1 Introduction

1.1 Setting a menu to music

Setting a menu to music. This expression may have its origins in a scene from the American version of an operetta originally called Das Dreimäderlhaus, first performed in Vienna in 1916. The plot of the operetta revolved around Schubert and some of his real-life circle in a fictitious story of unrequited love. In a scene unique to the American adaptation of the operetta, Schubert is portrayed as a “podgy, love-lorn Bohemian Schwammerl (mushroom) who scribble[s] gemütlich tunes on the back of menus in idle moments.” He is, of course, not literally setting the menu to music; in the scene, he is composing the music to the song Ständchen. However, as Brian Newbould says, “This concoction had an enduring influence on the popular conception of Schubert as man and composer which has still not been eradicated.”

Similar claims of text-setting sensitivity and prowess are told of (or by) other composers; reportedly, Rameau offered to set a Dutch newspaper to music, and

---

1 The American version is known as Blossom Time, and ran from 1921 until the Second World War. The beer garden scene does not appear in any other adaptation. Richard Morris, The Schubert Institute, from http://myweb.tiscali.co.uk/franzschubert/articles/mlbint.html and pages within.
Rossini may have been similarly disposed towards laundry lists. The scene in *Blossom Time* is undoubtedly the result of a history of similar stories, but it has crystallized into an image and meme that remains in the public consciousness.

The essence of the meme is that a composer with an understanding of text setting is capable of taking an otherwise inappropriate or unremarkable text and setting it to music in a way that at least camouflages the weaknesses in the text, and at best creates a masterpiece. Philip Lieson Miller invokes the meme in defending Schubert against the common criticism that he was not particularly discriminating in his choice of texts:

> Schubert, whose ability “to set a menu to music” is a popular myth, is often criticized for setting so many verses by his friends and contemporaries, verses which but for him would long ago have been forgotten.

### 1.2 What constitutes a good text setting?

When we speak of a composer’s sensitivity to aspects of text setting, it is often in an effort to describe how natural a composer’s setting of a text “feels.” It is common to read descriptions of compositional style that invoke abstract and qualitative terms such as those in the following quotes:

---

4 David Burrows, *Sound, Speech, and Music*, (Amherst: The University of Massachusetts Press, 1990), p. 78. The *setting a menu to music* meme has also been applied to Rossini and Strauss.
[Brahms] has also been criticized for his sometimes cavalier treatment of the words, for his occasional disregard of correct accentuation, and for breaking phrases. In this he was the antithesis of Hugo Wolf, who so successfully suggested natural speech.\(^6\)

Reyer’s melodic line reveals certain positive qualities, among them careful prosody, declamation that is often ingenious and subtle, and restrained use of the square phrase.\(^7\)

Within these two short quotes, there are a number of vague terms related to the composer’s skill at text setting. What is “correct” accentuation? What constitutes a “successful” setting of natural speech, “careful” prosody, or “ingenious and subtle” declamation?

Don Harrán’s *Word–Tone Relations in Musical Thought: From Antiquity to the Seventeenth Century* is a thorough discussion and compendium of theorists’ writings on text setting. Harrán describes the subject matter of his volume as falling into two categories:

Thus word-tone relations admit both a broad and narrower definition, the broad one referring to the total complex of associations between music and words in their structure and content, the narrower one to the problems and procedures of text placement – the alignment of pitches and syllables – in composition and performance. Statements of the theorists fall similarly into the one or the other category of general affinities between music and language or the specific correlation of their notes and syllables. Though the two categories are logically susceptible of differentiation, the second presupposes the first and

---

\(^6\) Miller, p. 19.
neither can be considered apart from the other for an inclusive view of the subject.\textsuperscript{8}

Harrán’s Appendix is a collection of quotes from primary sources, organized by topic; the categories are listed in Example 1.1. These categories provide an overview of aspects of text setting theorists and philosophers from the Ancient Greeks to the Renaissance found critical:

**Example 1.1: Table of Contents to the Appendix of Harrán (p. 360)**

I. Prerequisites  
II. Adapting Music to Speech  
III. Accentuation  
IV. Syntax  
V. Articulation  
VI. Pronunciation  
VII. Text Placement  
VIII. Number of Syllables per Phrase  
IX. Elisions  
X. Which Notes Carry Syllables  
XI. Position of Syllables in Phrase  
XII. Semiminims  
XIII. Series of Semiminims  
XIV. Notes Following Series of Semiminims  
XV. Dots  
XVI. Notes Following the Dot  
XVII. Dissonances, Syncopations  
XVIII. Leaps  
XIX. Repeated Notes  
XX. Repeated Motives or Phrases  
XXI. Textual Repeats

Many, if not most, of the categories in Harrán’s Appendix fall within his narrower category, dealing primarily with mechanics of text setting: the proper uses of accentuation, elisions, syllabification and melismas; appropriate setting of vowels and diphthongs; and specifics of setting text to various rhythmic patterns. All of

these topics fall under what linguists refer to as prosody, or the components that comprise the overall sound of a language. A sensitive text setting, then, must pay attention to the prosody of a language.

1.3 Prosody and paralinguistic characteristics

**Prosody** is a general term that includes the study of “intonation, rhythm, tempo, loudness, and pauses, as these interact with syntax, lexical meaning, and segmental phonology in spoken texts.”⁹ Prosody is closely intertwined with the paralinguistic aspects of language, which are messages expressed most effectively through interpersonal communication; for example, the attitude or emotional state of the speaker is carried primarily through paralinguistic elements. As a result, paralinguistic elements do not affect the meaning or understanding of written language, but have a tremendous effect on the spoken language.

Paralinguistic elements of language can be understood even when semantic meaning is masked. For example, even if an utterance is in a language the listener does not understand, or if there has been some sort of manipulation removing the semantic elements, a listener is often still able to understand information carried by paralinguistic aspects of the utterance, such as emotion and intent. Thus paralinguistic information is contained in a channel parallel to the linguistic information, and is often tightly coordinated with the linguistic information so as to reinforce or clarify its meaning.¹⁰

---


The *lexical* or *suprasegmental* phonetic features have traditionally been defined as *tone, stress, and quantity*. These features of the phonetic domain correspond to physical and psychophysical properties as illustrated in the table below: the phonetic property of *tone* corresponds to the physical property of *fundamental frequency* (also known as $F_0$) and the psychophysical property of *pitch; stress* corresponds to *intensity* and *loudness*, and *quantity* to *length* and *duration*. The terms are more or less equivalent across the domains, though they are represented differently.

### Example 1.2: Features of the phonetic domain and their correspondents in other domains

<table>
<thead>
<tr>
<th>Domains</th>
<th>Properties</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Phonetic:</strong></td>
<td>tone</td>
</tr>
<tr>
<td><strong>Physical:</strong></td>
<td>fundamental frequency ($F_0$)</td>
</tr>
<tr>
<td><strong>Psychophysical:</strong></td>
<td>pitch</td>
</tr>
</tbody>
</table>

Therefore, when an author speaks of a composer’s attention to the prosody of a language, the term can refer to any one of the three lexical properties of language, alone or in combination, and the composer’s skill in replicating those phonetic properties in music. For example, in the quote about Brahms and Wolf given above, the “occasional disregard of correct accentuation” likely refers to the
phonetic properties of stress and quantity, and Brahms’ occasional practice of aligning stressed beats with unstressed syllables to create metric ambiguity.

1.4 Project description and method

Theories regarding the relationship between the prosody of a composer’s native language and compositional style have been pervasive throughout history, and several recent studies have demonstrated aspects of the proposed relationship quantitatively.11 This study focuses on the paralinguistic properties of tone and quantity in spoken French and German to determine the effect of these two properties on the analogous properties in music, pitch/F₀ and length/duration, in the 19th-century art song repertory.

As is evident from the “setting a menu to music” meme, there is a widely held belief that composers of vocal music in this era were particularly sensitive to aspects of text setting. This study investigates several rhythmic and melodic aspects of linguistic prosody to determine what elements of French and German prosody are reflected in the musical rhythms and melodies written by French and German composers during that time.

---

1.4.1 Repertory and composers

The repertory for this study consists of French and German art songs composed in the nineteenth century; the cutoff date of composition was 1900. The composers selected for inclusion are all native speakers of either French or German. To avoid translation issues, the songs selected were those whose text was written by a poet who was a native speaker of the same language as the composer, and the text must have been originally written in that native language.\footnote{German and Austrian German are considered equivalent linguistically, therefore both German and Austrian composers and poets are included as “German” speakers.}

Other criteria for song selection included limiting the database to songs for a single voice with piano accompaniment; no multi-voice songs were included, as contrapuntal concerns may have taken precedence over melodic design. In addition, the study excludes songs with origins in the opera genre, songs of a deliberately evocative nature (thus eliminating a number of French Chinoise and other similar songs), and non-secular songs.

The French and German composers represented in the study and the total number of songs encoded for each composer are listed in Example 1.3. While some individual composers are better represented than others, there is an approximately equivalent amount of data for each language encoded for the study.
Example 1.3: Composers represented in the study

<table>
<thead>
<tr>
<th>French</th>
<th>German</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bizet</td>
<td>Brahms</td>
</tr>
<tr>
<td>(24)</td>
<td>(52)</td>
</tr>
<tr>
<td>Chausson</td>
<td>Franz</td>
</tr>
<tr>
<td>(20)</td>
<td>(61)</td>
</tr>
<tr>
<td>David</td>
<td>Schubert</td>
</tr>
<tr>
<td>(34)</td>
<td>(59)</td>
</tr>
<tr>
<td>Debussy</td>
<td>Schumann</td>
</tr>
<tr>
<td>(29)</td>
<td>(36)</td>
</tr>
<tr>
<td>Duparc</td>
<td>Strauss</td>
</tr>
<tr>
<td>(14)</td>
<td>(33)</td>
</tr>
<tr>
<td>Fauré</td>
<td>Wolf</td>
</tr>
<tr>
<td>(37)</td>
<td>(82)</td>
</tr>
<tr>
<td>Gounod</td>
<td></td>
</tr>
<tr>
<td>(29)</td>
<td></td>
</tr>
<tr>
<td>Lalo</td>
<td></td>
</tr>
<tr>
<td>(15)</td>
<td></td>
</tr>
<tr>
<td>Massé</td>
<td></td>
</tr>
<tr>
<td>(20)</td>
<td></td>
</tr>
<tr>
<td>Massenet</td>
<td></td>
</tr>
<tr>
<td>(29)</td>
<td></td>
</tr>
<tr>
<td>Reber</td>
<td></td>
</tr>
<tr>
<td>(30)</td>
<td></td>
</tr>
<tr>
<td>Reyer</td>
<td></td>
</tr>
<tr>
<td>(8)</td>
<td></td>
</tr>
<tr>
<td>Saint–Saëns</td>
<td></td>
</tr>
<tr>
<td>(7)</td>
<td></td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>Total</strong></td>
</tr>
<tr>
<td>(296)</td>
<td>(324)</td>
</tr>
</tbody>
</table>

1.4.2 Encoding process and method

This study uses the Humdrum Toolkit, an open source software program by David Huron consisting of tools, commands, and representations designed to assist with music research.\textsuperscript{13} In addition to using the existing Humdrum analysis tools, more than twenty new tools and commands were developed especially for this study.

The musical data was encoded in Humdrum’s **kern format, which is capable of representing pitch and rhythmic information as well as other musical information such as phrasing. The song text was encoded in the **silbe representation, which is a syllable-oriented text representation. Identifying information such as composer, song title, and opus number were also encoded with each song, in

\textsuperscript{13} Detailed information about the Humdrum Toolkit can be found online at http://dactyl.som.ohio-state.edu/Humdrum/index.html and pages therein.
addition to other relevant information such as the year of composition and tempo.

Example 1.4 illustrates a sample file.\(^\text{14}\)

Example 1.4: Humdrum representation of Schubert’s “Das Wandern” from

*Die schöne Müllerin*

\[
\begin{align*}
!!!COM: \text{Schubert, Franz} \\
!!!CNT: \text{Deutscher} \\
!!!TXO: \text{Deutsch} \\
!!!OTL: "Das Wandern" \\
!!!OPR: Die schoene Muellerin \\
!!!OPS: Opus 25 \\
!!!ONM: No. 1 \\
!!!ODT: 1823/ \\
!! NB: Vocal melody only encoded. \\
**kern** & **silbe** \\
*clefG2* & * \\
*k[b-e-]* & * \\
*B-:* & * \\
*M2/4* & * \\
*MM[allegretto]* & * \\
=4 & =4 \\
4r &  \\
8r &  \\
{8f} & {Das} \\
=5 & =5 \\
8f & Wan- \\
8b- & -dern \\
8a & ist \\
8ee- & des \\
=6 & =6 \\
16dd & Mu2l \\
16ff & | \\
16dd & -lers \\
16b- & | \\
8f & Lust, \\
8dd & das \\
=7 & =7 \\
8.cc & Wan- \\
16a & | \\
8b-} & -dern} \\
8r & . \\
=8 & =9 \\
4r & . \\
8r & . \\
{8f} & {das} \\
=9 & =9 \\
8f & Wan- \\
8b- & -dern \\
8a & ist \\
8ee- & des \\
=10 & =10 \\
16dd & Mu2l- \\
16ff & | \\
16dd & -lers \\
16b- & | \\
8f & Lust, \\
8dd & das \\
=11 & =11 \\
8.cc & Wan- \\
16a & |
\end{align*}
\]

\(^{14}\) Information about how to read a Humdrum file may be found online at [http://www.music-cog.ohio-state.edu/Humdrum/guide02.html](http://www.music-cog.ohio-state.edu/Humdrum/guide02.html) and associated pages.
1.5 Overview of the current study

Chapter 2 provides a brief overview of historical paradigms and analogies that have been applied to music, focusing on types of verbal paradigms and the relationship between language and music. It also draws from recent literature in the cognition and perception of both language and music to illustrate important cognitive links between the two disciplines.
Chapter 3 focuses on the rhythmic and durational aspects of the prosody of language and its effect on the rhythmic and durational aspects of musical melody. This study uses a measure designed to study rhythmic variability in language known as the *normalized Pairwise Variability Index*, or *nPVI*. The *nPVI* was first applied to music by Patel and Daniele (2003); this chapter modifies and replicates their study and discusses the results.

Chapter 4 studies melodic aspects of the prosody of language and how those aspects are reflected in music. Specifically, the range of songs and phrases are compared to the ranges of the spoken languages, and trends are identified across languages and for individual composers.