The Music of Speech:
Layering Musical Elements to Deliver Powerful Messages

Steven D. Cohen, Thomas E. Wei, Daniel C. DeFraia, and Christopher J. Drury

Steven D. Cohen is an award-winning speaker who teaches professional public speaking at Harvard University, delivers public speaking seminars at academic conferences, and helps high potential leaders improve their public speaking skills. Please direct correspondence to sdcohen@post.harvard.edu.

Thomas E. Wei is pursuing a Ph.D. in public policy at Harvard Kennedy School. He is an award-winning teacher and classically-trained violinist who has guest lectured on music and speech at Harvard.

Daniel C. DeFraia is pursuing a Master of Liberal Arts in English and American literature and language at Harvard Extension School.

Christopher J. Drury is pursuing a Master of Liberal Arts in biology at Harvard Extension School.
The Music of Speech:
Layering Musical Elements to Deliver Powerful Messages

Abstract

Music and speech are both performance arts. While these two art forms may seem distinct, they share a similar objective – to communicate a message that leaves a lasting impact. Our paper examines the relationship between music and speech and identifies the key musical elements that composers and professional speakers use to elicit specific emotional reactions from their listeners. Specifically, we explore five key musical elements – tempo, dynamics, pitch, timbre, and rhythm – in the context of selected musical compositions and public speeches and explain how speakers can layer these elements on top of one another to deliver powerful and memorable messages.

Keywords

Oral Communication, Music, Speech, Delivery, Images
The Music of Speech:
Layering Musical Elements to Deliver Powerful Messages

Introduction

Music and speech are both performance arts. While these two art forms may seem distinct, they share a similar objective – to communicate a message that leaves a lasting impact. Both music and speech depend upon a host of non-auditory factors, but it is the distinct mixture of sounds, both presented and perceived, that forms the primary basis for how these two art forms are understood.

From a cognitive standpoint, the purpose of music is “to arouse specific reactions or signify specific ideas” (Golomb 89). These reactions are not generated by random assortments of sounds, but rather by precise combinations of musical elements that skilled composers use to elicit particular emotional responses. We argue that professional speakers leverage musical elements in the same way to shape the “music” of their message. For example, Martin Luther King, Jr. leveraged specific musical elements to transform his “I Have a Dream” speech into a moving song. Had “[King] simply distributed printed flyers of his speech, the impact would have been minimal in comparison with his powerful vocal delivery” (Wennerstrom 60). Indeed, music is a “logical expression” of feeling, one that when used consciously in a speech “conveys to us what an emotion-characteristic ‘sounds’ like” (Langer 218; Davies 86).

While speech and music clearly have structural similarities, two important questions remain:

1. Is there a concrete relationship between music and speech?
2. If so, are there practical music-based techniques that speakers can use to deliver powerful messages?
We answer these questions by exploring the techniques behind powerful music and showing how these same techniques are used in speech. Specifically, we examine how composers and professional speakers use five key musical elements – tempo, dynamics, pitch, timbre, and rhythm – to evoke specific emotional responses.

This paper makes two primary contributions to the field of communications. Theoretically, we highlight the important technical linkages between music and speech that thus far have been underappreciated. Methodologically, we develop a concrete framework for using musical elements to enhance vocal delivery. We draw from multiple disciplines to bolster our theoretical arguments and provide numerous real-world music and speech examples to illustrate the practical relevance of our technical framework.

Our paper proceeds as follows: we first more fully develop the music-speech analogy, then present the key musical elements with detailed examples of how these elements are used in music, and conclude with guided speech examples that illustrate how professional speakers leverage these musical elements in practice.

The Music-Speech Analogy

Humans have the unique ability to communicate using a vocal structure that is remarkably efficient at creating complex sounds. The position of the larynx (where the vocal cords are located) is low in the neck, enabling a high degree of sound modification. Other mammals, to varying degrees, maintain a “high position of the larynx [that] severely limits the array of sounds that can be produced” (Gregg 431). In humans, “the vocal fold is adjustable in length, tension, and shape, giving the human larynx top honors for vocal versatility” (Gregg 432). Given this degree of versatility, one easily can appreciate the inherent musicality in everyday language.
There is also evidence that music has served an important evolutionary function. During the early stages of human development, “improvisation and novelty in a combined music/dance performance would indicate the cognitive flexibility of the dancer, signaling his potential for cunning and strategizing while on the hunt” (Levitin 254). That is, those humans with an aptitude for musical “performance” would be the most likely to survive, mate, and produce offspring. According to Levitin, this component of natural selection would enable humans to incorporate musical characteristics into speech (250-256). Levitin further asserts that “singing and instrumental activities might have helped our species to refine motor skills, paving the way for the development of the exquisitely fine muscle control required for vocal or signed speech” (260).

Research in vocal pedagogy suggests that the human voice is essentially an instrument. Ware argues that the voice is “a complete, unified instrument for human expression and communication” that facilitates a wide range of expression in song or everyday speech (66-67). Indeed, what makes the human voice such a powerful instrument “is that its resonances can be continually altered by movements of the larynx, jaw, tongue, and lips” (Jourdain 40). In addition, functional magnetic resonance imaging studies suggest that music involves virtually all regions of the brain, including many of the same neural mechanisms that are involved in speech processing (Levitin 86, 129-130).

Although many speakers may be consciously unaware of the associations between music and speech, they still intuitively understand the types of sounds that they find enjoyable. For instance, most individuals who find a live music performance pleasing have difficulty articulating what specifically made the performance so powerful. Although we may not
understand the intricate theory or technical structure of a song, “all of us are expert musical listeners, able to make quite subtle determinations of what we like and don’t like, even when we’re unable to articulate the reasons why” (Levitin 220-221). In other words, we can understand music even when we lack declarative knowledge about it. Consider the game of billiards. Proficient pool players may not be able to explain the detailed physics equations underlying the game, but they still are able to play the game as if they understand the principles. Similarly, audience members may not be able to explain what makes a speech powerful, but they instantly know whether or not they are moved by the speaker’s message.

The renowned composer, Leonard Bernstein, said that “music does possess the power of expressivity, and the human being does innately possess the capacity to respond to it,” because “in any sense in which music can be considered as a language . . . it is a totally metaphorical language” (“Bernstein at Harvard”). In essence, music is simply speech without words, a symphony of wordless sound. Great music, specifically classical music, encompasses nearly everything that speakers deliver other than the words themselves.¹ But just because music frequently lacks spoken words does not mean that it lacks meaning.

We sometimes recognize that there is a mood or emotion in music. However, this “theory of recognition alone is insufficient, and needs to be supplemented by an account also of arousal, that is, of evoking and communicating” (Allen 59). Music extends beyond sound and into a realm of arousal that has a specific, quantifiable role in speech. As Levitin explains, the perfect combination of music and language can create unparalleled impact:

The multiple reinforcing cues of a good song – rhythm, melody, contour – cause music to stick in our heads . . . as a tool for activation of specific thoughts, music is not as good as
language. As a tool for arousing feelings and emotions, music is better than language.

The combination of the two . . . is the best. (267)

The Techniques Behind Powerful Music

Before a piece of classical music is played in a symphony hall, it resides merely as a collection of unstructured musical elements within a composer’s mind. Individually, these musical elements represent nothing more than a set of unrelated qualities – volume, speed, and tone, for instance – but together, these qualities form the basis of music. Powerful music is represented by both the composer’s ability to layer musical elements in a way that is emotionally evocative and the performer’s ability to execute the composer’s vision.

Radwan proposes that “music presents us with an aural experience, a sonorous development over time, that expresses ideas about human life and emotion rather than propositions about reality or prescriptions for behavior” (4). Music evokes emotion because it mirrors the “main characteristics of emotional behavior, speech, and thought” (Lewis, Barrett, and Haviland-Jones 107). Levitin describes this idea in detail:

Abrupt, short, loud sounds tend to be interpreted . . . as [alert sounds]. Slow onset, long, and quieter sounds tend to be interpreted as calming . . . Composers [are aware of these differences], of course, and use hundreds of subtle shadings of timbre and note length to convey the many different emotional shadings of human experience. (92)

We have identified the five key musical elements that composers leverage to create musical images, emotion-inducing mental constructs shaped by musical elements (Table 1).

Table 1: The Five Key Musical Elements Used to Create Musical Images
<table>
<thead>
<tr>
<th>Element</th>
<th>Definition</th>
<th>Characterization</th>
</tr>
</thead>
</table>
| Tempo       | The general speed of the sound                                              | • Does the music have a fast, frenetic pace or is it very slow?  
• Does the pulse change suddenly?  
• Does the speed fluctuate a lot? |
| Dynamics    | The volume of the sound                                                     | • Is the volume always loud or always soft?  
• Does the volume suddenly become loud and then suddenly soft?  
• Is the change in volume very gradual? |
| Pitch       | The frequency of a sound                                                   | • Is the sound high-pitched like a typical female voice or bird or is it low-pitched like a typical male voice or lion?  
• Are there very high and very low notes?  
• Do pitches change gradually in small steps (as in a musical scale) or do they wildly jump up and down in large intervals? |
| Timbre      | The general quality of sound produced; the “character” or “personality” of a specific sound | • Is the sound rough, biting, sweet, warm, obtrusive, or sonorous?  
• Are certain notes heavily emphasized (accented) more than others?  
• Is there a sharp and alarming emphasis at the beginning of the note or is there a more subtle, dull, and leaning emphasis throughout most of the note’s duration? |
| Rhythm      | The structural spacing of the sounds                                        | • Do the sounds occur in regular and constant time intervals or is there significant variation and unpredictability among the sounds?  
• Do pauses occur suddenly after a fiendish flourish of notes or do the notes slowly ease into a natural break?  
• Are the pauses of significant duration or are they more like short breaths? |

Composers skillfully layer these musical elements on top of one another to form the phrase or cadence of a piece of music. A phrase refers to the general trajectory or contour of a piece. In other words, it addresses the shape and flow of a musical idea and identifies the main themes, the primary voices, and the auxiliary components. The ways in which a performer executes the phrasing of a piece of music will impact how the audience receives and processes the sound.
We have selected and analyzed a variety of classical music pieces to offer deeper insight into how composers use musical elements to evoke specific emotional responses. All five techniques are versatile insofar as they are used in a variety of ways to create dramatically different musical images. As we will demonstrate, there are not necessarily any strict formulaic rules; instead, composers decide how best to employ these techniques based on their particular objective.

The first musical element that we consider is tempo. Fast tempos typically generate excitement, energy, and action. Nikolai Rimsky-Korsakov’s *Flight of the Bumblebee* speeds along at a breakneck pace in order to capture the frenetic motions of a berserk insect. In contrast, slow tempos are often, but not always, used in solemn, deeply reflective, and brooding situations. Richard Strauss’ *Death and Transfiguration* crawls along at a lethargic pace, transporting the listener to a room housing a man on his deathbed.

Choosing a fast or slow tempo is not the only dimension of tempo technique – tempo contrasts are also vital. The lightning pace at the start of Modest Mussorgsky’s *Night on Bald Mountain* produces a turbulent seaside mountain image for listeners. After this first storm surge, a long pause unexpectedly fills the air. This frozen silence creates the feeling of a lull before the storm, which makes the suspense and dread that much more powerful, especially when suddenly, the same quick buzzing heard from the strings at the start is repeated to portend the onslaught of another storm.

The Mussorgsky example demonstrates how a sudden tempo change can capture the attention of listeners and produce maximum suspense. Some composers instead deliberately use a mechanically constant tempo to sustain tension. Gustav Holst’s *Mars* (from *The Planets*) begins at a fast but steady pace, evoking an image of a relentless army destroying everything in its path.
It is the unfailing constancy of the tempo that harbingers a sense of endless dread, which is precisely how Holst intends to characterize the blood-thirsty Roman God of War.

Yet another tempo contrast technique is gradually changing speeds to generate musical momentum in anticipation of a satisfying climax. An excellent example is Ludwig van Beethoven’s *Symphony No. 9 (fifth movement)*, which draws to a close with subtle tempo increases, soaring to one of the most exciting peaks ever attained in classical music.

Effective dynamics and pitch technique parallels proper tempo technique. Although there are no universal rules, loud sounds tend to be startling, alarming, high energy, and grand, whereas soft sounds tend to be sad, brooding, and ominous. Likewise, high pitches tend to be energizing and exciting whereas low pitches tend to be dark and morose. As with tempo technique, however, it is often the contrast in dynamics and pitch that provides the driving force behind powerful music. For example, the variations in pitch (alternating between low and high notes) and dynamics (soft to loud then back to soft) create a sense of unpredictable motion in Arnold Bax’s *November Woods*. These techniques transport the listener to a mysterious world of rustling leaves and howling winds. *Night on Bald Mountain* begins with a soft, foreboding, shimmering sound from the strings, which is then reinforced by thumping notes from low-pitched instruments. Throughout this initial musical statement, there is a rise in volume and then a sudden fall in volume. This dynamic contrast produces a heaving sensation as if waves are washing ashore in a violent fashion before receding and rising again. Similarly, *Mars* features low-pitched brass instruments with a deep, dark drone to state and restate the main theme, which starts soft and gets louder before fading away. Holst’s use of dynamic contrast and instrument timbre creates a menacing, ominous, and otherworldly musical image.
The prior discussion provides a natural segue to timbre technique. In *November Woods*, Bax creates suspense, magic, angst, and restlessness by using high strings, harp, and flutes, which feature thin, airy, ethereal, and rustling timbres. In *Death and Transfiguration*, the use of “sad” sounding instruments such as the bassoon and oboe reinforces the morose and resigned atmosphere. The choice of a shrieking violin in *Flight of the Bumblebee* mimics the buzzing of a pesky insect. And in *Gun Battle* (from *Billy the Kid*), Aaron Copland uses snare drums and trumpets – instruments with a sharp, metallic timbre – to create the essence of whizzing bullets.

Finally, composers use variations in rhythm to achieve their musical goals. *Mars* not only features a constant tempo, but also presents a regimented and steady rhythm repeated *ad nauseam*. This repetition enables the listener to visualize an army of soldiers marching in strict time. *Death and Transfiguration* begins with soft, low-pitched strings playing an uneven rhythmic pattern that is starkly different from Holst’s precise march-like rhythm. This irregular and quiet pattern evokes the image of a man experiencing heart arrhythmias, palpitations, and generally faint and unstable breaths. *Gun Battle* regularly involves pauses of varying lengths, as well as accented notes and quick irregular rhythmic flurries, in order to simulate a gun fight with bullets ricocheting off walls. The rhythmic inconsistency in particular generates a somewhat unpredictable but also exciting atmosphere that keeps the listener on edge and fully alert.

As these examples illustrate, composers use tempo, dynamics, pitch, timbre, and rhythm to construct an image, create a mood, tell a story, and evoke emotion. One need not be an auditory musical expert to observe these techniques since musicians often visually telegraph these musical elements through their body language. Levitin elaborates on this concept:
Studies have shown that nonmusician listeners are exquisitely sensitive to the physical gestures that musicians make. By . . . attending to things like the musician’s arm, shoulder, and torso movements, ordinary listeners can detect a great deal of the expressive intentions of the musician. Add in the sound, and an emergent quality appears – an understanding of the musician’s expressive intentions that goes beyond what was available in the sound or the visual image alone. (210)

While there is no strict formula for applying these techniques, speakers can layer these musical elements on top of one another to reinforce their desired message. In the next section, we discuss how speakers, through careful study and practice, can leverage these musical elements to create powerful musical images.

**Transmitting Musical Images Through Speech**

No two speakers or messages are exactly alike. Just as they use the same 26 letters to create vastly dissimilar content, speakers combine the five key musical elements to convey a variety of emotions. The human voice is a dynamic instrument capable of evoking every possible sentiment, and speakers throughout history have harnessed the music of their voice to transmit distinct and enduring messages. In their own unique ways, professional speakers have leveraged musical elements to achieve a variety of objectives such as inciting political change, calming public outrage, or inspiring a generation.

We have selected and analyzed a variety of speeches to illustrate the ways that professional speakers use musical elements to enhance their vocal delivery. We initiate the discussion with tempo technique and then explore the treatment of dynamics, pitch, timbre, and rhythm technique in well-known speeches.
Mario Savio’s “machine” metaphor at the *Sit-in Address at the Steps of Sproul Hall* is one of the most salient metaphors in public oratory, not only because of its powerful imagery but also because of the speech’s perpetually fast tempo. Savio speaks about the power of non-violent civil disobedience with a tensile rapidity, without dramatic pauses, thought breaks, or a moment to catch his breath. The panting Savio speaks quickly and assertively to generate excitement and inspire his listeners to “put [their] bodies upon the gears and upon the wheels, upon the levers, upon all the apparatus . . . to make [the machine] stop!” This quick tempo enables Savio to harness the energy of a crowd calling for change.

In his 1988 *Democratic National Convention Address*, Jesse Jackson employs tempo to a similar end, but instead varies the speed of his speech. This tempo contrast, similar to the contrast in Mussorgsky’s *Night on Bald Mountain*, generates a powerful feeling of suspense. Jackson begins by presenting the state of racial discrimination in America at a moderately slow pace. This is the status quo: indeed, Jackson factually paints the reality of the day with an ordinary delivery. But soon he begins to increase his tempo to express a passionate desire for racial equality. As his tempo quickens, Jackson seems to plead with the audience while strongly conveying his desire for a better tomorrow. Jackson’s use of tempo contrast emphasizes the temporal difference between the present and the future and enables him to build to a resounding climax. Jackson’s conclusion is not unlike the end of Beethoven’s *Symphony No. 9 (fifth movement)*, which makes use of tempo contrasts to build to an imposing finale.

Like tempo, dynamics dramatically influence the mood and excitement level. Speakers can use dynamics to excite or surprise, but they can also use dynamics to soothe and calm. Former President Bill Clinton illustrates the use of dynamics to assuage an audience in his 1995
Oklahoma Bombing Memorial Prayer Service Address when he speaks to a bereaving crowd and country four days after the bombing. Unlike Savio or Jackson, Clinton does not speak with a desire to inspire action and change; he speaks with a desire to comfort and reassure. To accomplish his objective, Clinton uses a particular dynamic pattern throughout his speech. At the beginning of each sentence, Clinton’s voice is strong and appropriately loud. As each sentence concludes, his voice becomes softer and tapers downward. This repeated dynamic pattern is congruent with the somber mood and the main themes of the speech – grief, acceptance, unity, and justice. By effectively using dynamics, Clinton succeeds at convincing the audience that he is taking the matter seriously while simultaneously encouraging the audience to move from thoughts of anger and hate to thoughts of acceptance and unity.

In contrast to Clinton, Mohandas Gandhi, in his 1931 Address at Kingsley Hall, focuses more on the element of pitch. Declaring his never-ending belief in a higher power, Gandhi generally maintains a constant tempo, rhythm, and timbre throughout his speech, while only subtly altering his dynamics. The most noticeable change, then, occurs as he modulates the pitch of his voice. By focusing almost exclusively on pitch, his voice lacks any semblance of pride and passion; instead, he speaks in a blunt, steady manner that conveys his firm and unalterable belief in God. Had Gandhi not altered his pitch from phrase to phrase, his speech would have been too bland and robotic; by varying just pitch, he effectively conveys his resolve and reverence. Baroque concertos – Johann Sebastian Bach’s Brandenburg Concertos with their methodical, almost mathematically precise form, for example – deliver some of the same images: strength, respect, and tenacity.
In her 1992 *Democratic National Convention Address*, Elizabeth Glaser takes advantage of the natural timbre of her voice to deliver an emotional message about living with AIDS. In her address, Glaser uses a natural and noticeable timbre to evoke a variety of emotional responses. Although dynamics and pitch are heavily involved in her speech, she carefully alters the timbre of her voice when telling stories about AIDS victims (sadness), discussing the ostracism of people with AIDS (anger), and delineating the need for a leader committed to addressing the AIDS crisis (pleading). Unlike the slight Southern drawl of Bill Clinton’s voice, Glaser’s vocal timbre is firm and deliberate – as is the conviction and passion that she demonstrates throughout her speech.

Although every speaker naturally employs some rhythmic technique (whether evenly-paced with similar tone and pause length, unevenly-paced with differing lengths, or some combination of the two), Tom Hanks, in his 2005 *Commencement Address at Vassar College*, carefully manages his rhythm to deliver a powerful message. Like Strauss’ *Death and Transfiguration*, Hanks’ rhythmic pattern constantly changes and forces the listener to hang on to each phrase and wait in anticipation for the next one. Hanks’ rhythm is steady and constant throughout most of the speech, but when he wants to emphasize a crucial point – such as removing four cars to eliminate gridlock, a powerful element of his “power of four” theme – his words become unevenly spaced with different length pauses separating each word. The changing rhythm adds significant prominence to this particular message. When he finishes the point, he reverts to his original rhythm – that is, until he reaches another crucial point.

Hillary Rodham Clinton, in her 1995 *Remarks to the U.N. 4th World Conference on Women*, further illustrates rhythmic variations but also combines these variations with distinct timbre and
pitch. Clinton emphasizes the importance of human rights for women through the rhythmic repetition of the phrase, “It is a violation of human rights,” but she also raises her pitch when she says the word, “human,” to draw attention to her theme that “human rights are women’s rights.” She makes this point with a hard and confident timbre, purposefully over-accentuating her words. Clinton’s blatant, almost staccato delivery makes her message even more meaningful as she explains that these issues must be heard “loudly” and “clearly” and that “we must move beyond rhetoric.”

As these examples illustrate, professional speakers deliberately use tempo, dynamics, pitch, timbre, and rhythm to elicit specific emotional responses. While a speaker may focus on a single musical element, he or she can also layer these elements on top of one another to form a powerful musical collage. Speakers who strategically layer musical elements transmit powerful musical images that are holistically much stronger than the sum of their various parts.

### Conclusion

Composers use a core set of building blocks – tempo, dynamics, pitch, timbre, and rhythm – to craft powerful, memorable pieces of music that evoke specific emotional responses. In this paper, we have introduced the five key building blocks and have explained how speakers can use these musical elements to create musical images that leave a lasting impact on their listeners.

Although many of the musical techniques that we have discussed may seem subtle and nuanced, they all serve vital purposes. As in the case of complex birdsong, listeners may not be able to immediately recognize or parse all of the subtle aspects involved in an effective speech, but every aspect is still relevant and has an immeasurable impact on the listener’s emotions and general impression.
It is our hope that speakers use our framework as a starting point to design and deliver powerful and meaningful messages. The goal of this paper is not to provide a formulaic playbook for speech success; indeed, speech delivery is an art form that requires extensive preparation and practice. Instead, our primary purpose is to stimulate discussion about musical elements and how speakers can leverage these elements to craft speeches that achieve a specific objective. Although we do not think that a rigid guidebook is very practical, we hope that our paper encourages communications scholars to study the music-speech analogy in more depth and design empirical studies to evaluate the effectiveness of using particular combinations of musical techniques in speech.
Notes

1. In practice, many forms of music include words, such as choral and operatic works. For the purposes of our music-speech analogy, however, we refer to musical selections that lack words – specifically classical music selections – to make it easier to focus on the musical elements rather than on the lyrics.

2. See Appendix A for a list of complete audio recordings of the classical music selections referenced in our paper. All music selections are also accessible via YouTube.

3. Appendix B provides a set of guided exercises to help readers understand what musical elements are and how composers vary these elements to elicit different emotional responses.

4. Text and video recordings of these speeches are accessible via the web links listed in Appendix C.
Works Cited


Appendix A: Classical Music Examples

Telarc Records, 1993. CD.


Appendix B: The Five Key Musical Elements Interactively Illustrated

We encourage readers to use a virtual keyboard (http://www.bgfl.org/bgfl/custom/resources_ftp/client_ftp/ks2/music/piano/index.htm) to explore how composers leverage the five key musical elements to elicit specific emotional responses.

The first five rows of the table below highlight the use and variation of each of the five key musical elements, holding all other elements constant. The final row of the table provides an illustration of how musical elements are layered on top of one another.

We hope that these exercises clarify the fundamental building blocks that speakers can use to create musical images that leave a lasting impact on their listeners.

<table>
<thead>
<tr>
<th>Element</th>
<th>Exercise 1</th>
<th>Exercise 2</th>
<th>Explanation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tempo</td>
<td>Press A, A, A, A. Wait 1 second between each key press.</td>
<td>Press A, A, A, A. Wait 0.5 seconds between each key press.</td>
<td>Holding all other elements constant, exercise 2 has a faster tempo than exercise 1.</td>
</tr>
<tr>
<td>Dynamics</td>
<td>Press A.</td>
<td>Increase speaker volume, then press A.</td>
<td>Holding all other elements constant, exercise 2 has a louder dynamic level than exercise 1.</td>
</tr>
<tr>
<td>Pitch</td>
<td>Press C.</td>
<td>Press F1.</td>
<td>Holding all other elements constant, exercise 2 has a higher pitch than exercise 1.</td>
</tr>
<tr>
<td>Timbre</td>
<td>Press A.</td>
<td>Click on the “pan pipes” button, then press A.</td>
<td>Holding all other elements constant, the character of the sound is different between exercise 2 (airy and soothing) and exercise 1 (sharp, pinging, and metallic).</td>
</tr>
<tr>
<td>Rhythm</td>
<td>Press A, A, A. Wait 1 second between each key press.</td>
<td>Press A. Wait 2 seconds, then press A, A, A. Wait 0.5 seconds between each key press.</td>
<td>Holding all other elements constant, exercise 2 has a more irregular rhythm than exercise 1.</td>
</tr>
<tr>
<td>Layered</td>
<td>Click on the “piano” button, then press C, C, G,</td>
<td>Click on the “pan pipes” button, decrease speaker</td>
<td>Relative to exercise 1, exercise 2 has a different timbre, lower dynamic level, higher pitch,</td>
</tr>
<tr>
<td>G, A, A, G, F, F, E, E, D, D, C. Wait 0.5 seconds between each key press.</td>
<td>volume, then press G, G, D1, D1, E1, E1, D1, C1, C1, B, B, A, A, G. Wait 1 second between each key press.</td>
<td>slower tempo, and the same rhythm. The mutually reinforced layering of elements makes exercise 2 more soothing and lullaby-esque, while exercise 1 is more sprightly, whimsical, and dance-like.</td>
<td></td>
</tr>
</tbody>
</table>
Appendix C: Speech Examples


