that in its persistent return it is an “unchanging and stubborn rhythm” one could say very same of the motivic cell that drives Beethoven’s scherzo movement (Figure 4.8).<sup>165</sup>

<sup>164</sup> A complete recording of *Die Ursonate* performed by Ernst Schwitters is available at <http://www.ubu.com/sound/schwitters.html>. My rhythmic transcription differs from that put forth by others who interpret the syllable “trr” as slightly shorter than a quarter note and the duration of “gll” as slightly longer than an 8<sup>th</sup> note. See Mizelle, 17. In either case there is a degree of participatory discrepancy, not to mention the ambiguous and malleable threshold for syllable onsets, which can change one’s perception of syllable timing.

<sup>165</sup> Schwitters, “erklärungen zu meiner ursonate,” reprinted in Mizelle, 71.

Figure 4.8 Beethoven Piano Sonata No. 29, Op. 106, ii, mm. 1–14.

The image shows the musical score for the second movement of Beethoven's Piano Sonata No. 29, Op. 106, in B-flat major, 3/4 time, marked 'Scherzo Allegro vivace'. The score consists of two systems of music, each with a treble and bass clef staff. The first system (measures 1-7) begins with a piano (*p*) dynamic and features a central rhythmic motive. The second system (measures 8-14) continues the piece, marked with a crescendo (*cresc.*) and a forte (*f*) dynamic. The score includes various musical notations such as slurs, accents, and dynamic markings.

While the connections between this musical motive and Schwitters’s opening motive of the *schерzo* might, at first, seem little more than coincidentally similar, the coincidence is, at least, relatively rare. In a survey of over 150 scherzo movements written between 1791–1905 by 25 different composers, not a single other scherzo employed so similar a musical shape and rhythm.<sup>166</sup> In this way, Schwitters’s *schерzo* shares another similarity with Beethoven’s scherzo movement: the central motive from each is distinctly different from all other motives around. This is not to suggest that the resemblance is necessarily indicative of derivation, but rather to suggest that upon *hearing* Schwitters’s motive, “lanke trr gll,” the most likely musical scherzo to come to mind would be, on the balance

<sup>166</sup> Although this investigation was limited to works for piano solo, I surveyed as many works as possible, including multiple pieces by both well-known and lesser known composers, namely, Pyotr Tchaikovsky, Edouard Frank, Emil von Sauer, Edouard Wolff, Lucien Vieuxtemps, Wilhelm Taubert, Julius Reubke, Alexander Scriabin, Robert Schumann, Clara Schumann, Franz Schubert, Xaver Scharwenka, Joachim Raff, Felix Mendelssohn, Franz Liszt, Rudolphe Kreutzer, Theodor Kirchner, Olaf Emil Erikson, Frederic Chopin, Johannes Brahms, and Ludwig van Beethoven, among others.

of probabilities, is that from Beethoven's Piano Sonata No. 29, Op. 106. Yet, while Beethoven uses this motive repetitively to form the entirety of the main theme, Schwitters immediately juxtaposes "lanke trr gll" with contrasting motivic fragments that bear no relationship to the Op. 106 scherzo. Instead, the subsequent motivic fragment, "pe pe pe pe pe" is recited as five equal durations in quick succession, of which the final "pe" is durationally emphasized by a following silence (Figure 4.9).

**Figure 4.9 Notated rhythms of motivic fragments 8a and 9a.**



The motivic fragment "pe pe pe pe pe" (and the closely related "pii pii pii pii pii") recall an entirely different scherzo, sounding both rhythmically and melodically similar to the characteristic opening motive from Chopin's Scherzo No. 3, op. 39 (Figure 4.10).

**Figure 4.10 Chopin Scherzo No. 3, op. 39, mm. 1–4.**

**Presto con fuoco**

And distinct yet again are the motivic fragments "ooka ooka ooka ooka" and (related) "züüka züüka züüka züüka." While these fragments from motives 8 and 9, respectively, bear some resemblance to each other in their syllabic content and formation, Schwitters recites them in such a way as to emphasize their differences more than their similarities.

The “ooka” fragment is shorter and give prominence to the first of the two syllables, while the “züüka” fragment involves longer durations on both syllables and a smoother connection between adjacent syllables (Figure 4.11).

**Figure 4.11** Notated rhythms of motivic fragments 8b and 9b from *Die Ursonate*, 3<sup>rd</sup> movement.

The figure displays two musical staves. The top staff contains four rhythmic units, each labeled 'ooka'. Each unit consists of two quarter notes followed by a quarter rest. The bottom staff contains four rhythmic units, each labeled 'züüka'. Each unit consists of two dotted half notes followed by a quarter rest.

Together, these motivic fragments recall certain aspects of the main motivic material of yet another scherzo composed by Beethoven (Figure 4.12).

**Figure 4.12** Beethoven Piano Sonata No. 15, Op. 28, iii, mm.1–16.

The figure shows a musical score for the first eight measures of the third movement of Beethoven's Piano Sonata No. 15, Op. 28. The tempo is marked 'Allegro vivace' and the dynamics are 'p' (piano). The score is in 3/4 time and features a melody in the right hand and a harmonic accompaniment in the left hand.

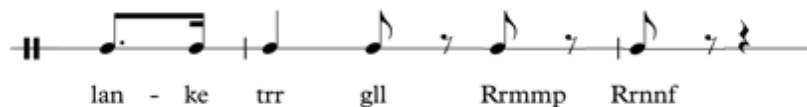
The order of the motivic fragments as Schwitters presents them syllabically differs from the motivic trajectory of the scherzo from Beethoven’s Piano Sonata No. 25, Op. 28.

Schwitters’s fragment 8b, “ooka,” recalls not the first measures of the scherzo for piano, but the second in mm. 5–8 and mm. 13–16. Schwitters’s fragment 9b, “züüka,” takes on



the sonic qualities of mm. 1–4 and mm. 9–12, as though mm. 5–8 have been cut out for use elsewhere and the two dotted half-note motives have been sewn together. The A section closes with motive III followed immediately by a statement of motive 4a, “Rrmmp Rrnnf.” While phonetically related to motive 4 from the first movement of the *Ursonate*, this motive 4 is not, as Schwitters is careful to indicate in the accompanying notes to the score, intended to receive the same performative delivery as in previous iterations.<sup>167</sup> While Schwitters borrows phonetic material from the original motive 4 (Rrumpff tillff too) but alters it substantially enough both phonetically and in performance that we can think of the first occurrence of the motive as 4a, a variant of motive 4 related only conceptually. In his performance notes, Schwitters instructs that motive 4a be performed as an extension of motive III (Figure 4.13).

**Figure 4.13** Notated rhythms of motivic fragments III and 4a from *Die Ursonate*, 3<sup>rd</sup> movement.



The fused phonetic utterance “lanke tr gll rrmmp rrrnnf” occurs twice in the movement, both times to conclude the A section theme. The utterance is eerily similar to the

<sup>167</sup> Schwitters, “erklärungen zu meiner ursonate,” reprinted in Mizelle, *Kurt Schwitters’s Ursonate and Quanta and Hymn to Matter*, 71. Schwitters writes “in *rrmmmp* and *rrrnnff* there is a reminder of the *rummpff tillff too* in the first movement. But it no longer sounds lamblike and tender, but short and commanding, quite masculine” (»in *rrmmmp* und *rrrnnf* ist eine erinnerung an das *rummpff tillff too* zom ersten satz. doch klingt es jetzt nich mehr lammhaft zart, sondern kurz und befehlend, durchaus männlich«).

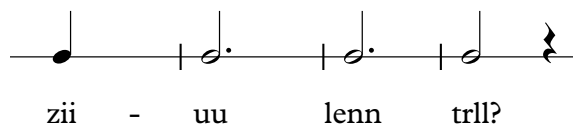
cadential material Beethoven employs at the end of the compound basic idea in m. 5–7 and to close the first theme of his Op. 106 scherzo in mm. 12–14 (Figure 4.14).

**Figure 4.14 Beethoven Piano Sonata No. 29, Op. 106, ii, mm. 5–7 and mm. 12–14.**



The new material Schwitters brings into the B section of the *Scherzo* departs from the musical referents called to mind by the opening period theme. Here, the phonetic motives seem to allude to a second scherzo by Chopin, Op. 54, interspersed with statements of the persistently ubiquitous motive III. Chopin’s fourth and final scherzo comes to mind with “ziiuu lenn trll?” as Schwitters delivers the line in such a way as to render the initial syllable, “zii,” anacrustic in nature. He sustains the durations of syllables “uu” and “lenn” while placing prominence on syllables “uu” and “trll?” (Figure 4.15).

**Figure 4.15 Notated rhythm of motivic fragment 3, “ziiuu lenn trll?” from *Die Ursonate*, 3<sup>rd</sup> movement.**

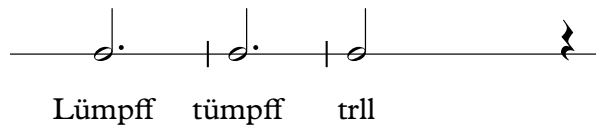


The durations presented for “ziiuu lenn trll?” in example 4.15 are meant to suggest the rhythmic relationship between each syllable in the utterance, and could have been written in different durational note values. The point is that the first syllable is approximately one-third the length of the subsequent syllables which, in turn, are approximately equivalent in length (including the final rest following “trll?”. The choice to represent the motive according to these particular durational note values is, in large part, deliberate, as it highlights the resemblance between the “ziiuu lenn trll?” motive and a recurrent cadential motive in Chopin’s Scherzo No. 4, Op. 54, mm. 22–25 (Figure 4.16).

**Figure 4.16 Chopin Scherzo No. 4, Op. 54, mm.1–25.**

Perhaps in no small part because Chopin’s fourth scherzo is presently on the mind, motive 10, “Lümpff tümpff trll” may also bear some resemblance to yet another fragment from the same piece, albeit one far less prominent in the work. The phonetic motive as Schwitters utters it is rhythmically simple, comprised of three long durations of roughly equivalent value (Figure 4.17).

**Figure 4.17** Notated rhythm of Motive 10, “Lümpff tümpff trll” from *Die Ursonate*, 3<sup>rd</sup> movement.



While such a generic rhythm might be related to or derived from countless musical referents, additional similarities of ascending melodic contour reinforce the relationship between Schwitters’s delivery and perhaps obscure passages in Chopin’s Scherzo Op. 54 Scherzo (Figure 4.18).

**Figure 4.18** Chopin Scherzo No. 4, Op. 54, mm. 197–201.

Just as the melodic fragment in mm. 199-201 ascend chromatically through sustained and equivalent durations, so too does Schwitters perform motive 10 in sustained durations with chromatically ascending pitch. Although the last duration in m. 201 is a quarter-note, and the piece moves on quickly to rapid arpeggiated quarter-notes, the additional pedal marking causes the F# major sonority to continue to ring out, rendering it rhythmically similar to that of Schwitters’s delivery of motive 10.

The final motive to consider is the central motive of the *Scherzo*, “Rrumpff tillff too,” which in the context of this movement I have called motive 4b, so as to differentiate it from its variant, “rrmmp rrnf.” This motive stands apart from the others, being the only motive to occur only one time in the piece and the only motive to correspond directly and precisely to motivic material existing elsewhere in the *Ursonate*. For this reason, it is perhaps the most interesting motivic moment, in that it is not only phonetically identical to motive 4 as it appears in the first movement of the *Ursonate*, but also delivered in the same way rhythmically and melodically.

In Schwitters’s recording of the first movement, motive 4 first occurs approximately one minute into the performance and Schwitters delivers the utterance in three sustained durations, with each syllable leaping upwards (Figure 4.19).<sup>168</sup>

**Figure 4.19 Notated pitch inflections for motive 4, “Rrumpff tillff too” from *Die Ursonate*, 3<sup>rd</sup> movement, as recited by E. Schwitters 1954.**



<sup>168</sup> A complete recording of *Die Ursonate* performed by Ernst Schwitters is available at <http://www.ubu.com/sound/schwitters.html>. Motive 4 occurs at 1:04-1:06 in the first movement and at 0:27-0:29 in the third movement. Pitch identification was established by ear and verified by Mauch and Dixon’s pYIN plugin for Sonic Visualiser, a probabilistic variant of the YIN algorithm for measuring fundamental frequencies in recorded speech (see fn.32).

When the motive returns in the *Scherzo*, Schwitters not only uses the same rhythmic durations, but the very same melodic contour and, indeed, nearly identical pitches (save for the semitonal distinction between B and B $\flat$  on the final syllable). In the context of my musical fragment collage analysis, it is as though amidst the various musical referents Schwitters inserts a referent to the first movement of the *Ursonate*, the only fragment directly “cut out” of existing phonetic material and pasted into the *scherzo* movement.

### 4.3.3 A Visual Collage Score

Identifying these familiar musical works in E. Schwitters’s recitation of the *Scherzo* and piecing them together involves processes not dissimilar from that which Schwitters himself employs in the vast majority of his creative work: collage and assemblage. By cutting out the various musical incipits suggested above from their respective places in their full scores and arranging them in such a fashion as to emphasize the formal design of the *Scherzo*, we can emulate in musical notation the musical trajectory of the work (Figure 4.20).

From a visual standpoint, the use of colour differentiates the four musical pieces from each other so as to demonstrate how the various motives move in and between the distinct and contrasting musical works. Such an endeavor might produce an interesting visual result, in which the colored score fragments juxtapose and interweave with one another, while the single referent to motive 4, rendered in text, finds its place in a central position.

Figure 4.20 Visual representation of musical collage allusions in *Scherzo*, overlaying digital fragments of paper materials used in the 1922 *Merz* collage, *Er* (He), by Kurt Schwitters with score fragments from Beethoven Piano Sonata No. 15, Op. 28, iii and No. 29, Op. 106, ii; and Chopin Scherzo No. 3, op. 39 and No. 4, Op. 54.



Although this particular design is perhaps only one way to envision a score for my musical interpretation of the *Scherzo*, it highlights a compelling relationship between this unique movement of the *Ursonate* and one of Schwitters's unique collages created around the same time as the artist toured performing the movement: the *Merz* collage called *Er*, which Schwitters created in 1922. *Er* (He) demonstrates both the development of Schwitters's method of assemblage and the growing influence of constructivist thinking on his work and philosophy—in particular, *de Stijl* influence through his association with Nelly (or, by her stage name, Petro) and Theo Van Doesburg.

In its prime in the early 1920s, *de Stijl*, which literally means “the style” developed concurrently with (and in close geographical proximity to) Dadaism. But where Dadaists gravitated towards chaos and irrationality in their artwork, artists associated with *de Stijl* favored abstraction and universality, emphasizing form and color in their work through the use of clean vertical and horizontal lines as well as black, white and primary colors almost exclusively. From this simple description of the style, even the most casual of onlookers would ascertain the influence of Van Doesburg and *De Stijl* on Schwitters's aesthetic development.

In regard to *Er* (He), Elderfield corroborates, writing that the “limitation not only to horizontal and vertical forms but almost entirely to primary colors ... derives from the vocabulary of De Stijl painting.”<sup>169</sup> While the Van Doesburgs played a key role

---

<sup>169</sup> John Elderfield, *Kurt Schwitters* (London: Thames and Hudson, 1985), 127. *Er* (He) by Kurt Schwitters (1922), fully titled ‘Mz’ “er” (so as to emulate in the very name of the work the fragmented extract of “er” not only within the work but from within *Merz* itself), presently resides in the provenance of an unknown American owner, and was last



Schwitters's artistic and philosophic development at this time, Schwitters maintains a distinctive element of collage in these works by constructing them out of independent, somewhat stained, and slightly irregular hand-cut pieces that he cut and pasted together. Similarly, the inclusion of fragmented and partially legible texts amidst the geometric blocks of color is uniquely Schwitters's developing style (Figure 4.21).

Schwitters's visual collages provide a great deal of insight into his attitude towards innovation and tradition in all of his artistic works, the *Ursonate* included. Particularly in his works from the early 1920s, Schwitters's literary work closely interweaves with this visual work, as scholar Dorathea Dietrich argues:

...both exist in a symbiotic relationship in which techniques and subject matter are developed in one genre and carried over into the other. His new literary language consisted of textual collages of found and invented parts; his new visual language, of ready-made forms, which he substituted for conventional drawing or painting. Sharing the collage process, the literary and visual works delineate a common approach to the modern, but in playing off the new against the established, they eloquently tell of a struggle, the difficulty of establishing boundaries for old and new and of defining their exact relationship.<sup>170</sup>

In order to approach the *Ursonate* with this interplay in mind—to search for a way to understand the work without lamenting its loss or failures but rather celebrates its newness and innovation—collage ought be at the heart of one's mind.

---

on exhibit in Chicago at the Museum of Contemporary Art in the Spring of 2002. The piece appears as an illustrated plate in Elderfield, 145.

<sup>170</sup> Dorathea Dietrich, *The Collages of Kurt Schwitters* (Cambridge: Cambridge University Press, 1993), 71.

Figure 4.21 *Er (He)* by Kurt Schwitters, 1922.<sup>171</sup>



<sup>171</sup> *Er (He)* by Kurt Schwitters (1922), fully titled 'Mz' "er" (so as to emulate in the very name of the work the fragmented extract of "er" not only within the work but from within *Merz* itself), presently resides in the provenance of an unknown American owner, and was last on exhibit in Chicago at the Museum of Contemporary Art in the Spring of 2002. The piece appears as an illustrated plate in Elderfield, *Kurt Schwitters*, 145.

Schwitters's ubiquitous use of collage, assemblage, and found objects permeates all the artistic forms in which he worked, in no small part because these techniques capture the key tenets of his artistic *Merz* philosophy. As early as 1919 he writes:

Essentially the word *Merz* means the assemblage for artistic purposes of all imaginable materials; technically, as a matter of principle, it means that each of these materials has the same value... It does not matter whether the materials used were already designed for some other purpose or not. The artist creates by choosing, arranging, and deforming the materials. The deforming of the materials may already come about through their arrangement on the picture surface. It may be further emphasized by breaking them up, twisting them, covering them over, or painting them over.<sup>172</sup>

This is, then, the hypothesis that drives my interpretation of the *Scherzo* from Kurt Schwitters's *Ursonate*: that the materials from which it derives are themselves found within other scherzos, broken up, twisted, covered over and refashioned into linguistic sound. Of course, to propose this involves entering into the imaginary possibilities of sound and sound meaning, an uncomfortable space to occupy for the truth-seeking music theorist. But, as Martyn Hudson so eloquently reminds us, Schwitters's art constantly calls us to precisely such a space:

The *Ursonate* does not provide a home for language or meaning or dwelling but rather a way of thinking about movements, boundaries and intersections on an imaginary plane. Each journey with and within the interior of the assemblage of the *Ursonate*, each articulation of its elements and each traverse of the whole piece by the listener is situated in the specific site of the imaginary. In this sense it is an imaginary

---

<sup>172</sup> Quoted in Dachy, *The Dada Movement, 1915-1923*, 116.

world–building, a fiction full of magical precipices in which, for the duration of the experience, the quotidian world is suspended.<sup>173</sup>

Thinking of Schwitters’s *Scherzo* as a collage of musical fragments imitative of preexisting scherzos by Beethoven and Chopin is a kind of “imaginary world–building,” where a visual score-collage presented in the style of one of Schwitters’s own visual collages intentionally serves to enhance the immersion or enchantment of that world. But a visual score-collage, however artistically rendered, cannot fully capture the musical world summoned by Schwitters’s *Scherzo*. An audio collage, on the other hand, could sway even the most skeptical.

#### 4.3.4 An Audio Collage

The following examples take Ernst Schwitters’s 1954 recording and overlay each motivic fragment with fragments from recordings of each proposed music work. Beethoven’s Piano Sonata No. 29 in B♭ Major, Op. 106, *ii*, is represented by a recording of the 1970 CBC broadcast (CBC Studios, Toronto, Canada) of Glenn Gould performing the work.<sup>174</sup> Similarly, Gould also performs Beethoven’s Piano Sonata No. 15 in D Major, Op. 28, *iii*, and the recording used here is the remastered 1980 record produced by Sony Music Entertainment. Both Chopin Scherzi recordings are of pianist Yulianna Avdeeva and are

---

<sup>173</sup> Martyn Hudson, “Schwitters’s Ursonate and the Merz Barn Wall,” *Leonardo Music Journal* 25 (2015): 91.

<sup>174</sup> Glenn Gould and Ludwig van Beethoven, “Piano sonata no. 24 in f-sharp major, op. 78: ‘A Therese’; Piano sonata no. 29 in B-flat major, op. 106: ‘Hammerklavier’ / Ludwig van Beethoven” [Sound recording] New York, NY: Sony Classical, 1993. SMK 52645. Track 5, Piano Sonata No. 29 in B-Flat Major, Op. 106 “Hammerklavier”: II. Scherzo. Assai vivace (CBC Radio Broadcast, CBC Studios, Toronto, Canada, November/December 1970).

her prize-winning performances in the 16<sup>th</sup> International Chopin Piano Competition in 2010. Avdeeva performed Scherzo in E major, Op. 54 in the competition's First Stage, and Scherzo in C# minor, Op. 39 in the Second Stage. Both recordings were made in collaboration by The Fryderyk Chopin Institute (NIFC), Polish Television (TVP), National Audiovisual Institute (NInA).<sup>175</sup>

While the recording of Schwitters's performance of the *scherzo* is presented exactly as he performed it, the musical fragments from the piano recordings are somewhat altered in two domains: tempo and amplitude. Tempo adjustments occur largely because Schwitters's recitation of these motives is slower than most performances of the musical pieces.<sup>176</sup> Gould's performance of Beethoven's Piano Sonata No. 29 in B $\flat$  Major, Op. 106, *ii*, for example, necessitated an approximate 20% reduction in tempo to match Schwitters's performance of the phonetic fragment "lanke trr gll." While mm. 1-4 and 9-12 in Gould's performance of Beethoven's Piano Sonata No. 15 in D Major, Op. 28, *iii*, are presented as performed, mm. 5-8 required an approximate 20% reduction in tempo. In Avdeeva's performance of Chopin's Scherzo in C# minor, Op. 39, the musical fragments were also slowed slightly by approximately 15%, while her performance of

---

<sup>175</sup> Yulianna Avdeeva, first prize winner First stage Scherzo in E major, Op. 54 All rights reserved 2010 The Fryderyk Chopin Institute (NIFC), Polish Television (TVP), National Audiovisual Institute (NInA). <https://www.youtube.com/watch?v=ORwDDsn2jTU>. Second stage Scherzo in C sharp minor, Op. 39 All rights reserved 2010 The Fryderyk Chopin Institute (NIFC), Polish Television (TVP), National Audiovisual Institute (NInA). <https://www.youtube.com/watch?v=RGyfroxys4I>.

<sup>176</sup> Both tempo and amplitude adjustments, as well as fragment extraction, were executed through Audacity, a free multi-track audio editor and recorder distributed under the GNU General Public License.

Chopin's Scherzo in E major, Op. 54 was reduced in tempo by approximately 5%.

Amplitude adjustments occur largely for passages that were relatively quiet in the piano recordings as compared to the analogous passage in Schwitters's performance. Avdeeva's performance of Chopin's Scherzo in C# minor, Op. 39 necessitated an increase in amplitude by approximately 4 dB, as did Gould's performance of Beethoven's Piano Sonata No. 15 in D Major, Op. 28, *iii*.

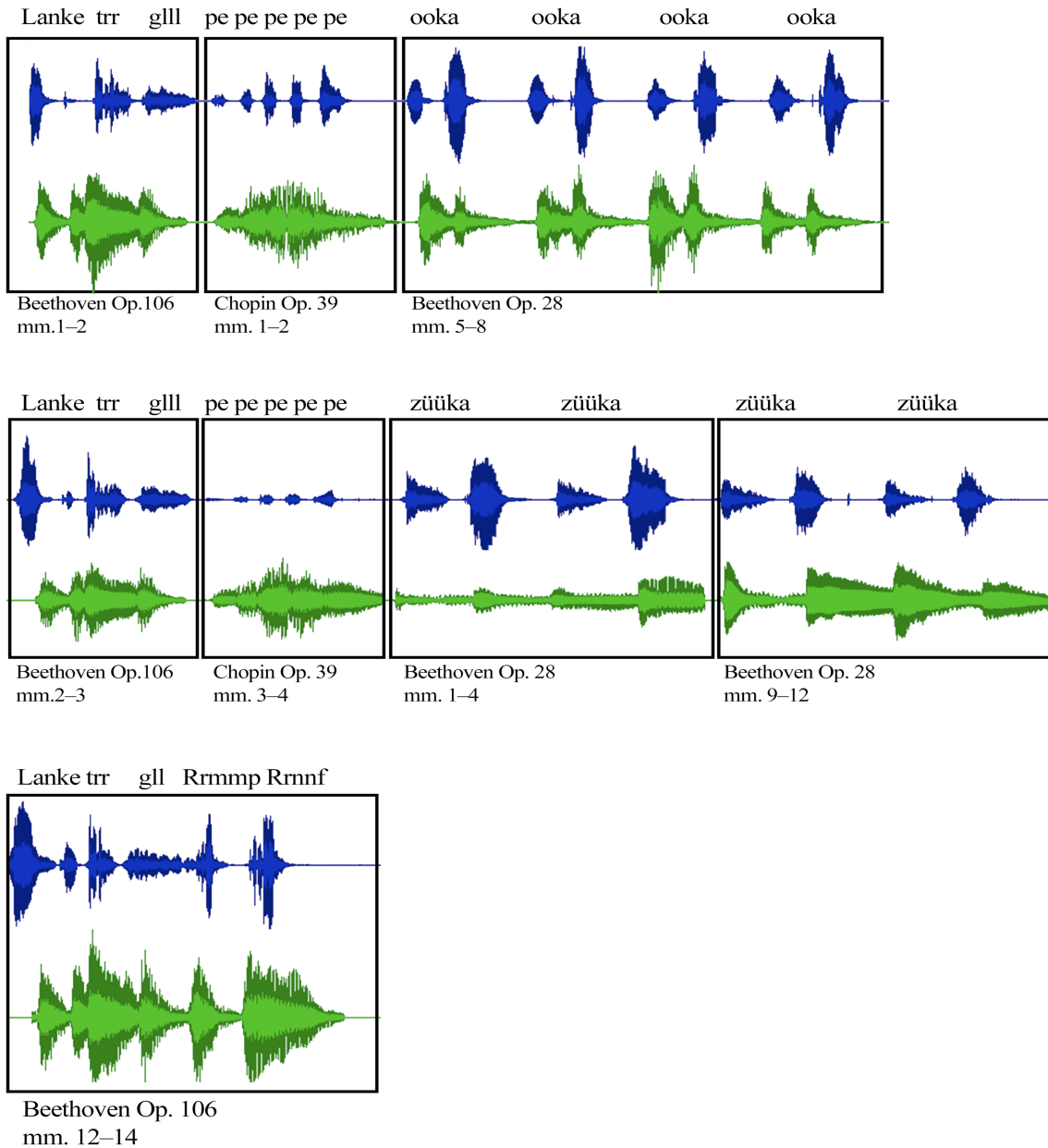
With these adjustments, the audio for each musical fragment could be laid over the recording of Schwitters's performance and compared phrase by phrase. While nothing is exactly the same, there are strong relationships between Schwitters's performance and the piano recording collage of the scherzo's A section (Figure 4.22 and [Audio 4.1](#)).<sup>177</sup>

The overarching similarities should, I hope, be evident both aurally and visually. Event onsets, for example, correspond to waves of amplitude in Figure 4.22 and for each motivic event, such waves occur in relatively analogous positions. A similarity not readily apparent in Figure 4.22—but certainly perceptible in Audio 4.1—are the melodic contour similarities between Schwitters's pitch inflections across the various phonetic motives and the melodic contours of the musical fragments. Similarities aside, there remain several notable differences between Schwitters's performance and the piano recording collage that merit discussion.

---

<sup>177</sup> Audio 4.1 is available for download at <https://ln2.sync.com/dl/d443caf60/uzekst9c-4szuhjcr-uxtb96v9-vnpk94iu>. Alternative file formats (.mp3, .aiff, .wav, and .ogg) are available as supplemental content appended to this dissertation.

**Figure 4.22 Audio Collage of *Die Ursonate*, 3<sup>rd</sup> movement, Scherzo A section.**



In the opening motive, “Lanke trr gll,” the waveform of Schwitters’s the first syllable, “lan” indicates a longer duration than that of the opening piano chord and the best alignment of motives places the piano onset slightly after Schwitters’s first syllable begins to sound. One explanation for this comes from articulatory differences between piano and speech. The voiced consonant /l/ phoneme that initiates this first syllable takes

the formal name of the alveolar lateral approximant. To produce this sound one places the tongue against their upper teeth and pushes air around the sides of their mouth. In studies on speech rhythm, particularly those interested in measuring the durational properties of syllable units, the standard approach to syllable length takes vowel duration as the essential measurement, while “the length of consonant clusters is irrelevant, whether in the onset or the coda, and whether constant or arbitrarily variable.”<sup>178</sup> Thus in this case, aligning the piano onset not with the start of the initial consonant of “lanke” but with the onset of the subsequent vowel sound is not unreasonable; and, as the audio demonstrates, sounds quite plausible. In the second motive of the *scherzo*, “pe pe pe pe pe,” the waveform of Schwitters’s recitation differs from Avdeeva’s performance of Chopin’s Scherzo in C sharp minor, Op. 39, mm.1–2 for reasons of articulation as well. The initial phoneme, /p/ is a voiceless bilabial stop, which means that it is produced without vibration of the vocal chords (voiceless), articulated with both lips (bilabial), and obstructs the airflow in the vocal tract (stopped). As such, the five iterations of the syllable “pe” appear as discrete attacks separated by this voiceless bilabial stop, while Avdeeva’s performance of Chopin’s Scherzo in C sharp minor, Op. 39, mm.1–2 appears without the same attack distinction. Nevertheless there is a similarity of shape between the two and, as was also the case in the previous motive, the audio comparison demonstrates that their similarities may overcome this articulatory difference. The

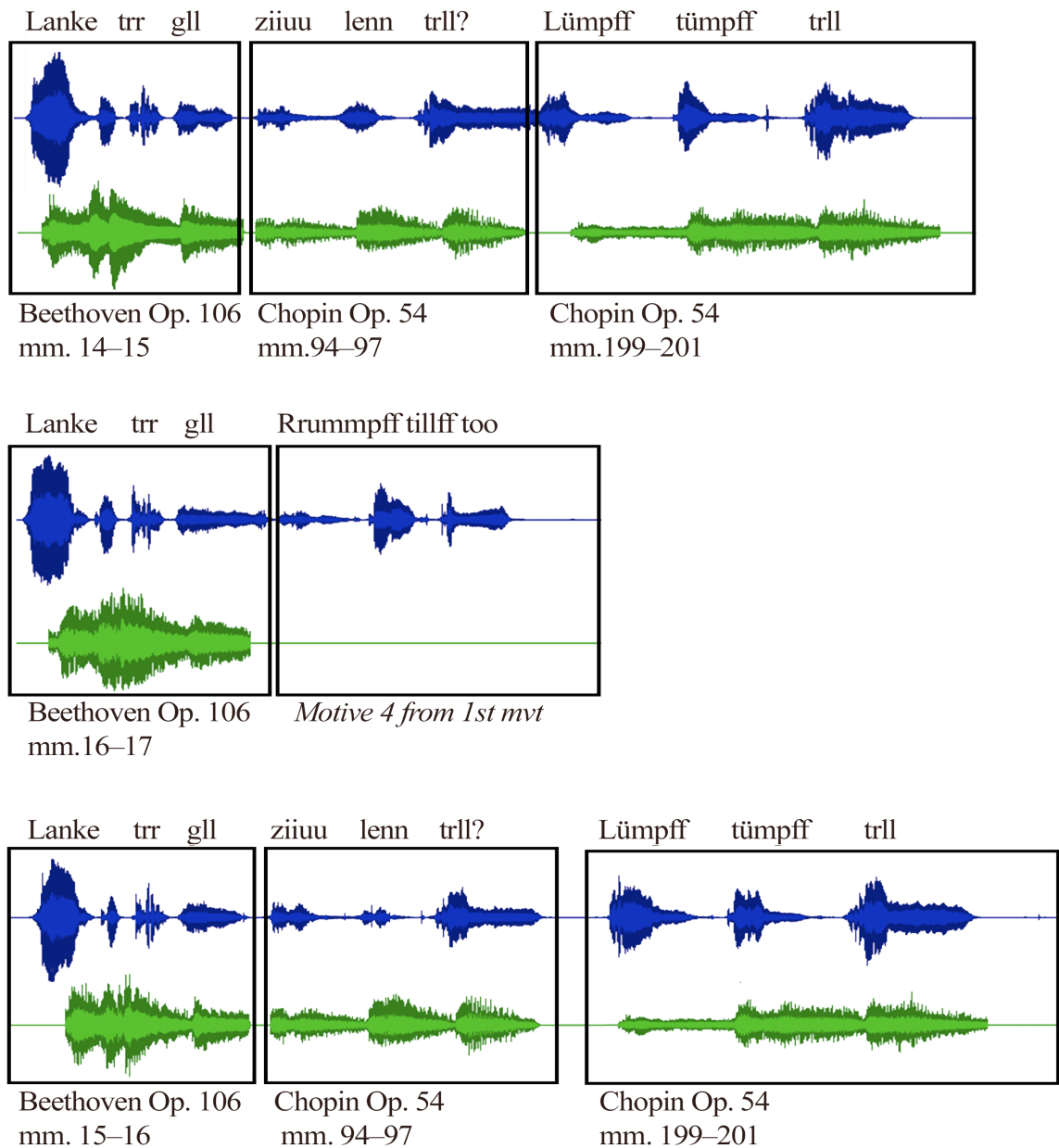
---

<sup>178</sup> Dafydd Gygbon and Ulrike Gut, “Measuring Speech Rhythm,” in *EUROSPEECH 2001 Scandinavia: Proceedings from the 7th European Conference on Speech Communication and Technology*, ed. Paul Dalsgaard et al. (Aalborg, Denmark, 2001), 96.



similarities continue between Schwitters's performance and the piano recording collage of the scherzo's B section (Figure 4.23 and [Audio 4.2](#)).<sup>179</sup>

**Figure 4.23 Audio Collage of *Die Ursonate*, 3<sup>rd</sup> movement, Scherzo B section**



<sup>179</sup> Audio 4.2 is available for download at <https://ln2.sync.com/dl/89aec4630/ne5rkdnb-wcbkqh8c-4kzrz78h-6wq2hnyd>. Alternative file formats (.mp3, .aiff, .wav, and .ogg) are available as supplemental content appended to this dissertation.

While there are occasions in which attack onset and/or decay differ slightly, the profiles of the piano performance fragments consistently parallel those of the vocalized syllables such that, taken together, the piano collage provides a musical analog to Schwitters's performance. And despite the four piano pieces being in entirely different keys (C# minor, E Major, D Major, and B♭ Major) the harmonic motion between fragments is surprisingly palatable.

When Kurt Schwitters writes that “the third movement is a real scherzo,” he means so not just in formal structure but in motivic design, writing “note the rapid succession of the three themes: *lanke trr gll*, *pe pe pe pe pe*, and *Ooka*, which are very different from one another, whereby the scherzo character and bizarre form arises.”<sup>180</sup> To be a real scherzo is perhaps to take the term quite literally—in other words, to understand the work as embodying both the traditional concept of a scherzo (a light or playful composition typically comprising a movement in a symphony or sonata) as well as the literal meaning of the word in the original Italian, as a joke or a jest.

Schwitters was known for hiding (or claiming to hide) objects or images in his artwork so why not in *Die Ursonate* as well? There are scholars who may know the answer:

Schwitters had begun sketches of *Die Ursonate* in musical notation and if those sketches include the *scherzo*, then one could certainly refute all I have proposed thus far. But to

---

<sup>180</sup> Schwitters, “erklärungen zu meiner ursonate,” reprinted in Mizelle, *Kurt Schwitters's Ursonate and Quanta and Hymn to Matter*, 71.

me, the point is not whether Schwitters really did base the motives in the *scherzo* on found object piano works. The point is that it is *possible*.

In his visual collages, Schwitters aimed for a temporally engaged viewing experience in which the surface presents half-hidden remnants of the past, reformed in new and creative ways. But while Schwitters transforms and repurposes found objects and fragments from the past, he rarely obscures their found-ness or past-ness altogether. He seems, rather, to want his audience to enjoy the tension between things lost and found, new and old, art and trash, or – as it is the case in *Die Ursonate* – speech and music.

## 4.4 Conclusions

This analysis, which focuses on an overlooked movement from Kurt Schwitters's *Ursonate* and finds productive method in Schwitters's artistic practice of found-object collage, offers a distinctly different way to explore the musicality of Spoken Word from those strategies employed in previous chapters. The strategy presented here is not unlike percussionist David Dockery's way of finding musical underpinnings in dramatic and/or theatrical performances. Dockery's works, which take short speeches from popular movies and television shows and rhythmically interpret them as duets for percussion and spoken word, powerfully demonstrate the rhythmic nature of language and the relationship between our ability to speak emotively and our ability to understand and internalize musical rhythms.<sup>181</sup> In scenes by Gene Wilder, John Cleese, Dennis Reynolds,

---

<sup>181</sup> For examples of David Dockery's work, visit his YouTube Channel, [https://www.youtube.com/user/daviddockery20?fbclid=IwAR2tCQEGPSl\\_0JJothv\\_ds6S\\_ClhgekpUTSnySNtPQ0QLenpx4MRTL3Gxss](https://www.youtube.com/user/daviddockery20?fbclid=IwAR2tCQEGPSl_0JJothv_ds6S_ClhgekpUTSnySNtPQ0QLenpx4MRTL3Gxss).

Jack Black, Seth MacFarlane, and Dwayne “the Rock” Johnson, Dockery makes explicit through his drumming the complex timing and structural organization of these actors’ performances. Whether these performances were music before Dockery got his hands on them, who knows? Whether the actors themselves consciously timed their lines or monitored their pitch and dynamic inflections, who cares? What matters is that, through his interpretations, Dockery uncovers uniquely musical qualities latent in these performances and, by presenting his interpretation not separate from but in collage with their original sources, allows the listener to discover such musical qualities too.

Such discovery has been the primary aim of this chapter. But there have been other aims as well. Through found objects, scores, and recordings that shared sonic properties with Ernst Schwitters’s well-known performance of the *scherzo*, I aim to bring focus to a movement that may not, at first glance, seem complex enough to merit thorough analytical investigation, and find merit in the possibilities at play in this “bizarre form.”<sup>182</sup> Through adopting an analytical approach that directly models Schwitters’s artistic approach and captures something of *Merz* philosophy, I aim for an analysis that emulates Schwitters’s own working method of first choosing raw materials with which to work, then transforming them in such a way as to prompt an entirely new experience of familiar materials. And finally, through exploring the sonic possibilities of collage and the metamorphosis of motivic fragments, I aim to situate the listener in the site of the

---

<sup>182</sup> Schwitters, “erklärungen zu meiner ursonate,” reprinted in Mizelle, *Kurt Schwitters’s Ursonate and Quanta and Hymn to Matter*, 71.

imaginary, and convince the reader that Schwitters's *scherzo* offers a vantage point from which one can play the game "what if I called this music" at its most expert level.

*Die Ursonate* by Kurt Schwitters is far from the only sound poem from the inter-war avant-garde period to interrogate the boundaries between language and music and in turn, speech and song. Indeed, every chapter in the present work to this point has addressed singular examples within a large body of works worthy of deeper musical consideration. In the following, final chapter, I discuss other spoken-word songs, slam poems, and sound poems that have potential to respond productively to musical perspectives and music-theoretical modes of analysis.

## Chapter 5

### 5 Last Words

Throughout this dissertation, I approached a variety of spoken-word songs, slam poems, and sound poems from a musical perspective, playing the game “what if I were to call this music?” In analyses of selected spoken-word song recordings by singer-songwriter Tom Waits, I showed how the rhythms and pitch inflections of a speaker work in musical ways to establish natural sounding speech as a musical instrument. In analyses of selected slam poems by slam founder Marc Kelly Smith, I showed how theories of musical form can shed light on the structural design of slam poetry and how the emergent design of a slam performance can emulate familiar musical forms. Finally, in my analysis of the scherzo from Kurt Schwitters’s *Die Ursonate*, I showed how “the assemblage for artistic purposes of all imaginable materials” can serve not only to create sound poetry but as a mode of musical engagement and analysis.<sup>183</sup> The array of analytical approaches and listening strategies presented offer a basis for future investigations of Spoken Word as an artistic practice that has a legitimate place within musicological and analytical research.

While establishing a stronger foothold for Spoken Word within musical scholarship has been my primary goal, a secondary objective has also been to consider the implications of adopting a musically informed perspective towards speech and language arts more generally. With both these aims in mind, this final chapter unfolds in two parts, the first of which reconsiders the legitimacy of a principled boundary between speech and song

---

<sup>183</sup> Quoted in Dachy, *The Dada Movement, 1915-1923*, 116.

and explores the implications of doing so not only for musical research but for the many disciplines within which language and music invariably meet and intertwine. The second section considers other avenues for further music-analytical research on spoken word, culminating with a discussion of a number of other notable artists and spoken-word works.

## 5.1 Implications for Cognitive Research

Thus far in the present work, I have referred to “what if I called this music?” as a game; yet asking such a question beyond the realm of things one already calls music is much more than a game—it is a mindset. To play the game with Spoken Word is to push the ontological and epistemological boundaries of language and music, challenging *a priori* assumptions that predetermine the constituents of and relationship between speech and song. In the context of music theory, accepting such a challenge is a creative enterprise. For the fields of music perception and cognition—where such a binary starting point continues to constrict the spectrum of variability and intersection, often reproducing normative assumptions in the evidence—it is an imperative.

Such efforts are evident in the growing body of speech-to-song illusion experiments since Diana Deutsch first discovered the effect in 2003.<sup>184</sup> These studies establish a clear connection between speech and song and contribute to the substantial evidence for a close link between language and music processing in general. In a 2017 study investigating the perceptual consequences of musical and linguistic listening on timing

---

<sup>184</sup> Deutsch, *Phantom Words and Other Curiosities*.

perception and absolute pitch memory, for example, Graber, Simchy-Gross and Margulis construct an experiment that attempts to distinguish between musical and linguistic listening modes without influencing participants to adopt one or the other.<sup>185</sup> In her 2013 book, *On Repeat: How Music Plays the Mind*, Elizabeth Hellmuth Margulis proposes that speech becomes music by crossing the language-music divide and that repetition is the gatekeeper between. While a dichotomy between language and music remains essential to her argument, the divide between the two domains narrows substantially by attending to the role of repetition. “When language is being repetitive,” Margulis writes, “language is being musical.”<sup>186</sup>

Despite the methodological differences of the analyses presented in chapters 2, 3, and 4, each analytical approach ultimately exposes the crucial role of repetition for establishing a viable musical interpretation of speaking songs. The rhythmic and melodic structures revealed through recording analyses of spoken-word songs are significant musically because they recur—either chronologically as the songs unfold, or concurrently in the melodic and harmonic settings in which the spoken-word part is enveloped. Similarly, the formal structures of the slam poems, revealed through generalized form-functional analysis, are centrally (though not exclusively) predicated upon repetition, be it of a particular word, a motive (or *reč*), a rhythm, or a sentiment. Repetition is most at the heart of my referential collage analysis of the Scherzo from *Die Ursonate*, not only

---

<sup>185</sup> Graber, Simchy-Gross, and Margulis, “Musical and Linguistic Listening Modes in the Speech-to-Song Illusion Bias Timing Perception and Absolute Pitch Memory.”

<sup>186</sup> Elizabeth Hellmuth Margulis, *On Repeat: How Music Plays the Mind*, ProQuest E (Oxford: Oxford University Press USA, 2013), 162.



because the work involves a great degree of motivic repetitions, but because the very process of hearing in a given sound a mimetic relation to sounds once heard before is to invoke repetition—the doing or saying again—across temporal, aesthetic, an authorial bounds. Moreover, repetition was integral to my analytical process, regardless of the method employed. Whereas speech-to-song illusion studies distinguish between transforming and non-transforming utterances based on a capped number of repetitions, for the purposes of the present project I listened to these works thousands of times.<sup>187</sup>

While every piece struck me as uniquely musical from the outset, it may well be the case that after a thousand repetitions, any spoken utterance would begin to transform into song for any listener.

As research on the speech-to-song illusion continues to grow, the current thought is that the phenomenon occurs to a greater degree for some stimuli than for others because of particular cues which prompt listeners to perceive a spoken phrase as song. Yet it may be plausible that the phenomenon occurs not to a greater degree in some stimuli, but at a swifter *rate* for some stimuli. It may be of interest to explore whether utterances currently designated “non-transforming” would transform without a cap on repetitions. Another consideration for comparative research on language and music is the degree to which the

---

<sup>187</sup> Tierney et al., “Speech versus Song: Multiple Pitch-Sensitive Areas Revealed by a Naturally Occurring Musical Illusion,” 250. This study determined a set of 24 transforming utterances (TUs) and 24 non-transforming utterances (NTUs) gave listeners 8 repetitions to determine whether what they heard more closely resembled speech or song. This particular set of utterances has since become the stimuli for several studies, including Graber, Simchy-Gross, and Margulis, ‘Musical and Linguistic Listening Modes in the Speech-to-Song Illusion Bias Timing Perception and Absolute Pitch Memory’; Vanden Bosch der Nederlanden, Hannon, and Snyder, ‘Finding the Music of Speech: Musical Knowledge Influences Pitch Processing in Speech’; Tierney, Patel, and Breen, ‘Acoustic Foundations of the Speech-to-Song Illusion.’

perceptual distinction between speech and song is influenced by external primers built into experimental design. Specifically, does the very task that would ask participants to distinguish between speech and song necessarily prime a certain listening approach?

While the results of such experiments can certainly reveal a great deal about our perceptual experiences, the underlying bias potentially limits any exploration of our perceptual potential. Before asking what the difference is between speech and song, ought we not first ask whether the categorical distinction itself is valid? How might my perception of a speech act differ if I call it song at the outset?

In each of the cases analysed in this dissertation, not only is the resultant difference indeed significant, but the approach is reasonably justified, given that the artists at work all express a similar perspective towards the somewhat arbitrary distinction between music and language. If researchers were to draw from works by spoken-word artists in order to form stimuli, for example, not only would speech be placed on a more even playing field with musical stimuli but it might be that such artistically composed speech could be informative about the limits of the language-music dichotomy altogether.

## 5.2 Avenues for Further Research and Analysis

Spoken Word has recently received some attention in musicological circles through the work of Marian Wilson Kimber's 2017 book, *The Elocutionists: Women, Music, and the Spoken Word*. While her work focuses primarily on a historical account of spoken-word song practices from an earlier era, and on the women engaged in these practices, Wilson Kimber recognizes an aesthetic relationship between the elocutionists she studies and the

styles of spoken-word poetry that the present work investigates.<sup>188</sup> She rightly points out, however, that works by women practicing Spoken Word consistently go overlooked, underappreciated, and unpublished/recorded, even when the number of female practitioners outnumber their male counterparts, writing “many of the famous performers who combined music and the spoken word were male, but I came to understand these figures as similar to high-profile chefs in a world where most cooking is done by women.”<sup>189</sup>

Where female spoken-word artists are the central subjects of Wilson Kimber’s work, the “high-profile chefs” conspicuously form the central subject of mine. While fantastic examples of spoken-word artists whose works respond well to a music-analytical approach, Tom Waits, Marc Kelly Smith, and Kurt Schwitters are far from the only artists worth studying and future music-analytical studies of spoken-word song, slam poetry, and sound poetry would benefit from including women and minorities occupying the southern hemisphere of voiced texts.

### 5.2.1 Spoken-Word Songs

American avant-garde artist, composer, musician and film director Laurie Anderson is as, if not more, prolific when it comes to spoken-word songs as Tom Waits. Analyses of the myriad spoken-word songs that make up much of Anderson’s oeuvre could alone fill a book. Nearly all of Anderson’s studio albums include some degree of speech, whether it

---

<sup>188</sup> Wilson Kimber, *The Elocutionists: Women, Music, and the Spoken Word*, 230.

<sup>189</sup> *Ibid.*, 10.

be exclusively spoken songs (as in “From the Air” and “Born, Never Asked” from the 1982 album *Big Science*), spoken verses and sung choruses (as in “Let X=X/It Tango” from *Big Science* and “Language is a Virus” from the 1986 album *Home of the Brave*), ongoing alternations between speaking and singing (as in “Excellent Birds” from the 1984 album *Mister Heartbreak*), or full-on *sprechstimme*, the kind of vocals for which Anderson is most famous (as in “Dark Angel” from the 2001 album *Life on a String*).

The 1960s American pop girl group, The Shangri-Las, led by singer Mary Weiss, also merits mention for their commercial success with spoken-word songs, though their output in no way compares to that of Waits or Anderson. Nevertheless, their 1965 single, “I can Never Go Home Anymore” reached #6 on the U.S. pop chart and features verses spoken by lead singer Weiss which alternate with sung choruses. Their 1966 single, “Past, Present, Future” is an exclusively spoken song in which teenage angst emotes over Beethoven’s “Moonlight Sonata.” In both these songs, Weiss’s recitation becomes an integral part of the musical whole and recording analysis techniques such as those applied to Tom Waits’s spoken-word songs would shed light on the subtle intricacies of Weiss’s most musical recitation.

The study and analysis of spoken-word songs opens up new possibilities for engaging with the new direction American singer-songwriter Lana Del Rey takes in her most recent project, *Violet Bent Backwards Over the Grass* (2020). The project features Del Rey performing spoken-word vocals and Jack Antonoff (her writing and producing partner on the 2019 album *Norman Fucking Rockwell*) contributing musical soundscapes ranging from lo-fi and Lynchian (“The Land of 1,000 Fires”) to slow-building and orchestral (“Bare Feet on Linoleum”).

Depending on how one looks at it, the audio release of *Violet Bent Backwards Over the Grass* is either an album of spoken-word songs or an “audiobook-with-music” and its mixed critical reception seems correlative to whether one approaches it as music or not. The following review from *New Musical Express (NME)*, a British music journalism website and former magazine publishing since 1952, for example, treats *Violet Bent Backwards Over the Grass* as first and foremost a record, despite its publication as an audiobook:

Lana’s poetry is beautiful and vivid – she’s always been a master lyricist and long spoken of her love of the likes of Walt Whitman and Allen Ginsberg, after all. What makes ‘Violet...’ so brilliant, though, are the little details that light up her passages, giving glimpses into her (or someone else’s) life...[Jack Antonoff] is a great collaborator – his backings creating the perfect atmosphere for Lana’s tales, accompanying her with soft, smoky jazz lilt or, as with ‘Bare Feet On The Linoleum’s’ discordant, chaotic layers of strings. The tone-setting takes are just as important to the ‘Violet...’ experience as Lana’s words themselves – too much or too little would throw the whole balance off, but these foundations get it just right.<sup>190</sup>

Ultimately, the review is positive and recognizes both words and music in its response to the work. On the flip side, a review of *Violet Bent Backwards Over the Grass* published in *The Guardian* poetry section is less flattering:

I listened to the 39-minute [Violet Bent Backwards Over the Grass], in which Del Rey reads 13 poems set to soft piano chords, occasional saxophone riffs and a staticky sound that could be the sea... This first poem, LA Who Am I to Love You, is the best of the collection... The

---

<sup>190</sup> Rhian Daly, “Lana Del Rey’s new poetry audiobook ‘Violet Bent Backwards Over The Grass’” *NME* 28<sup>th</sup> July 2020. <https://www.nme.com/features/lana-del-rey-new-poetry-book-violet-bent-backwards-over-the-grass-2716586>. Accessed October 12, 2020.

other 33-and-a-half minutes are not as strong. The writing is sometimes cliched, often rambling and consistently solipsistic... This is not only exasperating, it's boring. There's nothing wrong with artists drawing on their own experiences. Del Rey's mistake is forgetting that part of being a poet is to tell us about the world, not just focus on herself.<sup>191</sup>

Where the review from *New Musical Express* speaks to musical underpinnings of each track and treats the album as a collaborative work between Del Rey and Antonoff, the review from *The Guardian* minimizes the role of the musical setting and is negatively critical of Del Rey's writing. I bring this example forward not to offer my own opinion of Del Rey's attempt at Spoken Word but rather to demonstrate the degree to which an aesthetic valuation of a work differs depending on the perspective one adopts when listening to spoken-word songs. Del Rey's project is relevant to music theorists and musicologists alike who have an interest in the ontological and phenomenological properties of music.

Various other artists whose spoken-word songs merit further music-analytical study should not go unmentioned include William Shatner, Gil Scott-Heron, Ken Nordine, Lawrence Ferlinghetti, David Bowie, James Brown, Patti Smith, Gorillaz, Morrissey, Joy Harjo, and the Velvet Underground. In every case, the method of recording analysis is not only plausible but recommended, as these works tend to take on object status in the form of studio and/or live recordings.<sup>192</sup>

---

<sup>191</sup> Yara Rodrigues Fowler, "Lana Del Rey's poetry debut review – sometimes cliché, always solipsistic," *The Guardian* 29<sup>th</sup> July 2020. Accessed October 12, 2020.

<sup>192</sup> A selected list of spoken-word songs by the aforementioned artists can be found in Appendix A.

## 5.2.2 Performance and Slam Poetry

Music-analytical approaches to Spoken Word are presently nominal within the field of music theory and analysis, resulting in ample opportunities to explore what happens when we think musically about the flow and patterns of spoken-word songs. Even more abundant are the opportunities to consider the musical properties of slam and performance poetry. The central objective of my study of slam founder Marc Kelly Smith's two slam poems, "Kiss It" and "My Father's Coat," was to demonstrate the ways in which theories of musical form and form functionality can shed light on the structural design of slam poetry, drawing parallels between the poetic devices that shape, expand, and guide the trajectory of poetic utterances and the functional devices that govern the shaping of musical forms.

Such an approach may offer unique insights into the work of other slam and performance poets whose performances have been explored via ethnopoetic methods, such as Lynne Procope's performance of "elemental woman," analyzed using ethnopoetics by John Miles Foley; "Hurly Crowbirds at Mockingbars" by Buddy Wakefield and of "Like" by Mike McGee, which served as Cara Losier Chanoine's central performances for ethnopoetic "slam-scriptions," as she calls them; and "Coffee or Tea?" by U.K. based performance poet Jackie Hagan, which Julia Novak has analyzed using a musically informed ethnopoetic transcription method.<sup>193</sup>

---

<sup>193</sup> Foley, *How to Read an Oral Poem*, 87–120. Chanoine's textual source for "Hurling Crowbirds at Mockingbars" comes from Wakefield's website, while a 2009 video (a performance excerpted from "The Elephant Engine High Dive Revival" multi-poet show) serves as the performance record. Her analysis of "Like" considers three versions of the

Other slam and performance poets whose work merit further music-analytical study include Nikki Giovanni, who is not a typical slam artist, but is a spoken-word poet and an artist that many modern poets will cite as an inspiration; Saul Williams, who cannot be classified as only a spoken-word artist, as he is also a poet, a performer and a musician; Olivia Gatwood, best known for her collaborative slam poem, “Say No,” and is the author of *New American Best Friend*, a collection of poems; Melissa Lozada-Oliva, a U.S. National Poetry Slam Champion, a Brenda Moosey Video Slam Champion, and the author of the chapbook *Plastic Pájaros*; Lisa Buscani, who got her start in Chicago's performance poetry scene and ultimately became a U.S. National Poetry Slam Champion as well as a Pushcart Prize nominee for poetry; and Brandan Leake, a spoken-word poet who is the most recent winner of America's Got Talent and the show's first-ever spoken-word artist.<sup>194</sup>

### 5.2.3 Sound Poetries

While Kurt Schwitters's *Die Ursonate* is among the most well-known sound poems composed by Futurist and Dadaist artists operating in the early years of the twentieth century, it is far from the only work from that particular historical moment that merits music-analytical study. Other sound poems, such as *K' perioum* (1918) by Raoul Hausmann, *Karawane* (1916) by Hugo Ball, *Eigasing rin Jalamand* (1920) by Baroness

---

poem: a text from McGee's website in 2004; a video record of a 2006 performance of the poem during season five of *Def Poetry Jam*; and the version of the poem that appears in his 2009 print collection, *In Search of Midnight*. Chanoine, *Poetry, Points, and Performance: Expanding the Scope of Slam Poetry Analysis*, 103–11; 120–26. Novak, *Live Poetry: An Integrated Approach to Poetry in Performance*, 221–31.

<sup>194</sup> A selected list of slam performances by the aforementioned artists and other relevant artists can be found in Appendix B.



Elsa von Freytag-Loringhoven, *L'amiral cherche une maison à louer* (1916) by Richard Huelsenbeck, Marcel Janko, Tristan Tzara, and Toshiro Sawa, and *Zang Tumb Tumb* (1914) by F.T. Marinetti might similarly respond to a collage analysis method that probes the uniquely musical qualities latent in these works and presents an interpretation not separate from but in collage with original sources, allowing listeners to discover such musical qualities anew.

By no means did the early twentieth-century avant-garde artists who pioneered sound poetry exhaust its possibilities. Many others active in post-War years to the present day are worth mentioning. John Giorno and Jackson Mac Low, American poets and performance artists active from the 1960s onwards both created numerous works of sound poetry that problematize the bounds between language and music. Similarly, Henri Chopin, a French avant-garde poet and musician who practiced concrete poetry and sound poetry throughout the second half of the twentieth century, was an early adopter of tape and studio technologies to manipulate the sound of the human voice and prioritize the orality (read: musicality) of language.<sup>195</sup> The Canadian sound poetry group Four Horsemen (bpNichol, Rafael Barreto-Rivera, Paul Dutton and Steve McCaffery) active primarily between 1972 and 1988 offer unique examples where speech itself is the primary musical instrument at play. Sound poetry continues to thrive in the works of Dutch avant-garde performance artist Jaap Blonk (b. 1953), Canadian experimental poet

---

<sup>195</sup> Concrete poetry—a poetic form in which the typographical arrangement of linguistic elements and resulting visual effect are the central conveyors of meaning—should not be confused with *musique concrète*, a type of composition that utilizes recorded sounds as raw material. What a music scholar would call *musique concrète* more closely resembles Chopin's sound poetries than it does his work in concrete poetry.

Christian Bök (b. 1966), and American text-sound composer Charles Amirkhanian (b. 1945).

The list of composers, perchance more familiar to the reader, whose experimentation with text-sound composition produced works that problematize the bounds between language and music, is too long to list. The speech works of Stockhausen, Berio, Ligeti, Babbitt, Reich, Lucier, Aperghis, and Cage (to name only a few) offer numerous opportunities for the adventurous analyst to explore the musical properties of spoken word. John Cage's interest in language as a compositional medium alone could fill a book. Throughout his career, Cage composed myriad works that feature speaking parts and increasingly gravitated towards spoken-word composition. The second movement of *Living Room Music* (1940), for example, transforms four performers of unspecified instruments into a spoken-word quartet. The movement comprises speech fragments of Gertrude Stein's 1938 poem "The World is Round" and is akin to the Dadaist simultaneous poems of Huelsenbeck, Ball, and Tzara. Cage's interest in the relationship between speech and music reveals itself again in his *Lecture on Nothing* (1949-50), which was composed in rhythmic patterns and notated in a proto-metric form, as well as in his *Lecture on Something* (1951) composed the following year.<sup>196</sup> While often treated as 'experimental texts,' the game "what if I called this music?" would approach these works as musical scores, exploring possibilities for performance and analysis.

---

<sup>196</sup> Ironically, the *Lecture on Something* (1951) incorporates considerably more passages of silence than his *Lecture on Nothing* (1949-50).

In the late 1960s, Cage became increasingly interested in work with language, letters, syllables, words, and phrases, particularly when freed from syntax and meaning. *Moreau* (1970) for solo voice and *62 Mesostics Re Merce Cunningham* (1971) for solo amplified voice are early examples of Cage's interest in becoming more and more a spoken-word composer.<sup>197</sup> Cage recorded many of these works himself, and while these recordings are easily accessible to music theorists, to my knowledge they have yet to be explored via music-analytic methods such as recording analysis, nor examined from any generalized theory of formal structures.

### 5.3 Last Last Words

While this chapter, at times, reads much like a series of lists, the objective here is to demonstrate that spoken word, and the musical exploitation of speech sounds, is far from exhausted in the present study. Artists across all disciplines have been exploring the relations between spoken and musical sound systems throughout the twentieth century, and while musicologists and music theorists are finally beginning to catch up, the sheer volume of examples mentioned above suggests something of an unmapped territory that is ready and waiting to be explored.

The game “what if I call this music?” as applied to voiced texts allows for an exploration of the sounds of the speaking voice that intentionally connects this largely unmapped province of artistic expression to the more thoroughly charted province musical

---

<sup>197</sup> I have mentioned only a small sliver of Cage's spoken-word works here, in the interest of sparing the reader from another lengthy list to read. A more comprehensive account of Cage's spoken-word compositions appears in Appendix C.

composition. The particular path taken in this dissertation is similar, metaphorically speaking, to hiking a loop trail—starting relatively near the musical domain with spoken-word song, moving further afield into the realm of slam and performance poetry, and circling back towards the musical domain by way of sound poetics and sound-text compositions. This is but one of many possible routes one can take to begin to expand the region of artistic expression that one calls music. Another loop trail might explore the relations between spoken and musical sound systems by traversing the realm of film and theatre, examining the musical properties inherent in famously orated soliloquies and monologues, both accompanied by a musical score or performed acapella. Along this route one could move further afield into political speeches, then circle back towards the traditional domain of music by way of sprechstimme or dramatic narrations like Copeland's *Lincoln Portrait*, Stravinsky's *L'histoire du Soldat*, or Prokofiev's *Peter and the Wolf*.

Whatever route taken, each analytical investigation along the way can be thought of as a singular landmark selected from many possible landmarks that would have equally merited time and consideration. By traversing three different regions at the outskirts of the musical domain in this dissertation, I covered a large amount of ground, but necessarily overlooked many other possible landmarks in each respective area. Further research on the relations between spoken and musical sound systems might benefit most by focusing on a single type of Spoken Word and explore what insights might materialize from a comparative study of spoken-word songs by different artists and across different artistic traditions. Such a thorough examination of this particular region, which is arguably nearest to the traditional domain of music, would illuminate the categorical

overlap that exists among Spoken Word and music, strengthening the connections between the realms of music and language and bringing Spoken Word further into the view of mainstream musicological research and analysis.

## Appendices

The following three appendices provide the reader references to other spoken word works that may be of interest for further analysis or investigation. Wherever possible, hyperlinks to publicly available recordings are provided. Entries are sorted in ascending chronological order according to year of release or composition (provided in the right-most column of each table). These appendices are not an attempt to provide exhaustive lists. The selections included represent those spoken-word works which are among my personal favorites and which may be of interest for future research on the relationship between spoken and musical sound systems.

### Appendix A. Table of selected spoken-word songs, focusing on songs by popular music artists for further research and analysis.

Spoken Word Song	Artist	Album	Year
<a href="#">Abraham</a>	Jack Kerouac & Steve Allen	Poetry For The Beat Generation	1959
<a href="#">The moon and her majesty</a>	Jack Kerouac & Steve Allen	Poetry For The Beat Generation	1959
<a href="#">Four Song Poems</a>	Kenneth Patchen	Kenneth Patchen Reads Jazz In Canada	1959
<a href="#">I can Never Go Home Anymore</a>	Shangri-Las	Golden Hits of The Shangri-Las	1966
<a href="#">Past Present Future</a>	Shangri-Las	Golden Hits of The Shangri-Las	1966
<a href="#">Go Cry on Somebody Else's Shoulder</a>	The Mothers of Invention	Freak Out!	1966
<a href="#">The End</a>	The Doors	The Doors	1967
<a href="#">Horse Latitudes</a>	The Doors	Strange Days	1967
<a href="#">The Gift</a>	Velvet Underground	White Light/White Heat	1968
<a href="#">The Murder Mystery</a>	Velvet Underground	The Velvet Underground	1969

<a href="#"><u>A Boy Named Sue</u></a>	Johnny Cash	At San Quentin	1969
<a href="#"><u>Whitey On the Moon</u></a>	Gil Scott-Heron	Small Talk at 125th and Lenox	1970
<a href="#"><u>When the Revolution Comes</u></a>	The Last Poets	The Last Poets	1970
<a href="#"><u>On the Subway</u></a>	The Last Poets	Single	1970
<a href="#"><u>I Found a Reason</u></a>	Velvet Underground	Loaded	1970
<a href="#"><u>The Revolution Will Not Be Televised</u></a>	Gil Scott-Heron	Pieces of a Man	1971
<a href="#"><u>Stranger in Blue Suede Shoes</u></a>	Kevin Ayers	Whatevershebringswesing	1971
<a href="#"><u>No Knock</u></a>	Gil Scott-Heron	Free Will	1972
<a href="#"><u>I Can't Write Left Handed</u></a>	Bill Withers	Live at Carnegie Hall	1973
<a href="#"><u>March to the Witch's Castle</u></a>	Funkadelic	Cosmic Slop	1973
<a href="#"><u>Future Legend</u></a>	David Bowie	Diamond Dogs	1974
<a href="#"><u>The Jeweler</u></a>	John Cale	Slow Dazzle	1975
<a href="#"><u>Emotional Weather Report</u></a>	Tom Waits	Nighthawks at the Diner	1975
<a href="#"><u>A Feast of Friends</u></a>	Jim Morrison	An American Prayer	1978
<a href="#"><u>Romeo is Bleeding</u></a>	Tom Waits	Blue Valentine	1978
<a href="#"><u>Wave</u></a>	Patti Smith	Wave	1979
<a href="#"><u>Seen and Not Seen</u></a>	Talking Heads	Remain in Light	1980
<a href="#"><u>Private Investigations</u></a>	Dire Straits	Lover Over Gold	1982
<a href="#"><u>Let X=X/It Tango</u></a>	Laurie Anderson	Big Science	1982
<a href="#"><u>From the Air</u></a>	Laurie Anderson	Big Science	1982
<a href="#"><u>Born, Never Asked</u></a>	Laurie Anderson	Big Science	1982
<a href="#"><u>Shore Leave</u></a>	Tom Waits	Swordfishtrombones	1983
<a href="#"><u>Touble's Braids</u></a>	Tom Waits	Swordfishtrombones	1983
<a href="#"><u>Frank's Wild Years</u></a>	Tom Waits	Swordfishtrombones	1983
<a href="#"><u>Armageddon Man</u></a>	Black Flag	Family Man	1984

<a href="#"><u>History Lesson Part II</u></a>	The Minutemen	Double Nickels on the Dime	1984
<a href="#"><u>And Dream of Sheep</u></a>	Kate Bush	Hounds of Love	1985
<a href="#"><u>Paintwork</u></a>	The Fall	This Nation's Saving Grace	1985
<a href="#"><u>9th and Hennepin</u></a>	Tom Waits	Rain Dogs	1985
<a href="#"><u>Language is a Virus</u></a>	Laurie Anderson	Home of the Brave	1986
<a href="#"><u>In The Kingdom #19</u></a>	Sonic Youth	EVOL	1986
<a href="#"><u>Belong</u></a>	R.E.M	Green (25th Anniversary Deluxe Edition)	1988
<a href="#"><u>Fourth of July</u></a>	Galaxie 500	This Is Our Music (Deluxe Edition)	1990
<a href="#"><u>Justify My Love</u></a>	Madonna	The Immaculate Collection	1990
<a href="#"><u>The Ocean Doesn't Want Me</u></a>	Tom Waits	Bone Machine	1992
<a href="#"><u>Marbles</u></a>	Tindersticks	Tindersticks	1993
<a href="#"><u>Park Life</u></a>	Blur	Parklife	1994
<a href="#"><u>My Sister</u></a>	Tindersticks	Tindersticks (Deluxe Edition)	1995
<a href="#"><u>Susan's House</u></a>	Eels	Beautiful Freak	1996
<a href="#"><u>A Space Boy Dream</u></a>	Belle and Sebastian	The Boy with the Arab Strap	1998
<a href="#"><u>You Are Invited</u></a>	Dismemberment Plan	Emergency & I	1999
<a href="#"><u>What's He Building In There</u></a>	Tom Waits	Mule Variations	1999
<a href="#"><u>I am Waiting</u></a>	Lawrence Ferlinghetti	Coney Island of the Mind	1999
<a href="#"><u>Dark Angel</u></a>	Laurie Anderson	Life on a String	2001
<a href="#"><u>Love Detective</u></a>	The Arab Strap	The Red Thread	2001
<a href="#"><u>Ahoy There!</u></a>	Mr. Scruff	Trouser Jazz	2002
<a href="#"><u>Positive Jam</u></a>	The Hold Steady	Almost Killed Me	2004
<a href="#"><u>Children's Story</u></a>	Tom Waits	Orphans: Brawlers, Bawlers & Bastards	2006
<a href="#"><u>Be Safe</u></a>	The Cribs	Men's Needs, Women's Needs, Whatever	2007



---

<a href="#"><u>Vox Humana</u></a>	Deerhunter	Weird Era Cont.	2008
<a href="#"><u>A Machine for Loving</u></a>	Iggy Pop	Préliminaires	2009
<a href="#"><u>Moonbeams</u></a>	Yoko Ono	Take Me to the Land of Hell	2013
<a href="#"><u>#Selfie</u></a>	Chainsmokers	The Duff (Original Motion Picture Soundtrack)	2015

**Appendix B. Table of selected slam and performance poems of interest for further research and analysis.**

<b>Slam/Performance Poem</b>	<b>Artist</b>	<b>Year</b>
<a href="#"><u>Matisse</u></a>	Gertrude Stein	1911
<a href="#"><u>You Walk, Resembling Me</u></a>	Maria Tsvetaeva	1913
<a href="#"><u>If I Told Him: A Completed Portrait of Picasso</u></a>	Gertrude Stein	1923
<a href="#"><u>I Carry Your Heart With Me (I Carry It In My Heart)</u></a>	ee cummings	1952
<a href="#"><u>Let's From Some Loud Unworld's Most Rightful Wrong</u></a>	ee cummings	1953
<a href="#"><u>Canto LI</u></a>	Ezra Pound	1958
<a href="#"><u>The Weary Blues (1925)</u></a>	Langston Hughes	1958
<a href="#"><u>Heaven's Worry</u></a>	Ezra Pound	1970
<a href="#"><u>Amethyst Rocks</u></a>	Saul Williams	1997
<a href="#"><u>elemental woman</u></a>	Lynne Procope	2000
<a href="#"><u>Poem for Magic</u></a>	Quincy Troupe Jr.	2002
<a href="#"><u>So Edgar Allan Poe Was In This Car</u></a>	Terry Jacobus	2002
<a href="#"><u>Like</u></a>	Mike Mcgee	2006
<a href="#"><u>Coffee or Tea</u></a>	Jackie Hagan	2007
<a href="#"><u>Hurly Crowbirds at Mockingbars</u></a>	Buddy Wakefield	2009
<a href="#"><u>Mother</u></a>	Two Tongues (Cristin O'Keefe Aptowicz & Aimee Nezhukumatathil)	2009
<a href="#"><u>Banshee</u></a>	Lisa Buscani	2009
<a href="#"><u>Skinheads</u></a>	Patricia Smith	2010
<a href="#"><u>Ohm</u></a>	Saul Williams	2010
<a href="#"><u>Thanx for Nothing</u></a>	John Giorno	2011
<a href="#"><u>Full House</u></a>	Tom Hanks	2012
<a href="#"><u>What Kind of Asian Are You?</u></a>	Alex Dang	2013
<a href="#"><u>Dog</u></a>	Lawrence Ferlinghetti	2013

---

<a href="#"><u>Say No</u></a>	Olivia Gatwood & Megan Falley	2014
<a href="#"><u>Paper People</u></a>	Harry Baker	2014
<a href="#"><u>When Love Arrives</u></a>	Sarah Kay & Phil Kaye	2015
<a href="#"><u>My Honest Poem</u></a>	Rudy Francisco	2015
<a href="#"><u>Like Totally Whatever</u></a>	Melissa Lozada-Oliva	2015
<a href="#"><u>Beginning, Middle &amp; End</u></a>	Phil Kaye	2018

---

**Appendix C. Table of selected sound poems and text-sound compositions for further research and analysis.**

<b>Sound Poem(s)</b>	<b>Composer</b>	<b>Year</b>
<a href="#"><u>Karawane</u></a>	Hugo Ball	1916
<a href="#"><u>L'amiral cherche une maison à louer</u></a>	Richard Huelsenbeck, Marcel Janko, Tristan Tzara, & Toshiro Sawa	1916
<a href="#"><u>Totenklage</u></a>	Hugo Ball	1916
<a href="#"><u>K' perioum</u></a>	Raoul Hausmann	1918
<a href="#"><u>(brüllt)</u></a>	Tristan Tzara	1918
<a href="#"><u>Eigasing rin Jalamand</u></a>	Baroness Elsa von Freytag-Loringhoven	1920
<a href="#"><u>Living Room Music</u></a>	John Cage	1940
<a href="#"><u>Lecture on Nothing</u></a>	John Cage	1950
<a href="#"><u>Karlheinz Stockhausen</u></a>	Gesang der Junglinge	1955
<a href="#"><u>Thema-Omaggio á Joyce</u></a>	Luciano Berio	1958
<a href="#"><u>Schnurmilch</u></a>	Jean/Hans Arp (1887-1966)	1961
<a href="#"><u>Aventures</u></a>	György Ligeti	1962
<a href="#"><u>Philomel</u></a>	Milton Babbitt	1964
<a href="#"><u>Nouvelles Aventures</u></a>	György Ligeti	1965
<a href="#"><u>5th Bluebird Asymmetry</u></a>	Jackson Mac Low	1967
<a href="#"><u>The Sinking of the Titanic</u></a>	Gavin Bryars	1969
<a href="#"><u>Moreau</u></a>	John Cage	1970
<a href="#"><u>Mr Smith in Rhodesia</u></a>	Åke Hodell	1970
<a href="#"><u>I am sitting in a room</u></a>	Mary and Alvin Lucier	1970
<a href="#"><u>62 Mesostics Re Merce Cunningham</u></a>	John Cage	1971
<a href="#"><u>Heavy Aspirations</u></a>	Charles Amirkhanian	1973
<a href="#"><u>Speech Songs</u></a>	Charles Dodge	1974
<a href="#"><u>Empty Words</u></a>	John Cage	1974
<a href="#"><u>Writing for the second time through Finnegans Wake</u></a>	John Cage	1977

<a href="#"><u>Récitations</u></a>	Georges Aperghis	1978
<a href="#"><u>Six Fantasies on a Poem by Thomas Campion</u></a>	Paul Lansky	1979
<a href="#"><u>Jeux de presence</u></a>	Henri Chopin	1980
<a href="#"><u>History of Collage</u></a>	Charles Amirkhanian	1981
<a href="#"><u>Monotony</u></a>	Four Horsemen	1984
<a href="#"><u>On Speaking Terms II: Poem No. 5</u></a>	Lars-Gunnar Bodin	1986
<a href="#"><u>Just a Thought</u></a>	Jaap Blonk	1992

## Bibliography

- Adams, Kyle. "Aspects of the Music/Text Relationship in Rap." *Music Theory Online* 14, no. 2 (2008).
- Adams, Kyle. "On the Metrical Techniques of Flow in Rap Music." *Music Theory Online* 15, no. 5 (2009).
- Attas, Robin. "Form as Process: The Buildup Introduction in Popular Music." *Music Theory Spectrum* 37, no. 2 (2015): 275–96.
- Bäckström, Per. "The Intermedial Cluster: Åke Hodell's Lågsniff." *Acta Universitatis Sapientiae, Film and Media Studies* 10, no. 1 (2016). <https://doi.org/10.1515/ausfm-2015-0025>.
- Bateson, Gregory. *Steps to an Ecology of Mind: Collected Essays in Anthropology, Psychiatry, Evolution, and Epistemology*. New York: Ballantine Books, 1972.
- Benadon, Fernando. "Time Warps in Early Jazz." *Music Theory Spectrum* 31, no. 1 (2009): 1–25.
- Bernhart, Walter. "Rhythmical Ambivalence of Poetry Performance: The Case of Elizabethan Verse." In *Word and Music Studies: Essays on Performativity and on Surveying the Field*, edited by Walter Bernhart and Michael Halliwell, 87–100. Amsterdam: Rodopi, 2011.
- Besson, Mireille, and Angela D. Friederici. "Language and Music: A Comparative View." *Music Perception: An Interdisciplinary Journal* 16, no. 1 (1998): 1–9.
- Bickford, Tyler. "Music of Poetry and Poetry of Song: Expressivity and Grammar in Vocal Performance." *Ethnomusicology* 51, no. 3 (2007): 439–76.
- Bjorklund, Beth. "Form, Anti-Form, and Informality: Reinventing Free Verse." *Poetics Today* 16, no. 3 (1995): 547–67.
- Boyle, Antares. "Formation and Process in Repetitive Post Tonal Music." PhD Dissertation, University of British Columbia, 2018.
- Burger, Jeff, and Tom Waits. "Waits: The Beats Go On..." In *Tom Waits on Tom Waits: Interviews and Encounters*, edited by Paul Jr. Maher, 39–42. Chicago: Chicago Review Press, 2011.
- Burnham, Bo. *Egghead: Or, You Can't Survive on Ideas Alone*. New York: Grand Central Publishing, 2013.
- Butler, Mark. *Unlocking the Groove: Rhythm, Meter, and Musical Design in Electronic Dance Music*. Bloomington: Indiana University Press, 2006.
- Butterfield, Matthew W. "Participatory Discrepancies and the Perception of Beats in

- Jazz.” *Music Perception: An Interdisciplinary Journal* 27, no. 3 (2010): 157–76.
- Cannam, Chris, Christian Landone, and Mark Sandler. “Sonic Visualiser: An Open Source Application for Viewing, Analysing, and Annotating Music Audio Files.” In *Proceedings of the ACM Multimedia International Conference*, 2010.
- Caplin, William E. *Classical Form: A Theory of Formal Functions for the Instrumental Music of Haydn, Mozart, and Beethoven*. Oxford: Oxford University Press, 1998.
- . “Teaching Classical Form: Strict Categories vs. Flexible Analyses.” *Dutch Journal of Music Theory* 18, no. 3 (2013): 119–35.
- . “The ‘Expanded Cadential Progression’: A Category for the Analysis of Classical Form†.” *Journal of Musicological Research* 7, no. 2–3 (1987): 215–57.
- Carroll, Cath. *Tom Waits*. New York: Thunder Mouth’s Press, 2000.
- Cason, Nia, Corine Astésano, and Daniele Schön. “Bridging Music and Speech Rhythm: Rhythmic Priming and Audio-Motor Training Affect Speech Perception.” *Acta Psychologica* 155 (2015): 43–50.
- Cason, Nia, and Daniele Schön. “Rhythmic Priming Enhances the Phonological Processing of Speech.” *Neuropsychologia* 50, no. 11 (2012): 2652–58.
- Cavanagh, Clare, Roland Greene, Stephen Cushman, Jahan Ramazani, and Paul Rouzer, eds. *The Princeton Encyclopedia of Poetry and Poetics*. 4th ed. Princeton, N.J.: Princeton University Press, 2012.
- Chanoine, Cara Losier. *Poetry, Points, and Performance: Expanding the Scope of Slam Poetry Analysis*. PhD dissertation, Indiana University of Pennsylvania, 2018.
- Chomsky, Noam, and Morris Halle. *The Sound Pattern of English*. New York: Harper and Row, 1968.
- Cook, Nicholas. *Beyond the Score: Music as Performance*. Oxford: Oxford University Press, 2013.
- . “Methods for Analysing Recordings.” In *The Cambridge Companion to Recorded Music*, edited by Nicholas Cook, Eric Clarke, Daniel Leech-Wilkinson, and John Rink, 221–45. Cambridge: Cambridge University Press, 2009.
- . *Music as a Creative Practice*. Oxford: Oxford Scholarship Online, 2018.
- Cooper, Grovesnor, and Leonard Meyer. *The Rhythmic Structure of Music*. Chicago: Chicago University Press, 1960.
- Cummins, Fred. “Speech and Song: A Joint Perspective.” *Music Perception* 37, no. 4 (2020): 339–50.
- Cummins, Fred. “Interval Timing in Spoken Lists of Words.” *Music Perception: An Interdisciplinary Journal* 22, no. 3 (2005): 497–508.

- Dachy, Marc. *The Dada Movement, 1915-1923*. Geneva; New York: Skira; Rizzoli, 1990.
- Davies, Matthew E.P., and Mark D. Plumbley. "Context-Dependent Beat Tracking of Musical Audio." *IEEE Transactions on Audio, Speech and Language Processing* 15, no. 3 (2007).
- Dayan, Peter. "The Music of Dada Follows Kurt Schwitters to Holland (Bypassing Berlin) and Ends in the Ursonate." In *The Music of Dada: A Lesson in Intermediality for Our Times*, 129–56. New York: Routledge, 2019.
- Deutsch, Diana. *Musical Illusions and Paradoxes*. La Jolla: Philomel Records, 1995.
- . *Phantom Words and Other Curiosities*. La Jolla: Philomel Records, 2003.
- . "The Perceptual Organization of Streams of Sound." In *Musical Illusions and Phantom Words*. New York: Oxford University Press, 2019.
- Deutsch, Diana, Trevor Henthorn, and Rachael Lapidis. "Illusory Transformation from Speech to Song." *The Journal of the Acoustical Society of America* 129, no. 4 (2011): 2245–52.
- Dietrich, Doratheia. *The Collages of Kurt Schwitters*. Cambridge: Cambridge University Press, 1993.
- Dilley, Laura, and J. Devin McAuley. "The Fabb-Halle Approach to Metrical Stress Theory as a Window on Commonalities Between Music and Language." In *Language and Music as Cognitive Systems*, edited by Patrick Rebuschat, Martin Rohrmeier, John A. Hawkins, and Ian Cross, 22–31. Oxford: Oxford University Press, 2012.
- Duxbury, C, Jp Bello, Mike Davies, and Mark Sandler. "Complex Domain Onset Detection for Musical Signals." *Proceedings of the 6th Conference on Digital Audio Effects*, 2003.
- Elderfield, John. *Kurt Schwitters*. London: Thames and Hudson, 1985.
- Eliot, T.S. *The Music of Poetry: The Third W.P. Ker Memorial Lectures Delivered in the University of Glasgow 24th February 1942*. Folcroft, PA: The Folcroft Press, 1969.
- Everett, Todd, and Tom Waits. "Tom Waits: In Close Touch With the Streets." In *Tom Waits on Tom Waits: Interviews and Encounters*, edited by Paul Jr. Maher, 45–49. Chicago: Chicago Review Press, 2011.
- Fabb, Nigel, and Morris Halle. "Grouping in the Stressing of Words, in Metrical Verse, and in Music." In *Language and Music as Cognitive Systems*, edited by Patrick Rebuschat, Martin Rohermeier, John A. Hawkins, and Ian Cross, 4–21. Oxford: Oxford University Press, 2012.
- Falk, Simone, Tamara Rathcke, and Simone Dalla Bella. "When Speech Sounds like Music." *Journal of Experimental Psychology: Human Perception and Performance*



- 40, no. 4 (2014): 1491–1506.
- Feinsod, Henry. “Sound Poetry.” In *The Princeton Encyclopedia of Poetry and Poetics*, edited by Roland Greene, Stephen Cushman, Clare Cavanagh, Jahan Ramazani, and Paul Rouzer, 1327–29. Princeton, N.J.: Princeton University Press, 2012.
- Feinstein, Sascha. *Jazz Poetry: From the 1920s to the Present*. Westport, Conn.: Greenwood Press, 1997.
- Foley, John Miles. “Word-Power, Performance, and Tradition.” *The Journal of American Folklore* 105 (1992): 275–301.
- . “Signs, Texts, and Oral Tradition.” *Journal of Folklore Research* 33, no. 1 (1996): 21–29.
- . *How to Read an Oral Poem*. Urbana: University of Illinois Press, 2002.
- Frost, Corey. “The Omnidirectional Microphone: Performance Literature as Social Project.” Ph.D. diss., City University of New York, 2010.
- Goldsmith, John. *Autosegmental and Metrical Phonology*. Oxford: B. Blackwell, 1990.
- Gordon, Reyna L., Cyrille L. Magne, and Edward W. Large. “EEG Correlates of Song Prosody: A New Look at the Relationship between Linguistic and Musical Rhythm.” *Frontiers in Psychology* 2, no. NOV (2011): 1–13.
- Gregory, Helen. “(Re)Presenting Ourselves: Art, Identity, and Status in U . K . Poetry Slam.” *Oral Tradition* 2, no. 23 (2008): 201–17.
- Graber, Emily, Rhimmon Simchy-Gross, and Elizabeth Hellmuth Margulis. “Musical and Linguistic Listening Modes in the Speech-to-Song Illusion Bias Timing Perception and Absolute Pitch Memory.” *The Journal of the Acoustical Society of America* 142, no. 6 (2017): 3593–3602.
- Gybbon, Dafydd, and Ulrike Gut. “Measuring Speech Rhythm.” In *EUROSPEECH 2001 Scandinavia: Proceedings from the 7th European Conference on Speech Communication and Technology*, edited by Paul Dalsgaard, Børge Lindberg, Henrik Benner, and Zheng-hua Tan, 95–98. Aalborg, Denmark, 2001.
- Halle, Morris. “The Stress of English Words 1968-1998.” *Linguistic Inquiry* 29, no. 4 (1998): 539–68.
- Halle, Morris, and Jean-Roger Vergnaud. “Stress and the Cycle.” *Linguistic Inquiry* 18, no. 18 (1987): 45–84.
- Hannon, Erin E. “Perceiving Speech Rhythm in Music: Listeners Classify Instrumental Songs According to Language of Origin.” *Cognition* 111, no. 3 (2009): 403–9.
- Hausen, Maija, Ritva Torppa, Viljami R. Salmela, Martti Vainio, and Teppo Särkämö. “Music and Speech Prosody: A Common Rhythm.” *Frontiers in Psychology* 4 (2013): 1–16.

- Hayes, Bruce. *Metrical Stress Theory: Principles and Case Studies*. Chicago: University of Chicago Press, 1995.
- Hayes, Bruce, and Abigail Kaun. "The Role of Phonological Phrasing in Sung and Chanted Verse." *The Linguistic Review* 13, no. 3–4 (1996): 1–60.
- Hepokoski, James, and Warren Darcy. *Elements of Sonata Theory: Norms, Types, and Deformations in the Late-Eighteenth Century Sonata*. Oxford: Oxford University Press, 2008.
- Hoffman, Tyler. *American Poetry in Performance: From Walt Whitman to Hip Hop*. Ann Arbor: University of Michigan Press, 2011.
- Hoskyns, Barney. *Low Side of the Road: A Life and Times of Tom Waits*. New York: Broadway Books, 2009.
- Howland, Patricia. "Formal Structures in Post-Tonal Music." *Music Theory Spectrum* 37, no. 1 (2015): 71–97.
- Hudson, Martyn. "Schwitters's Ursonate and the Merz Barn Wall." *Leonardo Music Journal* 25 (2015): 89–92.
- Huelsenbeck, Richard., Malcom Green, and Alastair Brotchie, eds. *The Dada Almanac*. 2nd ed. London: Atlas Press, 1993.
- Humphries, Patrick. *Small Change: The Life of Tom Waits*. London: Omnibus Press, 1989.
- Humphries, Patrick. *The Many Lives of Tom Waits*. London: Omnibus Press, 2007.
- Hunter, Christina Rosine. "Romantic Fragments: Kurt Schwitters's Collages." Columbia University, 2003.
- Huron, David, and Ann Ommen. "An Empirical Study of Syncopation in American Popular Music, 1890–1939." *Music Theory Spectrum* 28, no. 2 (2006): 211–31.
- Hymes, Dell. *Narrative Inequality: Toward an Understanding of Voice*. London: Taylor & Francis, 1996.
- Jackendoff, Ray. "Parallels and Nonparallels between Language and Music." *Music Perception: An Interdisciplinary Journal* 26, no. 3 (2009): 195–204.
- Jacobs, Jay S. *Wild Years: The Music and Myth of Tom Waits*. Toronto: ECW Press, 2000.
- Jescheniak, Jörg D., Anja Hahne, and Angela D. Friederici. "Brain Activity Patterns Suggest Prosodic Influences on Syntactic Parsing in the Comprehension of Spoken Sentences." *Music Perception: An Interdisciplinary Journal* 16, no. 1 (1998): 55–62.
- Jones, Dafydd. *Dada Culture: Critical Texts on the Avante-Garde*. Amsterdam: Rodopi, 2006.

- Jones, M.E. "African-American Jazz Poetry: Orality, Prosody and Performance." Ph.D. diss., Stanford University, 2001.
- Keil, Charles, and Stephen Feld. *Music Grooves: Essays and Dialogues*. Chicago: University of Chicago Press, 1994.
- Kessel, Corinne. *The Words and Music of Tom Waits*. Westport, Conn.: Praeger, 2009.
- King, Stephen. *On Writing: A Memoir of the Craft*. New York: Scribner, 2000.
- Kivy, Peter. *The Performance of Reading: An Essay in the Philosophy of Literature*. Malden: Blackwell Publishers, 2006.
- Klein, Jeff. "Tom Waits and the Right of Publicity: Protecting the Artist's Negative Voice." *Popular Music and Society* 37, no. 5 (2014): 583–94.
- Krebs, Herald. *Fantasy Pieces: Metrical Dissonance in the Music of Robert Schumann*. New York: Oxford University Press, 1999.
- Krims, Adam. *Rap Music and the Poetics of Identity*. Cambridge: Cambridge University Press, 2000.
- Lach, Friedhelm. *Der Merz Künstler Kurt Schwitters*. Cologne: Verlag M. Dumont, 1971.
- Large, Edward W., and Caroline Palmer. "Perceiving Temporal Regularity in Music." *Cognitive Science* 26, no. 1 (2002): 1–37.
- Last, Rex William. *German Dadaist Literature: Kurt Schwitters, Hugo Ball, Hans Arp*. New York: Twayne Publishers, 1973.
- Lemire, Chantal. "At the 'Crossroads': Interactions Between Speech Rhythm and Musical Rhythm in Tom Waits's Spoken Word Songs." MA Thesis, University of British Columbia, 2013.
- Lerdahl, Fred. "The Sounds of Poetry Viewed as Music." *Annals of the New York Academy of Sciences* 930, no. 1 (2001): 337–54. <https://doi.org/10.1111/j.1749-6632.2001.tb05743.x>.
- Lewer, Debbie. "Hugo Ball, Iconoclasm and the Origins of Dada in Zurich." *Oxford Art Journal*, 2009.
- Liberman, Mark. "The Intonational System of English." Massachusetts Institute of Technology, 1975.
- Liberman, Mark, and Alan Prince. "On Stress and Linguistic Rhythm." *Linguistic Inquiry* 8, no. 2 (1977): 249–336.
- Lichtenstein, Sabine. "Listening to Kurt Schwitters' 'Ursonate'" 38 (2003): 276–85.
- List, George. "The Boundaries of Speech and Song." In *Music, Words, and Voice: A Reader*, edited by Martin Clayton, 24–32. Manchester: Manchester University Press, 2008.

- London, Justin. *Hearing in Time: Psychological Aspects of Time*. 2nd ed. Oxford: Oxford University Press, 2012.
- Lord, Albert B. *The Singer of Tales*. Cambridge: Harvard University Press, 1960.
- Losada, Catherine C. "Between Modernism and Postmodernism: Strands of Continuity in Collage Compositions by Rochberg, Berio, and Zimmermann." *Music Theory Spectrum* 31, no. 1 (2009): 57–100.
- Luke, Megan R. *Kurt Schwitters: Space, Image, Exile*. Chicago: Chicago University Press, 2013.
- Macarthur, John. "Schwitters and Benjamin: The Modernity of the Baroque and Romanticism." *Journal of Architecture* 15, no. 3 (2010): 283–300.
- Maher, Paul Jr. *Tom Waits on Tom Waits: Interviews and Encounters*. Chicago: Chicago Review Press, 2011.
- Mantell, James T., and Peter Q. Pfordresher. "Vocal Imitation of Song and Speech." *Cognition* 127 (2013): 177–202.
- Margulis, Elizabeth H., Rhimmon Simchy-Gross, and Justin L. Black. "Pronunciation Difficulty, Temporal Regularity, and the Speech-to-Song Illusion." *Frontiers in Psychology* 6 (2015): 1–7.
- Margulis, Elizabeth Hellmuth. *On Repeat: How Music Plays the Mind*. ProQuest E. Oxford: Oxford University Press USA, 2013.
- Margulis, Elizabeth Hellmuth. "Repetition and Emotive Communication in Music versus Speech." *Frontiers in Psychology* 4, no. APR (2013): 1–4.
- Marinetti, Filippo Tommaso. *Selected Poems and Related Prose*. Edited by Elizabeth R. Napier and Barbara R. Studholme. New Haven and London: Yale University Press, 2002.
- Mauch, Matthias, Chris Cannam, Rachel Bittner, George Fazekas, Justin Salamon, Jiajie Dai, Juan Bello, and Simon Dixon. "Computer-Aided Melody Note Transcription Using the Tony Software: Accuracy and Efficiency." *Proceedings of the First International Conference on Technologies for Music Notation and Representation*, 2015.
- Mauch, Matthias, and Simon Dixon. "PYIN: A Fundamental Frequency Estimator Using Probabilistic Threshold Distributions." In *IEEE International Conference on Acoustics, Speech and Signal Processing*, 2014.
- McCarthy, Eugene B., and Fran Quinn. *Sound Ideas: Hearing and Speaking Poetry*. Brookline: Hobblebush Books, 2013.
- McGowan, Rebecca W., and Andrea G. Levitt. "A Comparison of Rhythm in English Dialects and Music." *Music Perception: An Interdisciplinary Journal* 28, no. 3 (2011): 307–14.

- Merrill, Julia, Daniela Sammler, Marc Bangert, Dirk Goldhahn, Gabriele Lohmann, Robert Turner, and Angela D. Friederici. "Perception of Words and Pitch Patterns in Song and Speech." *Frontiers in Psychology* 3, no. 76 (2012).
- Mizelle, Dary John. *Kurt Schwitters' Ursonate and Quanta and Hymn to Matter*. PhD Dissertation, University of California, 1977.
- . "Kurt Schwitters's Ursonate and Quanta and Hymn to Matter." University of California, 1977.
- Montandon, Mac. *Innocent When You Dream: The Tom Waits Reader*. New York: Thunder Mouth's Press, 2005.
- Moore, Allan. *Rock: The Primary Text: Developing a Musicology of Rock*. 2nd ed. Burlington, VT: Ashgate, 2001.
- Neill, Annegrith. "Decoding Merz: An Interpretive Study of Kurt Schwitters's Early Work, 1918-1922." PhD dissertation, University of Texas at Austin, 1990.
- Novak, Julia. *Live Poetry: An Integrated Approach to Poetry in Performance*. Amsterdam: Rodopi, 2011.
- Olsen, Lance. "Notes Toward the Musicality of Creative Disjunction, Or: Fiction by Collage." *Symploke* 12, no. 1/2 (2004): 130–35.
- Olsson, Jesper. "Kneaded Language: Concrete Poetry and New Media in the Swedish 1960s." *Modernism - Modernity*, 2011. <https://doi.org/10.1353/mod.2011.0038>.
- Ox, Jack. "Creating a Visual Translation of Kurt Schwitters's Ursonate." *Leonardo Music Journal* 3, no. 1993 (1993): 59–61.
- Palmer, Caroline, and Sean Hutchins. "What Is Musical Prosody?" In *Psychology of Learning and Motivation: Advances in Research and Theory*, edited by B.H. Ross, 245–278. Amsterdam: Elsevier Press, 2006.
- Patel, Aniruddh D. "Syntactic Processing in Language and Music: Different Cognitive Operations, Similar Neural Resources?" *Music Perception* 16, no. 1 (1998): 27–42.
- . "Musical Rhythm, Linguistic Rhythm, and Human Evolution." *Music Perception: An Interdisciplinary Journal* 24, no. 1 (2006): 99–104.
- . *Music, Language, and the Brain*. New York: Oxford University Press, 2008.
- Patel, Aniruddh D., John R Iversen, and Jason C Rosenberg. "Comparing the Rhythm and Melody of Speech and Music: The Case of British English and French." *The Journal of the Acoustical Society of America* 119, no. September 2005 (2006): 3034–47.
- Patel, Aniruddh D., and Joseph R. Daniele. "An Empirical Comparison of Rhythm in Language and Music." *Cognition* 87, no. 1 (2003): 35–45.

- . “Stress-Timed vs. Syllable-Timed Music? A Comment on Huron and Ollen (2003).” *Music Perception: An Interdisciplinary Journal* 21, no. 2 (2003): 273–76.
- Perloff, Marjorie. *The Futurist Moment: Avant-Garde, Avant Guerre, and the Language of Rupture*. Chicago: Chicago University Press, 2003.
- Perloff, Marjorie, and Craig Dworkin, eds. *The Sound of Poetry/The Poetry of Sound*. Chicago: University of Chicago Press, 2009.
- Perloff, Nancy. “Sound Poetry and the Musical Avant-Garde: A Musicologist’s Perspective.” In *The Sound of Poetry/The Poetry of Sound*, edited by Craig Dworkin and Marjorie Perloff, 97–117. Chicago: Chicago University Press, 2009.
- . “Schwitters Redesigned: A Post-War Ursonate from the Getty Archives.” *Journal of Design History* 23, no. 2 (2010): 195–203.
- Peterson, James Braxton. “The Depth of the Hole: Intertextuality and Tom Waits’s ‘Way Down in the Hole.’” *Criticism* 52, no. 3–4 (2010): 461–85.
- Petrusich, Amanda, and Tom Waits. “Interview: Tom Waits.” In *Tom Waits on Tom Waits: Interviews and Encounters*, edited by Paul Jr. Maher, 416–24. Chicago: Chicago Review Press, 2011.
- Pierrehumbert, J. “The Phonetics and Phonology of English Intonation.” PhD diss., Massachusetts Institute of Technology, 1980.
- . “Tonal Elements and Their Alignment.” In *Prosody: Theory and Experiment*, edited by M. Horne, 11–36. Dordrecht, The Netherlands: Kluwer, 2000.
- Price-Styles, Alice. “MC Origins: Rap and Spoken Word Poetry.” In *The Cambridge Companion to Hip-Hop*, edited by Justin A. Williams, 11–21. Cambridge: Cambridge University Press, 2015.
- Quené, Hugo, and Robert F. Port. “Effects of Timing Regularity and Metrical Expectancy on Spoken-Word Perception.” *Phonetica* 62, no. 1 (2005).
- Ratz, Erwin. *Einführung in Die Musikalische Formenlehre: Über Formprizipien in Den Inventionen Und Fugen J.S. Bach Und Ihre Bedeutung Für Die Kompositionstechnik Beethovens*. 3rd ed. Vienna: Universal, 1973.
- Rings, Steven. “A Foreign Sound to Your Ear: Bob Dylan Performs ‘It’s Alright, Ma (I’m Only Bleeding).’” *Music Theory Online* 19, no. 4 (2013).
- Rodgers, Stephen. “Sentences with Words: Text and Theme-Type in *Die schöne Müllerin*.” *Music Theory Spectrum* 36, no. 1 (2014): 58–85.
- Rubin, Edgar. *Synsoplevede Figurer: Studier i Psykologisk Analyse*. København og Kristiania: Gyldendal, Nordisk Forlag, 1915.
- Schechner, Richard. *Performance Theory, Revised and Expanded Edition*. London; New York: Routledge, 2006.

- Schmalenbach, Werner. *Kurt Schwitters*. Cologne: Dumont Schauberg, 1967.
- Schoenberg, Arnold. *Fundamentals in Music Composition*. Edited by Gerald Strang and Leonard Stein. London: Faber & Faber, 1967.
- Scholz, Christian. "Relations between Sound Poetry and Visual Poetry: The Path from the Optophonetic Poem to the Multimedia Text." *Visible Language* 35, no. 1 (2001): 92.
- Schulz, Isabel. *Kurt Schwitters: Color and Collage*. Houston: Menil Foundation, 2011.
- Schwitters, Kurt. *Die Reihe Merz 1923–1932*. Edited by Ursula Kocher and Isabel Schulz. *Die Reihe Merz 1923–1932*. Berlin/Boston: Walter de Gruyter GmbH, in cooperation with the Sprengel Museum Hannover, 2019.
- Selkirk, Elisabeth. *Phonology and Syntax: The Relation between Sound and Structure*. Cambridge, Mass: MIT Press, 1984.
- Selkirk, Elisabeth. "The Role of Prosodic Categories in English Word Stress." *Linguistic Inquiry* 11, no. 3 (1980): 563–605.
- Shepard, Roger N. *Mind Sights: Original Visual Illusions, Ambiguities, and Other Anomalies, with a Contemporary Commentary on the Play of Mind in Perception and Art*. New York: W.H. Freeman & Co., 1990.
- Silverton, Pete, and Tom Waits. "The Sultan of Sleaze." In *Tom Waits on Tom Waits: Interviews and Encounters*, edited by Paul Jr. Maher, 153–58. Chicago: Chicago Review Press, 2011.
- Simchy-Gross, Rhimmon, and Elizabeth Hellmuth Margulis. "The Sound-to-Music Illusion: Repetition Can Musicalize Nonspeech Sounds." *Music & Science* 1 (2018): 1–6.
- Smay, David. *Swordfishtrumpets*. New York: Continuum International Publishing Group Inc, 2008.
- Smith, Marc Kelly. *Crowdpleaser*. Chicago: College Press, 1996.
- . *Take the Mic: The Art of Performance Poetry, Slam and the Spoken Word*. Naperville: Sourcebooks MediaFusion, 2009.
- Solis, Gabriel. "'Workin' Hard, Hardly Workin': Hey Man, You Know Me" - Tom Waits, Sound, and the Theatrics of Masculinity." *Journal of Popular Music Studies* 19, no. 1 (1974): 26–59.
- Somers-Willett, Susan B.A. *The Cultural Politics of Slam Poetry: Race, Identity, and the Performance of Popular Verse in America*, 2009.
- Starky, David, and Richard Guzman, eds. *Smokestacks and Skyscrapers: An Anthology of Chicago Writing*. Chicago: Wild Onion Books, 1999.

- Stockman, Steve. "Tom Waits: Misery Is the River of the World." In *The Rock Cries Out: Discovering Eternal Truth in Unlikely Music*. Lake Mary, FL.: Relevant Books, 2004.
- Tchir, Stephen Patrick. "Tom Waits's Rain Dogs: Influences and Musical Genre." M.A. thesis, University of Alberta, 2013.
- Tedlock, Dennis. *The Spoken Word and the Work of Interpretation. The Journal of American Folklore*. Philadelphia: University of Pennsylvania Press, 1985.
- Temperley, Nicholas, and David Temperley. "Music-Language Correlations and the 'Scotch Snap.'" *Music Perception: An Interdisciplinary Journal* 29, no. 1 (2011): 51–63.
- Thomas, Margaret E. "Text and Temporality: Toward an Understanding of Rhythmic Irregularities in the Music of Tom Waits." *Music Theory Online* 22, no. 4 (2016).
- Thompson, Stephen, ed. "Tom Waits." In *The Tenacity of the Cockroach: Conversations with Entertainment's Most Enduring Outsiders*, 125–30. New York: Three Rivers Press, 2002.
- Tierney, Adam, Fred Dick, Diana Deutsch, and Marty Sereno. "Speech versus Song: Multiple Pitch-Sensitive Areas Revealed by a Naturally Occurring Musical Illusion." *Cerebral Cortex* 23, no. 2 (February 1, 2013): 249–54.
- Tierney, Adam, Aniruddh D. Patel, and Mara Breen. "Acoustic Foundations of the Speech-to-Song Illusion." *Journal of Experimental Psychology: General* 147, no. 6 (2018): 888–904.
- Tsatsoulis, Christos Iraklis, and Markus Hofmann. "Focusing on Maximum Entropy Classification of Lyrics by Tom Waits." In *IEEE International Advance Computing Conference (IACC)*, 664–67, 2014.
- Vanden Bosch der Nederlanden, Christina M., Erin E. Hannon, and Joel S. Snyder. "Finding the Music of Speech: Musical Knowledge Influences Pitch Processing in Speech." *Cognition* 143 (2015): 135–40.
- Vanden Bosch Der Nederlanden, Christina M., Erin E. Hannon, and Joel S. Snyder. "Everyday Musical Experience Is Sufficient to Perceive the Speech-to-Song Illusion." *Journal of Experimental Psychology: General* 144, no. 2 (2015): 43–49.
- Vanhandel, Leigh. "The Classic(AI) Dada Work: Kurt Schwitters' Ursonate." In *WCCMTA*, 2017.
- Waits, Tom. *Dime Store Novels Vol. 1: Live from Ebbets Field (1974)*. NMC Music Ltd 8 00945 00822 8, 2001.
- Ward, Robert, and Tom Waits. "Play It Again, Tom." In *Tom Waits on Tom Waits: Interviews and Encounters*, edited by Paul Jr. Maher, 51–59. Chicago: Chicago Review Press, 2011.



- Webster, Gwendolen. *Kurt Merz Schwitters: A Biographical Study*. Cardiff: University of Wales Press, 1997.
- Wendt, Larry. "Sound Poetry: I. History of Electro-Acoustic Approaches II. Connections to Advanced Electronic Technologies." *Leonardo* 18, no. 1 (1985): 11–23.
- Wheeler, Lesley. *Voicing American Poetry: Sound and Performance from the 1920s to Present*. Ithaca: Cornell University Press, 2008.
- Wilke, Tobias. "Da-Da: 'Articulatory Gestures' and the Emergence of Sound Poetry." *Modern Language Notes* 128, no. 3 (2013): 639–68.
- Williams, Justin A. "The Construction of Jazz Rap and High Art in Hip-Hop Music." *The Journal of Musicology* 27, no. 4 (2010): 435–59.
- Williams, Paul. "Why We Listen (Arrested Development, Pearl Jam, Elvis Costello, Tom Waits)." In *Back to the Miracle Factory: Rock Etc. 1990s*, 49–82. New York: Tom Doherty Associates, LLC, 2002.
- Wilson Kimber, Marian. *The Elocutionists: Women, Music, and the Spoken Word*. Urbana: University of Illinois Press, 2017.
- Wittgenstein, Ludwig. *Philosophical Investigations*. 2nd ed. Oxford: Basil Blackwell Ltd, 1958.
- Wright, Gavin. "Everything You Can Think of Is True: The Life and Music of Tom Waits." In *Waits Mini Biography, Narrated by Marianne Hoff Engesland*. Grieg Audio: Audiobook, 2009.
- Yust, Jason. *Organized Time: Rhythm, Tonality, and Form*. Oxford: Oxford University Press, 2018.
- Zumthor, Paul. *Oral Poetry: An Introduction*. Edited by Kathryn Murphy-Judy. Minneapolis: University of Minnesota Press, 1990.
- Zurbrugg, Nicholas. "Poetry of the Contemporary Avante-Gard." *Journal of European Studies* 9 (1979): 121–143.

## Discography

- Anderson, Laurie. *Big Science*. Warner Bros. Records, 1982. Compact Disc.
- Avdeeva, Yulianna. *Scherzo in E major, Op. 54*. All rights reserved 2010 The Fryderyk Chopin Institute (NIFC), Polish Television (TVP), National Audiovisual Institute (NInA). <https://www.youtube.com/watch?v=ORwDDsn2jTU>.
- Avdeeva, Yulianna. *Scherzo in C sharp minor, Op. 39*. All rights reserved 2010 The Fryderyk Chopin Institute (NIFC), Polish Television (TVP), National Audiovisual Institute (NInA). <https://www.youtube.com/watch?v=RGyfxoys4I>.
- Gould, Glenn. *Piano Sonata No. 24 in F-sharp major, Op. 78: 'A Therese'; Piano Sonata No. 29 in B-flat major, Pp. 106: 'Hammerklavier' / Ludwig van Beethoven*. Sony Classical, 1993. SMK 52645. Compact Disc.
- Kerouac, Jack, and Steve Allen. *Poetry for the Beat Generation*. Hanover Records LP5000, 1959. Compact Disc.
- Schwitters, Kurt, John Abram, Martin. Arnold, Jim. Tate, Erica. Goodman, Rick. Sacks, Marie. Bérard, Roberto. Occhipinti, Anthony. De Mare, and Christopher Butterfield. *Ursonate / Schwitters; Music for Klein and Beuys (1987; Pillar of snails (1984-87) / Butterfield*. Victoria, B.C: C. Butterfield, 1993.
- Schwitters, Kurt. *Kurt Schwitters–Urwerk*. Zweitausendeins 230 069, 2007. Compact Disc.
- Schwitters, Ernst. *Die Sonate in Urlauten and Other Works*. Recorded October 1958. Lord's Gallery Limited Edition 100. Compact Disc.
- Various Artists. *Arp / Schwitters / Hausmann – Dada > Antidada > Merz*. Belgium 2005. Sub Rosa. SR 195. Compact Disc.
- Waits, Tom. *The Black Rider*. Island Records CIDM314-518-559-2, 1993. Compact Disc.
- Waits, Tom. *Nighthawks at the Diner*. Asylum (Elektra) CDA2008, 1975. Compact Disc.
- Waits, Tom. *The Heart of Saturday Night*. Asylum CD1015-2, 1974. Compact Disc.
- Waits, Tom. *Real Gone*. Anti/Epitaph CD86678-2, 2004. Compact Disc.
- Waits, Tom. *Dime Store Novels Vol. 1: Live from Ebbets Field (1974)*. NMC Music Ltd 8 00945 00822 8, 2001.

## Curriculum Vitae

**Name:** Chantal Lemire

**Post-secondary Education and Degrees:**

University of British Columbia  
Vancouver, British Columbia, Canada  
2003-2006 B.Mus.

McGill University of Western Ontario  
Montreal, Québec, Canada  
2007-2010 M.Mus.

University of British Columbia  
Vancouver, British Columbia, Canada  
2010-2013 M.A.

The University of Western Ontario  
London, Ontario, Canada  
2013–2021 Ph.D.

**Honours and Awards:**

Province of Ontario Graduate Scholarship  
2015

Social Science and Humanities Research Council (SSHRC)  
Doctoral Scholarship  
2015–2017

National Educational Association of Disabled Students (NEADS)  
National Student Award  
2017

**Related Work Experience**

Teaching Assistant  
University of Western Ontario  
2013–2020

Teaching Assistant  
University of British Columbia  
2010-2013