Spring 2012

The Impact of Vowels on Pitch Finding and Intonation in the Movable-Do Solmization System

Thomas Neuerburg
University of Nebraska-Lincoln, tnburg@gmail.com

http://digitalcommons.unl.edu/musicstudent/51

This Article is brought to you for free and open access by the Music, School of at DigitalCommons@University of Nebraska - Lincoln. It has been accepted for inclusion in Student Research, Creative Activity, and Performance - School of Music by an authorized administrator of DigitalCommons@University of Nebraska - Lincoln.
THE IMPACT OF VOWELS ON PITCH FINDING AND INTONATION
IN THE MOVABLE-DO SOLMIZATION SYSTEM

by

Thomas Neuerburg

A THESIS

Presented to the Faculty of
The Graduate College at the University of Nebraska
In Partial Fulfillment of Requirements
For the Degree of Master of Music

Major: Music

Under the Supervision of Professor Stanley Kleppinger

Lincoln, Nebraska

May, 2012
THE IMPACT OF VOWELS ON PITCH FINDING AND INTONATION
IN THE MOVABLE-DO SOLMIZATION SYSTEM

Thomas Robert Neuerburg, M.M.

University of Nebraska, 2012

Adviser: Stanley Kleppinger

The modern movable-Do solmization system based on syllables devised by Guido d'Arezzo was modified and pieced together over centuries by various scholars and pedagogues, each with their own rationale. To date, considerations of the movable-Do system have not sufficiently examined the effects of the vowels contained within its solfège syllables. While vowels have been thoroughly analyzed among vocal pedagogues, that information has not been adequately transferred to the realm of aural theory. Individual vowels contain perceptual qualities and intonational tendencies, due to their physiological articulation and acoustic properties. This document relates vowel characteristics with the solfège syllables used in the movable-Do solmization system, and explores potential implications contained therein.
Author’s Acknowledgements

This thesis would not have been possible without Mrs. Lois Nassen, who provided my musical foundation; Dr. Donald Simonson, who taught me to understand my voice; Dr. James Rodde, who inspired my love of choral music; and Dr. Jeffrey Prater, who fostered my passion and deep appreciation for aural skills by being a teacher, mentor, and friend.

Above all, I would like to thank Dr. Stanley Kleppinger for his patience, guidance, and unwavering enthusiasm during the creation of this thesis.
Table of Contents

Abstract......................................................................................................................... ii
Author’s Acknowledgements......................................................................................... iii
List of Multimedia Objects.......................................................................................... v

Chapter 1: Literature Review and Introduction............................................................... 1
Chapter 2: Vowel Physiology and Pronunciation............................................................ 7
Chapter 3: General Physiological Effects of Vowels and Diphthongs on Timbre and Intonation ................................................................. 13
Chapter 4: Diphthongs and Vowel Unification............................................................... 20
Chapter 5: The Acoustic and Perceptual Qualities of Syllables Found in Solfège......................... 27
Chapter 6: The Vowels Used in Solfège and Their Individual Pitch Tendencies............................ 37
Chapter 7: Application in a Choral Setting.................................................................... 45
Chapter 8: Avenues for Further Research...................................................................... 53

Reference List.................................................................................................................. 55
List of Multimedia Objects

Figure 1: The physiological processes of speech and singing………………………………………7
Figure 2: The vowel trapezoid…………………………………………………………………………8
Figure 3: IPA symbols and their pronunciation…………………………………………………..9
Figure 4a: Placement of [i]………………………………………………………………………………….16
Figure 4b: Placement of [e]………………………………………………………………………………….16
Figure 4c: Placement of [ɔ]………………………………………………………………………………….16
Figure 4d: Placement of [o]………………………………………………………………………………….16
Figure 5: Different pronunciations of sung [eɪ]………………………………………………………21
Figure 6: Different pronunciations of sung [oo]………………………………………………………21
Figure 7: Frequency of F₁ and F₂ by vowel…………………………………………………………30
Figure 8: The relationship between Mels (Y-axis) and kilohertz (X-axis)……………………32
Figure 9: American English vowels based on perception………………………………………..33
Figure 10: Summary of vowel characteristics………………………………………………………34
Figure 11: Summary of vowel impact on pitch……………………………………………………….44
Figure 12: Chorale from *Ein feste Burg ist unser Gott*…………………………………46-47
Chapter 1: Literature Review and Introduction

This document was inspired by connecting two areas of the author’s educational background. One aspect was introduced through choral and solo vocal training, the core premise of which is that sung vowels contain inherent intonational tendencies and perceptual qualities due to their physiological generation and acoustic properties. The second main consideration concerns the movable-Do solmization system, and its role as a tool used to assist musicians with accuracy of pitch and intonation. By applying the vowel tendencies described in vocal literature, this document explores the impact vowels have on those singing the solfège syllables used in movable-Do.

Because this thesis combines two commonly separated musical fields of study, extant research is largely split into two categories: vocal production and solmization. Voice teachers and choral conductors often use solfège in their respective positions. Likewise, those using solmization systems are required to sing. The same instructions given to singers in the interest of healthy and accurate vocal production apply to musicians employing solfège. Despite this, there seems to be little extension of vocal pedagogy toward the use of solfège. The following literature review will explore some of the most necessary texts for “bridging the gap” between these two topics, and illuminate why this document fulfills a role that has thus far been missing.

The physiological generation and acoustic properties of vowels have been analyzed extensively in literature relating to speech (Delattre 1951; Denes 1973; Lieberman 1977; Lieberman and Blumstein 1988; Paget 1976; Titze 1993; Titze and
Scherer 1983). These authors connect raw acoustic data with physiological articulation to illuminate how vowels are perceived and differentiated.

Vocal pedagogy and diction texts intended for solo singers vary in their approach and focus, but share the intent of fostering healthy vocal production, proper diction, and resonant, beautiful, accurate tone. Although significant overlap occurs, some texts heavily emphasize physiology, often basing their methods and data on speech literature (Appelman 1967; Coffin 1976, 1980; Doscher 1988; Sundberg 1977, 1987; Taylor 1908; Vennard 1957; Ware 1998). Others rely primarily on mental imagery, musical examples, and archetypes of pronunciation (Christy 1965; Fracht 1978; Marshall 1953; Moriarty 1975; Wall 2005). In these texts, solmization, if it appears at all, is only mentioned in passing.

Literature on choral conducting and pedagogy approaches vowel tendencies differently than texts intended for soloists, due to the desire for unification of pitch and diction across multiple voices. Specifically, they illuminate many potential problems in intonation, including those arising directly from improperly produced vowels (Haaseman and Jordan 1991; Hammar 1984; Hylton 1995; Jordan 1996; Powell 1991). Some even recommend some vowels over others during warm-ups, for timbral and intonational reasons (Haaseman and Jordan 1991; Hylton 1995; Jordan 1996). In correlation, Sundberg (1987, 134-45) mentions several studies that analyze the effects of vowels on tuning in a choral setting. Steven Demorest (2001, 37-58) explores the merits of various solmization systems in a choral setting, but does so from a functional standpoint without mentioning the relative syllabic advantages on timbre or intonation. In every case,
however, solmization is either unmentioned, or is conspicuously disconnected from the treatment of intonational effects arising from syllabic content.

Solmization systems are used by musicians with extremely diverse backgrounds and specialties. Sometimes solfège syllables are introduced as part of the beginning music curriculum in school systems. Some high school or college choirs use it during the reading process of a new piece. Professional singers familiar with solfège may utilize it in rehearsal for their own accuracy.

Experienced singers routinely engage in vowel modification depending on circumstance and register, in both solo and choral contexts.¹ The extent to which both intonation and vowel production are emphasized in vocal and choral pedagogy acclimates singers to the modification process. With practice, the sensitivity and physical adjustments required to accurately sing various vowels across the entire vocal range can become habitual and comfortable. In other words, the processes involved with correctly singing a given pitch or phrase in tune regardless of the vowels involved becomes easier with experience.

Naturally, musicians who aren’t as comfortable using their voices are at a comparative disadvantage in terms of accuracy. For those with musical training who lack vocal experience, pitch finding problems don’t always lie with mental processes. The extra attention required for an inexperienced singer to control his or her voice detracts from the overall fluidity of accurate vocalization. In the same way as those first learning a solfège system spend a greater portion of their mental energy remembering the correct

¹ Vowel modification relating to solo singing can be found in Coffin 1980, Wall 2005, and Ware 1998. Vowel modification in choral settings can be found in Haasemann and Jordan 1991, and Jordan 1996.
sylables to use than those experienced with that system, inexperienced singers do not have the same reflexive good habits as those more vocally educated and comfortable.

To isolate and study the individual procedures necessary for singing, it is important to distinguish between the mental processes involved with imagining a pitch before generating sound, and the physiological mechanisms required to produce that sound. The ability to mentally “hear,” even with no sound present, is termed audiation. Musicians of all types are regularly required to audiate in some capacity. Conductors and soloists may think through a piece before a performance, to mentally prepare the correct tempo and style. Music students being tested on dictation may recall a recently heard musical excerpt. Choral singers may mentally establish their beginning pitches before singing to ensure readiness and accuracy. All of these are examples of audiation, which is an essential component of musical competency. Unfortunately, auditory ability does not equal singing prowess; musicians who are skilled at audiation are not necessarily vocally proficient enough to accurately represent their aural sagacity.

Aural skills texts, even those with an emphasis on sight singing, focus primarily upon the mental aspect of ear training, while omitting any detailed instruction on vocal production (Benjamin 1994; Carr 1991; Gottschalk 1997; Horacek 1989; Chosky 1999; Karpinski 2000; Lieberman 1959; Ottman 1996; Rogers 2004). Gary S. Karpinski (2000, 145-46) does emphasize the importance of proper vocal production, saying the goal is to make the voice “a tool, not an obstacle.” Michael R. Rogers (2004, 126-28) even recognizes the near-universal dependence on singing within aural skills curricula, and proposes alternatives for those with vocal impediments. Even in these cases, though,
there is no mention of solmization issues resulting from poor singing technique. Of course, the argument can be made that vocal pedagogy lies outside the domain of ear training manuals. However, the sheer amount singing required by those using aural skills, as a means of evaluation or otherwise, indicates that at least some baseline of vocal competency should be associated with ear training.

Even experienced singers who are intensely focused on the technical aspects of a piece, i.e. notes and rhythms, may revert to habitual but improper singing techniques (Hylton 1995, 71; Jordan 1996, 280-81). In such cases, the attention paid to breath support, tone, and vowel modification may suffer. Depending on the singer, the sound produced during the learning process may more closely resemble that of an amateur vocalist than it will as the singer becomes comfortable with the piece (Jordan 1996, 286-87). When not concentrating on proper vocal production due to either inexperience or inattention, singers become more susceptible to intonation issues that arise from problematic physiological elements, including those caused by vowel tendencies (Ware 1998, 179; Hylton 1995, 70-72; Hammar 1984, 112). Because solfège is commonly used when these issues are in effect, the treatment of vowels in this thesis is structured to reflect similar conditions.

In chapter 2, I will give an introduction to the physiological mechanisms involved with vocalization and vowel articulation, as well as the method by which vowels are organized, codified, and labeled. Preliminary knowledge of vocal physiology will assist the reader by establishing terminology and generating familiarity with vowel formation.

Chapter 3 examines how vocal literature approaches the articulation of vowels in
terms of resonance placement, and the effect that has on intonation. Additionally, it describes how choral practices reflect the intonational tendencies present in vowels through the usage and avoidance of certain vowels during vocalization warm-ups. Finally, the vowel tendencies are related to solfège, due to the similar goals of solmization and choral warm-ups.

Chapter 4 expands on the previous chapter by exploring vowels directly adjacent to each other in diphthongs. The destabilizing effects of diphthongs are related to issues of vocal clarity, unification, and perception before being applied directly to the diphthong [oo] found in Do and Sol.

In chapter 5, I outline the acoustic features upon which humans rely to differentiate between vowels in speech and singing. By relating raw data to how we interpret pitches, I explain how some vowels are naturally perceived as higher or lower.

Chapter 6 discusses the scale-degree functionality associated with solfège syllables in movable-Do, and consolidates the information presented in chapters 3 through 6 in figure 11.

The application of vowel tendencies on intonation and pitch finding is applied to a choral setting in chapter 7 by highlighting points at which those tendencies would either exacerbate or mitigate difficult passages. Additionally, I explain how many of the intonational issues caused by the vowels used in solfège syllables could be avoided through the implementation of three broad, easily communicated vocal guidelines.

In chapter 8 I offer avenues for further study built on the information collected for this thesis, and questions that arose in the process its creation.
Chapter 2: Vowel Physiology and Pronunciation

This chapter begins with an introduction to the physiology of vocalization and articulation, which serves to establish terminology. Following that is an explanation regarding the methods by which vowels are distinguished, organized, and represented as IPA symbols. Finally, the solfège syllables found in the movable-Do system will be associated with their respective vowels, along with a survey of those syllables’ potential varying pronunciations.

The human voice, like any other wind instrument, can be analyzed in terms of its mechanical elements and their functions. As shown in figure 1, there are three essential components to phonation. The lungs act as a compressor, which produce air flow. The vocal folds act as oscillators, which vibrate to induce pitch. The vocal tract refers to the interior and exterior of the mouth and nasal cavity, which together are used for

![Diagram of phonation process](image)

Figure 1. The physiological processes of speech and singing (from Sundberg 1987, 10)
articulation. Although there is extensive connectivity and interdependence between each stage of phonation, this document will primarily focus on the vocal tract, which is responsible for creating distinct vowel sounds.

Figure 2 shows a common organization of vowels based on their formation. The figure below, or something similar, appears in multiple texts concerning both speech and singing (Lieberman 1988, 164; Delattre 1951, 866; Coffin 1980, 10; Moriarty 1975, 18; Wall 2005, 15). This so-called “vowel trapezoid” can be read starting at the top left, moving counterclockwise around the outside of the trapezoid. The letters encased in brackets, such as [i], are symbols from the International Phonetic Alphabet, or IPA. A concise table for interpreting these symbols has been provided in figure 3.

![Vowel Trapezoid Diagram](image)

**Figure 2. The vowel trapezoid (from Ware 1998, 160)**

The vowels shown in figure 2 are organized according to tongue and lip position. The left side of the trapezoid, from [i] to [a], contains vowels that are modified by the tongue. This is easy to demonstrate by speaking them, gradually flowing from one vowel to the next counterclockwise along the trapezoid’s left edge. The resulting series of vowels, as shown in Figure 3, is as follows: [i] as in “seat,” [e] as in “day,” [ɛ] as in
“get,” [æ] as in “back,” and [a] as in “father,” spoken brightly. As can be observed when performing this exercise, when intoning [i] the tongue is placed very close to the hard palate, more colloquially known as the roof of the mouth. As the speaker smoothly shifts through [e], [ɛ], [æ], and finally [a], the tongue gradually lowers to the bottom of the mouth. The vowels are thus organized in terms of the amount of space between the tongue and the hard palate, and can be classified in terms of how “open” or “closed” they are. Referring again to figure 2, it is now possible to see that the vowels along the left side of the trapezoid are ordered by the increasing space between the tongue and the hard palate. Throughout the series of these “tongue” vowels, the lips remain inactive.

<table>
<thead>
<tr>
<th>Basic Phonemes</th>
<th>IPA Symbol</th>
<th>English</th>
<th>Italian</th>
<th>German</th>
<th>French</th>
</tr>
</thead>
<tbody>
<tr>
<td>ee</td>
<td>[i]</td>
<td>seat</td>
<td>si</td>
<td>sie</td>
<td>hiver</td>
</tr>
<tr>
<td>ih</td>
<td>[ɪ]</td>
<td>sit</td>
<td>immer</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ay</td>
<td>[ɛ]</td>
<td>day</td>
<td>vero</td>
<td>leben</td>
<td>été</td>
</tr>
<tr>
<td>eh</td>
<td>[ɛ]</td>
<td>get</td>
<td>belle</td>
<td>denn</td>
<td>clair</td>
</tr>
<tr>
<td>ah</td>
<td>[a]</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>aw</td>
<td>[ɔ]</td>
<td>father</td>
<td>casa</td>
<td>Mann</td>
<td>ras</td>
</tr>
<tr>
<td>uh</td>
<td>[ʌ]</td>
<td>shirt</td>
<td></td>
<td></td>
<td>folie</td>
</tr>
<tr>
<td>oh</td>
<td>[o]</td>
<td>tone</td>
<td>dove</td>
<td>so</td>
<td>hôtel</td>
</tr>
<tr>
<td>oo</td>
<td>[u]</td>
<td>look</td>
<td>sua</td>
<td>zu</td>
<td>boule</td>
</tr>
</tbody>
</table>

Figure 3. IPA symbols and their pronunciation (from Ware 1998, 157)

The trapezoid’s right edge contains vowels which follow a similar structure as those just discussed. Instead of tongue position, however, the vowels [ɒ], [ɔ], [o], and [u] are arranged by lip position. With all of these vowels, tongue placement is not the major determining physical condition. In every case, proper production can result with the
tongue relaxed and inactive at the bottom of the mouth. Instead of tongue position, the primary differentiating feature between these vowels is the openness and roundness of the lips. A similar exercise as that concerning the “tongue” vowels above can be used to demonstrate this, by smoothly intoning the vowels in order without pause. By beginning with the [u] found in “soothe” and slowly opening the lips while maintaining a wide space between the tongue and hard palate, the transition through the series of “lip” vowels is shown.

In summary, vowels are organized in figure 2 in terms of how “open” or “closed” they are, based on tongue and lip placement. Vowels that lie lower on the trapezoid are more “open,” while those positioned higher are more “closed.” At the very bottom center of the trapezoid lies [ɑ]. As might be inferred from Figure 2’s organization, [ɑ] is formed from both fully open tongue and lip positions.

In the movable-Do solmization system, only four distinct vowels or vowel clusters are used. They are as follows:

- [o] or [ou]: Do and Sol
- [a] or [a]: Ra, Fa, and La
- [i]: Di, Ri, Mi, Fi, Si, Li, and Ti
- [e] or [ei]: Re, Me, Se, Le, and Te

The precise pronunciation of these syllables is varied, somewhat complicating matters. The syllable [o], for example, often glides into [o] among American English speakers (Wall 2005, 64). The result is a diphthong, an unbroken transition between two distinct vowel sounds. The diphthong [ou] can be heard in words such as row and bone. Instead of speaking a pure [o] vowel, the tendency is to instead say “oh-oo.” This can be demonstrated by saying “row” very slowly. Instead of maintaining consistent lip position
during the vowel, American English speakers will tend to reduce the space between their lips toward the end of that vowel.

Similarly, the [ɛɪ] diphthong, found in “bait” or “weigh,” is, in the words of Joan Wall (2005, 28), “consistently used in place of the pure [e], and may be considered an allophone [a sound considered to be the equivalent of another] of [e].” Slowly pronouncing the words “bait” or “weigh” demonstrates the propensity for American English speakers to say “eh-ee” in place of a pure [e] vowel\(^2\).

Finally, the distinction between [a] and [ɑ] is also shown due to American speech patterns. In some languages, such as Italian and German, the “ah” vowel is generally spoken brightly, like a sigh. Americans, however, are inclined to pronounce words containing the “ah” sound gutturally, causing “ah” to more closely resemble the darker [ɑ] (Wall 2005, 47; Jordan 1996, 287). The [a] vowel resembles the sound produced when saying the phrase “park the car” using a Boston accent, whereas [ɑ] is the standard American “ah” used in words such as “water.”

In light of American English speakers’ characteristics above, the list of practically applied vowels in the movable-Do solfège system is as follows:

- [oo]: Do and Sol
- [ɑ]: Ra, Fa, and La
- [i]: Di, Ri, Mi, Fi, Si, Li, and Ti
- [ɛɪ]: Re, Me, Se, Le, and Te

---

\(^2\) Some singing diction texts, including later versions of Joan Wall’s own book, use [ɛɪ] in place of [ɛɪ]. To the best of the author’s knowledge, this adoption of [ɛɪ] arose from sensitivity to a process called “vowel elongation” employed in singing. While singing words containing [ɛɪ], such as “wait,” the [ɛ] vowel, as in “wet,” is held for most of the duration, and the singer doesn’t glide to [i] until toward the end of the word. Speakers, however, seem to be more accurately represented by [ɛɪ]. More information can be found in Marshall 1953.
The extent to which those using solfège syllables conform to the four vowels and diphthongs listed above will vary on an individual basis. For two reasons, this document will only consider [oʊ], [ɑ], [i], and [eɪ]. First, these four vowels and diphthongs above are more commonly used in American English speakers than their eliminated counterparts. Trained singers routinely modify vowels, which reduces their vulnerability to those vowels’ intrinsic differentiating intonational qualities and spoken idiosyncrasies. Less experienced singers, however, are not accustomed to changing vowel shapes when singing, and will thus more closely reflect spoken pronunciation (Jordan 1996, 286-87). Second, [oʊ], [ɑ], and [eɪ] pose more intense potential problems than [o], [a], and [e], as will be shown later. By focusing on pronunciations with more extreme intonational problems, the consequent pitch finding and intonational issues come to the fore with greater clarity.

In this chapter, terminology concerning the physiological generation, organization, and IPA representation of vowels has been introduced. Additionally, the four vowels or diphthongs used in movable-Do have been established based on American speaking habits. The next two chapters explore what intonational tendencies are associated with the physiological articulation of these vowels, and how vocal pedagogues facilitate accurate and healthy production.
Chapter 3: General Physiological Effects of Vowels and Diphthongs on Timbre and Intonation

Among pedagogical texts directed toward solo singers, there seems to be little literature directly connecting vowel pronunciation with intonation. Instead, instruction on this topic tends to concentrate on pronunciation, tone, and healthy production of vowels (Coffin 1980; Marshall 1953; Wall 2005; Christy 1965). Proper vocalization and tone production, however, are intimately linked with intonation, even among experienced singers (Ware 1998, 108, 179). The goal of maintaining proper pronunciation and physical technique, therefore, directly aids intonational accuracy, even if it is not often explicitly discussed in many solo singing texts.

Choral directors tend to discuss the connection between vowels and intonation more frequently, possibly due to the desire for a unified tone that allows for clarity of pitch and diction (Smith 2000, 139; Hylton 1995, 21-28; Jordan 1996, 280-82; Hammar 1984, 74-75). Current pedagogy emphasizes unifying vowel sounds within and across voice sections. Part of the attention given to vowel formation involves intonation.

Intonational problems pertaining to vowel production more commonly address flatting than sharpening. Even when not discussing vowel production, however, choral procedures tend to discuss intonation in terms of preventing or rectifying flatting (Hylton 1995, 70-72; Jordan 1996, 289-95; Hammar 1984, 112). Some of these problems are attributed directly to physical vocal production, such as improper posture and breathing, fatigue, or singing in an uncomfortable range (Hylton 1995, 70-72; Hammar 1984, 112). John B. Hylton (1995, 71) mentions that a choir tends to sing under pitch while first
learning a piece, and makes the following claim: “As the choir members learn their individual lines more solidly, their ability to sing in tune will improve.” By connecting the impact of physiological vocal production with the tendency to sing under pitch when unfamiliar with the music, it is possible to see how one can influence the other. That is, a singer who is devoting significant attention to merely singing the correct notes may be less focused on the proper physical production of sound. The resulting poor vocal technique can negatively affect the singers’ intonation.

With such an array of potential physiological and external causes for tuning difficulties, it is necessary to examine elements individually to appropriately study their effects. In the case of vowels, this has been accomplished through the use of individual syllables during warm-ups and reading. One such exercise is described by James Jordan (1996, 286-87):

Chant the text one syllable per pulse on a static whole tone chord, e.g., E-F#-G#-A#, with basses on the E, tenors on the F#, altos on the G#, and sopranos on the A#. By chanting on this static chord, the singers can hear which vowel sounds tend to go “out of tune” and make immediate adjustments to the color of the vowel. Remember that pitch problems are, for the most part, caused by vowel color problems and not “wrong pitches.”

One key concept, from a physiological perspective, for singing both healthily and in tune is that of vowel “placement” (Hammar 1984, 81-87; Jordan 1996, 282-85; Hylton 1995, 15, 22-23). The term “placement,” in this context, refers to the area of resonance being employed by the singer. As a demonstration of how one vowel or consonant can differ in its placement, one can slowly say the word “little,” with close attention to the two “L” sounds used in this word. The first L tends to be pronounced dentally, with the
tip of the tongue close to the front teeth, while maintaining a wide space between the rest of the tongue and the hard palate. This dental pronunciation promotes resonance in the front of the mouth, which can be proven by noticing which area of the mouth is vibrating. In contrast, the second L in “little” is generally pronounced less brightly. The tongue’s tip tends to be further back, and there is less space between the tongue and the roof of the mouth. As a consequence, the resonant area is placed far back in the mouth. By switching between the first and second L sounds in “little,” it is possible to notice the difference in resonance location. The first L is “placed” forward, and can be described as being a “bright” L. The second L is “darker” than its counterpart, and is “placed” much farther back, almost reaching the throat. The same concept can be applied to vowels, and can be demonstrated similarly, albeit sometimes in a less obvious manner. One further exercise is saying the [ɑ] vowel, as in “water,” while frowning and smiling. Saying [ɑ] while smiling promotes a bright, forwardly placed sound, while frowning causes dark, rear placement.

Figures 4a through 4d show one author’s representation of “correct” and “incorrect” placement, as represented by the light and dark areas. Note that the shaded dark areas represent the “incorrect” placement of resonance, not physiological structures. “Correct” centers of resonance are represented by the hollow oval shapes indicated by the arrows. A common feature across all four illustrated vowels is the relative positions of “incorrect” and “correct” vowels. In general, “correct” vowel placement results in resonance that is “high [and] forward,” near the top teeth (Hammar 1984, 75). “Incorrect” placement occurs farther back in the mouth, nasal cavity, and throat, in differing degrees,
Figure 4a. Placement of [i]  
Figure 4b. Placement of [e]  
Figure 4c. Placement of [a]  
Figure 4d. Placement of [o]  

(from Hammar 1984, 81-84)
among all vowels shown in figures 4a through 4d. Many vocal pedagogues espouse the use of mental imagery to promote desirable vowel placement, such as imagining the vowel resonating within the singer’s forehead, or focusing the tone to a point in front of one’s face (Vennard 1957, 68; Hammar 1984, 63-64; Haasemann 1991, 69-70; Ware 150-53, Hylton 1995, 15; Wormhoudt 1991, 39-40). Some vowels facilitate proper forward placement more naturally than others, which, from a choral perspective, makes them more desirable for use in reading and exercises (Jordan 1996, 281-87; Hylton 1995, 10-26).

A study described by Johan Sundberg (1987, 143-44) tested the ability of singers to steadily sing a pitch while shifting from one vowel to another. The results were then charted according to whether the vowel change raised, lowered, or had no effect on the singer’s pitch. On average, shifting from a closed vowel (based on either tongue or lip position) to a more open vowel resulted in flatting the pitch. Likewise, shifting from an open vowel to a more closed vowel, on average, caused the pitch to rise. Vowels with similar degrees of openness resulted in lesser change than those with more drastic differences in openness.

Although this study is far from comprehensive and is not graphically represented to allow for interpretive precision, the general effects it yields are reflected by choral pedagogy texts. James Jordan (1996, 282-83) strongly advocates the use of [u], [i], and [y] in warm-ups and reading, which also happen to be the three most closed vowels. The rationale behind [u] is its propensity for natural forward placement, while [i] is chosen for
both its “brilliance” and its forward placement. According to Jordan, [y], which can be formed by using the tongue position of [i] with the lip position of [u], is “…the most desirable vowel for most circumstances because it combines the best qualities of the above vowels” (Jordan 1996, 283).

Frauke Haasemann and James Jordan (1991, 33, 70, 110) employ visual imagery to modify vowels toward either [i] or [u] for resonance and intonational purposes. By instructing singers to use a “fishmouth,” Haasemann and Jordan (1991, 70) are in fact modifying their pronunciation to more closely resemble [u]. Its counterpart, “rabbit teeth,” helps singers “brighten” their sound by modifying the vowel toward [i]. Haasemann and Jordan (1991, 110) further reinforce the notion of vowels affecting intonation by saying, “It is best not to start with the ‘ah’ or ‘eh’ vowel because those vowels tend to be too open and flat, and to exert a bad influence on the other vowels.”

Although the varied connections among vowel openness, tone placement, and intonation are not clearly and reliably described, enough literature exists to tentatively accept the relationship between vowel openness and intonational tendencies. Specifically, closed lip vowels promote intonational accuracy due to their natural proclivity for proper placement of resonance. Closed tongue vowels also promote forward resonance, albeit to a lesser degree than closed lip vowels, and are additionally chosen for their perceived “brightness,” which will be discussed at length in chapter 5. Both categories of closed vowels are therefore desirable in a choral setting for their opposition to the seemingly

---

3 The term “brilliance,” in this case, refers to the acoustic perceptual quality of [i], which will be discussed further in chapter 5. Note that Jordan 1996 also warns against an improperly produced [i], which he claims is particularly susceptible to intonational difficulties.

4 Sundberg 1987 additionally states that closed lips tend to lower the larynx, which emulates the laryngeal motion associated with singing higher.
ubiquitous problem of flatting. Open vowels, bereft of advantageous upward tendencies derived from either natural placement or auditory perception, will be more likely to flatten.

This chapter has established, through choral practices and studies on intonation, that vowels conducive to proper resonance placement closely coincide with accurate intonation. Because, on average, singers tend to drift flat more than sharp, choral pedagogues advocate the use of vowels that counteract downward tendencies. Vowels used in choral warm-ups may, by extension, have similar qualities and levels of desirability in the context of solfège. That is, the same characteristics that cause a vowel to inherently flatten or sharpen may lend themselves equally to both solfège use and choral warm-ups, for similar reasons. After all, choral warm-ups and solmization share many similar goals: the development of careful listening and proper intonation, the ability to internally audiate, the development of proper vocal technique, and the capability to accurately recreate heard music (Karpinski 2000, 85-87, 145-46, 156, 169-71; Hylton 1995, 10-13, 21-26). The information presented in this chapter is used in the following chapter as it explores the effects of diphthongs, which combine two consecutive vowel sounds.
Chapter 4: Diphthongs and Vowel Unification

Diphthongs, such as the [oo] found in Do and Sol, can be separated into two distinct vowel sounds—in this case, [o] and [ʊ]. As described in chapters 3 and 5, each vowel taken individually has unique characteristics, which can be applied even in the context of diphthongs. The unbroken placement of vowels beside each other, however, requires additional attention.

When clarity of intonation is a primary goal, as is the case with solmization systems, diphthongs in general are problematic. In a choral setting, undirected singers may glide from the first vowel to the second at different times. Figures 5 and 6 show some common interpretations of sung diphthongs, featuring the ways in which a singer may choose to treat the shift between vowels. When addressing this issue, choral scholars agree that the loss of vowel uniformity that arises from singing diphthongs results in unclear diction and intonation (Jordan 1996, 281; Hammar 1984, 100). Hylton (1995, 25) even claims that “even though the sound may technically be in tune, it will sound as if it is not” when vowels are pronounced inconsistently."

---

5 The potential disparity in pronunciation between Do and Sol should be addressed. Specifically, due to its spelling, some may have a tendency to pronounce the “L” at the end of Sol. This may become particularly acute if Sol is followed by a syllable beginning with L, such as La. For the purposes of this document, the ramifications of that particular pronunciation of Sol will not be discussed, for two reasons. First, to the best of the author’s knowledge, pronouncing the L at the end of Sol is less common than omitting it. Second, L is a consonant, and as such lies outside the scope of this thesis. For the remainder of this paper, Do and Sol will be considered to have identical pronunciations past their initial consonants.
In speech, diphthongs often manifest themselves as a smooth “glide” between two vowel sounds. That is, the speaker doesn’t pronounce one distinct vowel sound before moving to another distinct vowel sound. Rather, the mouth shifts from the first vowel to the second without a break in vocalization. The result is a smooth gradient between those two vowels spanning the entire duration of the diphthong (Hammar 1984, 100; Haasemann 1991, 108). This glide can be demonstrated by saying the word “house” in a natural fashion. Although the diphthong for “house” is represented in IPA as [aʊ], the two individual vowels [a] and [ʊ] do not entirely represent what is actually said. As the mouth moves from [a] to [ʊ], every vowel formed using a shape between [a] and [ʊ] is briefly intoned. To move from [a] to [ʊ], the lips gradually become more closed. Because of this, the vowels lying between [a] and [ʊ] on the vowel trapezoid shown in figure 2 are
pronounced in the process. A more accurate, if impractically cumbersome, representation of the \[aʊ\] diphthong might be shown as \[aʌʊəʊ\], due to the vowels actually articulated during the process. Although the diphthong \[aʊ\] contains a more extreme shift between vowels than either \[eɪ\] or \[oʊ\], which appear in solfège syllables, it serves to illuminate the issue at hand.

Concerning solfège syllables sung in a group setting, the phenomenon of intonational issues arising from differing vowels due to diphthongs has obvious ramifications. Without consistent direction, multiple singers could easily approach a diphthong differently, despite singing it in the same context. Substituting Do in figure 6 or Re in figure 5, in place of the given words, shows the potential for solfège syllables to receive varied pronunciation. To address this issue, conventional singing practices for diphthongs employ a process known as vowel elongation (Hammar 1984, 100-101; Wall 2005, 30-35, 64-69, 108-27; Marshall 1953, 165-84; Hylton 1995, 25; Ware 1998, 166; Christy 1965, 70-72). With vowel elongation, the gradual glide between vowel sounds is minimized. Instead, there is a specific point at which the singer shifts from one vowel in the diphthong to the next, as represented in figures 5 and 6. The approach taken to the elongation of a particular vowel may vary, depending on the group and style of singing. The important point of vowel consistency is not to prescribe one correct universal method for singing diphthongs, but to unify vowels between singers within a group in order to avoid unclear diction and intonation. While diphthongs themselves do not cause intonational drift in one particular direction, articulatory inconsistency among singers obscures the true pitch center, resulting in aural destabilization. Providing clear,
homogeneous instruction in a choral or aural skills class setting preemptively mitigates problems relating to vowel treatment, and ultimately results in a clearer, more focused sound.

Diphthongs sung in a solo context provide some different challenges from those sung in a group. The destabilizing effect of diphthongs on aural perception and consequently intonation in solo singing is diminished by vowel elongation. However, for those inexperienced singers who display speech-like pronunciation in singing, diphthongs that gradually glide from one vowel to the next, instead of switching at a specific point, may be problematic. Most texts seem to focus on how to correctly approach diphthongs, rather than exploring the impact improperly sung diphthongs have on vocal production and intonation. Russell A. Hammar (1984, 100), however, says the following:

Failure to focus upon one of the two vowel sounds results in an indefinite, diffused sound. This writer refers to the mixing of these sounds in singing as “vowel migration,” i.e., the singers’ vowel focus leaves the core of the sound that should be produced. Diphthongs interfere with vocal production when they are mixed together.

Given the established connection between healthy vocal production, diction, and intonation, improperly treated diphthongs have the potential to cause intonational issues. Transitioning between two vowel sounds requires part of the vocal mechanism to change. Because the physiological processes that contribute to producing sound are so interdependent, modifying one area may undermine the stability of the entire system. The physical responses required to maintain a steady, accurate tone across vowel changes add an element of difficulty that is not present in singular vowel sounds. Interruption of comfortable, steady vocalization may be engendered intonationally, or as a distraction to
the singer. The potential problems applying to solo singers may also exist among individuals singing in a group, which reinforces the need for clear instructions concerning the treatment of diphthongs.

Returning to Do and Sol, the hazards of a diphthong within solfège syllables become apparent. In the movable-Do system, solfège syllables represent scale-degree functions rather than specific pitches. Do always represents tonic, regardless of the given key. Likewise, in the major mode, Re represents the supertonic, Mi represents the mediant, Fa represents the subdominant, Sol represents the dominant, La represents the submediant, and Ti represents the leading tone.

As tonic, Do plays a uniquely important role within the diatonic system. The ability to quickly infer or determine tonic is intimately linked with establishing one’s location within a diatonic collection (Karpinski 2000, 145-54). Additionally, maintaining a sense of tonic, even when tonic is not being sounded, is one of the crucial benefits to solmization systems relying on scale-degree functionality, such as movable-Do (Rogers 1996, 149-50; Karpinski 2000, 149). In this respect, tonic acts as the fundamental reference point used by musicians to find their location within a key. As such, its stability and accuracy is of paramount importance. To a lesser extent, the dominant, Sol, is used in a similar way. The Sol-Do relationship is highly indicative of the traditional V-I harmonic pattern that plays such a ubiquitous role in tonal music. Sol plays a supporting role to Do, and can act as a similar reference point to further strengthen a musician’s

---

6 In the interest of consistency, this document will use Do-based minor, rather than La-based minor. Karpinski 2000 gives an comparison of these two systems, among others.
loca
tional awareness within a diatonic collection (Karpinski 2000, 148-54). Because they share the diphthong [oo], the same difficulties affecting Do can be applied to Sol.

As users of the movable-Do solmization system become comfortable, the association between solfège syllables and their respective functions becomes stronger.

Gary Karpinski (2000, 85-86) writes:

Once listeners become fluent in a functional solmization system, the syllables become self-reinforcing and serve as immediate and facile means of communicating functional understanding. “Do” and “1” become more than just shorthand for “tonic”; they become intimately associated with tonic function, resulting in a personal knowledge of the tonic and how it feels.

Given the essential interchangeability between Do and tonic, and the significance of tonic’s function, it stands to reason that the destabilization of Do may have widespread consequences. Disruption of the reference point by which other pitches are measured jeopardizes the measurement itself. Loss of contextual orientation may manifest itself in degrees of severity, depending on the extent to which the sense of tonic has been disrupted. From a gradual drift in intonation to the immediate loss of correct pitch, any technical issue potentially arising from unclear tonality should be avoided.

From a musical standpoint, the potential for Do and Sol to be sung as diphthongs is problematic, due to physiological complications. The particular importance of Do and Sol within the tonal field serves to reinforce the extent of this issue. Certainly, singing Do and Sol without the diphthong is preferable in terms of stability and accuracy. This can be accomplished either by vocally educating those using solfège, or by the replacement of Do and Sol with syllables lacking the susceptibility to include a diphthong.

Pragmatically, the option of vocal education seems more immediately realistic than
replacing such well-established syllables. However, given the historically drastic successful changes to solmization systems, a potential improvement to the existing movable-Do system, however unlikely, should not be dismissed offhand.\textsuperscript{7}

Fortunately, even within academic settings such as an aural skills class, diphthongs are easily rectified and prevented. For instance, a quick introduction to the five Latin vowels, \([\varepsilon], [i], [a], [o], \text{ and } [u]\), provides students with quick, approachable examples of correct pronunciation. When needed, instructors can then remind students to sing only the “pure” Latin vowels, which serve as references that can be applied to solfège syllables.

This chapter has explored the negative effects that gliding uninterrupted from one vowel to another in a diphthong can have on physiological stability and aural clarity. The individual vowel characteristics described in chapter 3 still apply to diphthongs, and are simply added to the destabilizing diphthongal factor. The next chapter approaches vowels in the context of perception through acoustic analysis. It will help elucidate the differences between vowels described in chapters 3 and 4, and add the factor of perception to vowels’ combined intonational tendencies.

\textsuperscript{7} McNaught 1893 contains an interesting overview of solmization systems’ historical development.
Chapter 5: The Acoustic and Perceptual Qualities of Syllables Found in Solfège

Humans are able to distinguish between vowels with a remarkable degree of accuracy, as evidenced by the sheer number of discrete vowels included in dictionary literature. Primarily, the recognition of vowels is done subconsciously and automatically. At times, vowels are the only distinguishing factor between words with dissimilar meanings. For example, “beat,” “bait,” “bit,” “bat,” “bot,” “boat,” and “but” are differentiated from each other exclusively by their vowel content. All of these words have only minute articulatory distinctions between them, which are then received, interpreted, and comprehended by the listener almost instantly from a distance.

The biological and neurological operations that contribute to auditory discernment are complex and tangential to this work. Knowing how humans sense sound, however, is not necessary to study how those noises are perceived. Sounds, including those produced vocally, can be objectively analyzed outside the scope of human sensation through understanding how they are generated and propagated.

Sound is not comprised of some physical material that can be contained or visually observed. When objects interact, e.g. through collision or friction, the resulting physical disturbance causes vibrations to propagate outward from the point(s) of interaction, like waves in water. Sound is merely the name given to those vibrations, in

---

8 The table of contents in Wall 2005 comprehensively lists the recognized vowels in English, Italian, French, and German. Coffin 1980 deals almost exclusively with vowel modification, which demonstrates an even more refined degree of aural sensitivity.

9 In this chapter, perception does not refer to the biological function of sensing sound. Instead, it is associated with the interpretation or understanding of that sound.
relation to how they are perceived. Pitch is determined by the frequency of vibrations per second, known as hertz (Hz). High pitches vibrate quickly, and thus have a greater frequency than lower pitches.

For a combination of reasons, most noises generate more than one frequency. In the case of vocal production, this is due to the resonating chambers included in the vocal tract (Ware 1998, 135-42). Similar to the way brass instruments have multiple pitches that naturally resonate in any given position or fingering, certain frequencies “fit” in the vocal tract’s chambers. In addition to the fundamental frequency [the lowest, and usually most intense, frequency generated in a sound] created through the vibration of the vocal folds, multiple higher pitches are also produced due to their resonance within the vocal tract’s various cavities. These naturally resonating frequencies are known as formants. The fundamental frequency is labeled as F₀, while following formants are labeled as F₁, F₂, etc. in order of ascending frequency.

Any given vowel can be sung in multiple registers, and multiple vowels can be recognized on a single pitch. Therefore, fundamental frequency is not a primary distinguishing factor between vowels. When singing different vowels on a single pitch, the vocal folds maintain vibration at a single frequency. However, modifying the vocal tract also changes the size and shape of its resonating chambers. The changes in lip and tongue position (among other physiological possibilities) required for pronunciation directly impact which frequencies naturally resonate in the vocal tract (Ware 1998, 138-10

---

10 For a more thorough definition of terminology and an accessible introduction to the physics of sound, including how it relates to vocal production, see chapter 8 in Ware 1998.
11 Thinking of pitches as “higher” or “lower” is a cultural conception that does not accurately describe the physical properties of sound. They will serve for the purposes of this document, though, due to their virtually universal use among Western musicians.
40). In other words, formants above the fundamental frequency (specifically, the relationship between $F_1$ and $F_2$) are the principal determining factor in vowel recognition. $F_3$ and $F_4$ are also perceived in speech and singing, but are associated with timbre rather than pronunciation (Ware 1998, 158).

Frequencies generated by a given sound can be graphed according to their intensity. In this way, it is possible to visually represent which frequencies are present during the vocalization of individual vowels. Additional information can then be extracted by comparing the frequencies among multiple vowel sounds. Figure 7 shows, in hertz, the frequencies of $F_1$ and $F_2$ for multiple vowels. Due to variances in physiology and pronunciation, each vowel is represented in an area, rather than as a single point. As shown in figure 7, neither $F_1$ nor $F_2$ alone determine vowel recognition. For example, [i] and [u] are very similar in $F_1$ frequency, but are very different in $F_2$ frequency. Likewise, [i] and [e] have an almost identical range of $F_2$ frequencies, but vary along the $F_1$ axis. It is the combination $F_1$ and $F_2$ that determines vowel, in both independent frequency and relationship to each other.

---

12 One method for determining the natural resonant pitches of vowels without the factor of fundamental frequency involves the acoustic analysis of vowels breathed without phonation. Paget 1976 describes this process.

13 Ware 1998 provides easily interpreted examples on pp. 133 and 162.
Figure 7: Frequency of $F_1$ and $F_2$ by vowel (from Ware 1998, 159)

Notice the similarity of vowels’ relative positions between figure 7 and the vowel trapezoid shown in figure 2. If the vowel trapezoid was horizontally flipped then rotated 90 degrees counterclockwise, it could be superimposed on figure 7 fairly accurately. Although this connection reinforces the link between physiological generation and acoustic properties, there still remains the matter of aural perception. As Philip Lieberman (1988, 153) writes:

Although it is possible to perform precise analyses of speech signals using electronic instruments and computer programs that effect various mathematical transformations of the signal, these analyses are, in themselves, meaningless. We can never be certain that we have actually isolated the acoustic cues that people use to transmit information to each other unless we run psychoacoustic studies in which human listeners respond to acoustic signals that differ with respect to the acoustic cues that we think are relevant.
Although representing the frequency of vibrations per second in hertz is a useful scale of measurement, it does not proportionally represent the way humans interpret frequencies. Meaningful graphical representation of auditory perception relating to pitch relies on the “Mel” conversion scale, as shown in figure 8. The Mel scale was devised to proportionally represent auditory perception (Lieberman 1988, 154; Pederson 1965, 296). Thus, a pitch of 2000 Mel is perceived as being twice as high as a pitch of 1000 Mel, which in turn is perceived twice as highly as a pitch with 500 Mel.\textsuperscript{14} Using this conversion, graphs can be scaled to appropriately represent human perception, rather than pure mathematics. Figure 7, for example, is useful for easy visual interpretation of the relationships between vowel formants, but is not designed to proportionally show the perceived extent of those relationships.

\textsuperscript{14} Trained musicians may be tempted to say a pitch that has a 2:1 frequency ratio (an octave) to another pitch is twice as high. However, Mel's were developed outside of a musical context, and do not conform to musical preconceptions. Thus, in the context of subconscious pitch discrimination, such as $F_1$ and $F_2$ when perceiving vowels, the Mel scale is still useful in representing the differences in aural perception of vowels. Pederson 1965 gives more information on the process by which the Mel scale was determined, as well as an examination of the system’s merits and deficiencies.
Figure 8. The relationship between Mels (Y-axis) and kilohertz (X-axis) (from Lieberman and Blumstein 1988, 155)

Figure 9 shows the frequencies of F₁ and F₂ in Mels, rather than Hz. The vowels here are not shown as an area like figure 7, but instead as averaged points. Figures 9 and 7 look approximately the same, due to the general shape and direction of F₁ and F₂ frequencies plotted on their respective scales. However, the relative positions of the vowels between figures 7 and 9 are quite different, due to the differing information they are meant to convey. Figure 7 displays the vowels in hertz based on sung physiological modification, where figure 9 is intended to represent aural perception of speech. At times, though, information is inconsistent between them. For example, according to figure 7, [i] and [ɪ] lie in approximately the same F₂ range. However, figure 9 shows [i] lying significantly higher than [ɪ] along the F₂ axis. Regardless of which figure is more accurate, both sources agree on the affected perceptual result. The creator of figure 9,
Philip Lieberman, describes the vowels moving along the line from [a] to [i] as becoming more “acute,” while those moving from [a] to [o] become increasingly “grave.” Clifton Ware, the creator of figure 7, uses different terminology, and describes [i] as being “brilliant,” while [o] and [u] are considered “dark” (Ware 1998, 161).

From a strictly mathematical standpoint, the fundamental frequency’s pitch is not affected by $F_1$ or $F_2$. However, whether its perceived pitch is affected by the frequencies of $F_1$ and $F_2$ is another matter. The idea seems plausible, given that the combined frequencies of $F_1$, and $F_2$ vary among vowels sharing the same $F_0$. Ware claims that

![Figure 9. American English vowels based on perception (from Lieberman and Blumstein 1988, 182)](image-url)
vowels do contain differences not only in timbre, but also in perceived pitch (Ware 1998, 160-61). His conclusions are summarized in a table, reproduced here as figure 10.

According to Ware, the vowels described as brilliant or acute are perceptually higher, while those described as dark or grave are perceptually lower. Most vocal pedagogues avoid describing vowels as higher or lower in an effort to avoid inducing intonational tendencies and improper production, but often refer to vowels using terms such as brilliance, brightness, and darkness (Christy 1965, 63-64; Schmidt 1989, 10-11; Powell 1991, 42; Jordan 1996, 283; Hammar 1984, 86-87; Wall 2005, 47-48; Marshall 1953, 125; Hylton 1995, 22). These descriptive words certainly have their own connotations to height. The words brilliant and bright are reminiscent of sunlight and the sky, whereas darkness invokes depth and shade. Haasemann and Jordan (1991, 70) make an immediate association between brightness and height: “Rabbit teeth can be used to brighten an overly dark sound and to bring brightness into descending in lower-register singing. It also helps to create a higher pitch.” Even though these examples are in reference to vocal production, the imagery invoked arose from the perception of changes in pronunciation.

<table>
<thead>
<tr>
<th>Vowel Characteristics</th>
<th>[i]</th>
<th>[e]</th>
<th>[a]</th>
<th>[o]</th>
<th>[u]</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pitch perception</td>
<td>highest (forward)</td>
<td></td>
<td></td>
<td>lowest (back)</td>
<td></td>
</tr>
<tr>
<td>Timbre</td>
<td>brilliant</td>
<td>medium</td>
<td></td>
<td>dark</td>
<td></td>
</tr>
<tr>
<td>Tongue (and larynx)</td>
<td>raised</td>
<td>slightly raised</td>
<td></td>
<td>lowered</td>
<td></td>
</tr>
<tr>
<td>position</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Figure 10. Summary of vowel characteristics (from Ware 1998, 161)

---

15 Moriarty (1975, 9-17) does directly describe some vowels as higher or lower. However, this is an attempt to induce proper physiological production through imagery. All other observed authors are careful to associate directional terms with tone placement, as described in chapter 2, rather than the vowels themselves.
Assuming, then, that intonational perception is linked with vowel color and timbre, acoustic feature responsible for that connection can be determined. Notice each vowel’s position on the F2 axis in both figures 7 and 9 in reference to their position in figure 10. Vowels with the highest perception also have the highest F2 frequency, with a common descent from [i] to [a] between both perception and F2 frequency.16 F1 frequency is not a determining factor from [i] to [a], as common F1 values are shared among vowels with both high and low perceptual qualities. From [a] to [u] or [o], figures 7 and 9 disagree more strongly on the exact frequencies among vowels. However, both graphs show a lessening of F2’s impact along this line. Therefore, along the journey from [a] to [u] or [o], F1’s role becomes more critical in determining vowel color and perceived pitch.

Sundberg (1987, 137) describes an experiment that connects vowel perception to intonation. In this experiment, a choral bass section recorded a steady reference note against which other singers were asked to sing various intervals. The reference note was then varied by pitch and vowel, and the relative accuracy of intervals above it was measured. Although this experiment was structured to determine whether there was a link between lining up formant frequencies with a partial [a sounding member of a pitch’s overtone series] of the reference pitch and relative accuracy, it did discover predictable differences in intonational accuracy based on the reference pitch’s vowel.

A more applicable situation can be described in the exercise of tuning perfect 5ths between two choral sections while employing different solfège syllables. For instance,

16 In figure 9, no [a] is present. The author’s inference is that no distinction between [a] and [a] is made in this example, since it is solely representing American speech.
Do-Sol and La-Mi are both a perfect 5th apart. If these syllables were sung on the same two pitches, e.g. F-C, the perceived tuning would slightly change with vowel variation even if the fundamental frequencies remained constant. In other words, Do and Sol share [oo], and therefore have the same perceptual qualities, assuming uniform pronunciation. In contrast, La and Mi have quite different perceptual characteristics. The upper note, on [i], is perceived as a very high vowel due to its formant frequencies, whereas the lower note, on [ɑ], is perceived as slightly lower than average. So, even if the fundamental frequencies remain static, the perfect 5th between La and Mi might sound too large due to the opposite pull of their respective formants.

By combining the information in this chapter with the physiological tendencies of vowels described in chapters 3, 4 and 5, the interplay between vowel articulation and vowel perception can be examined. In some cases, such as [ɑ], the tendencies align. That is, [ɑ] trends downward in both articulation and perception. However, vowels such as [u], display interesting tension. Physiologically, [u] is extremely closed, and promotes healthy forward resonance and accurate intonation. Perceptually, however, [u] is the darkest and lowest vowel, due to the makeup of its formants. Measuring the relative strength of these conflicting characteristics is likely variable and beyond the scope of this document. However, the next chapter clearly lists the physiological and perceptual qualities of each vowel, so that reinforcing or opposing aspects can be compared. Additionally, the next chapter surveys functional expectations commonly attributed to the scale degrees each solfège syllable represents. In compiling each vowel’s attributes into one chapter, some of the forces affecting intonation can be more easily interpreted.
Chapter 6: The Vowels Used in Solfège and Their Individual Pitch Tendencies

This chapter surveys functional expectations attached to the scale degrees each solfège syllable represents. Additionally, each syllable is individually discussed in terms of its physiological tendencies and perceptual qualities, along with any additional relevant considerations. The cumulative information is compiled at the end of this chapter, in figure 11. For convenience, the solfège syllables used in movable-Do and their associated IPA symbols are provided below:

- [oo]: Do and Sol
- [α]: Ra, Fa, and La
- [i]: Di, Ri, Mi, Fi, Si, Li, and Ti
- [et]: Re, Me, Se, Le, and Te

As the tonic, Do is the most functionally stable of all scale degrees. Because the tonic chord is used as the starting point for most traditional harmonic patterns, Do can depart in any direction. With so many options available, there are no expectations concerning Do’s initial movement. However, as the point to which other notes often lead, Do acts as the reference pitch that causes other scale degrees to have tendencies. The first discrete vowel sound pronounced when singing Do is [o], which is then often followed by a closing of the lips toward [ʊ], forming the diphthong [oo]. As a moderately closed

---

17 The solfège syllables’ functional tendencies will be necessarily limited, due to the multiple contexts in which each scale degree can appear. The prominent tendencies are based on the author’s discrimination, with the consultation of traditional harmonic voice leading implications outlined by Kostka 2009 and Rogers 1996. The overall functional tendencies in this chapter are not intended to be comprehensive or universally applicable. Functions, and therefore functional tendencies, are subject to variability depending on context. The point of including scale degree tendencies is to present likely functional scenarios that result in the typical expected motion of a solfège syllable.

18 Because the relative strength of each factor’s tendencies are likely variable depending on context, scale degree tendencies, physiological tendencies, and perceptual qualities will all be considered as roughly equal.
lip vowel, [o] has natural upward tendencies due to resonance placement and laryngeal positioning, which are additionally strengthened during the glide to [ʊ]. In contrast, [o] is perceived as a low vowel due to the F$_1$ and F$_2$ frequencies that resonate during its production. The move to [ʊ] reinforces that low perception. The overall intonational impact of Do is therefore fairly balanced since it tends physiologically upward, perceptually downward, and functionally stable. However, the destabilization caused by the [oo] diphthong may create aural ambiguity or vocal inconsistency.

Sol, as dominant, is the second most mentally prominent scale degree behind Do. Although Sol is included in the tonic chord, it more notably acts as the root of V. Sol is heavily influenced by the role of V, and mirrors the root movement often associated with V. In a V-I progression, for instance, Sol can either jump up or down to Do. If Sol instead followed a common harmonic motion to Fa, La, or Le (e.g. V to V$^7$, IV, or vi), depending on the mode, the stepwise motion of Sol either up or down would also be expected. Sol acts as a secondary reference point, behind Do, and is permitted similar freedom of movement. The net expected motion, therefore, is largely nullified, due to the general balance in allowed movement. The physiological and perceptual characteristics of the [oo] in Sol are identical to Do, and will not be repeated in depth. To summarize: Sol trends moderately upward physiologically, moderately downward perceptually, and is potentially destabilized in either direction by its diphthong.

As the leading tone, Ti plays an integral role in the V-I relationship. Because of its close proximity to the tonic, Ti conventionally resolves upward by semitone to Do. Motion from Ti to Do is aurally reinforced due to the regular occurrence of V-I, which
often signifies the end of a phrase, and is particularly noticeable due to Do’s role as tonic. Although Ti can move downward away from Do, the associated harmonic patterns are not as functionally essential or conspicuous as V-I. Therefore, Ti is almost always thought of in relation to Do, and consequently has a strong upward mental association. Physiologically, Ti is formed using the most closed tongue position, which fosters good resonance placement and inhibits flatting. Due to the high frequency of F2, as well as the large distance between F1 and F2, Ti is perceived as very high and bright. In summary, Ti trends significantly upward functionally, significantly upward physiologically, and significantly upward perceptually.

The subdominant, Fa, is often found in two contexts. In the first, Fa fills the predominant role while acting as a member of ii or IV. These predominant chords naturally progress to V, the dominant, during which Fa is often pulled upward to Sol according to voice-leading convention. The second, and stronger, main functional context including Fa is the pervasive dominant-to-tonic motion of V7 or vii° to I. In both of these progressions, Fa is strongly pulled downward by resolving to Mi. The Fa-Mi motion reinforces and is mutually reinforced by Ti-Do, strengthening the tendencies involved in both progressions. Although the Fa-Sol and Fa-Mi tendencies counteract each other, the downward motion associated with Fa is stronger due to the harmonic context. So, while the upward Fa-Sol motion may mentally mitigate that affiliation, the overall expectation is for Fa to move downward. Physiologically, [ɑ] is generated using fully open tongue and lip position. The resulting natural relaxation and resonance placement give [ɑ] significant downward tendencies. The formant frequencies produced by [ɑ] cause it to be
perceived as slightly dark and low, which reinforces the downward physiological inclination. To summarize, Fa trends moderately downward functionally, significantly downward physiologically, and slightly downward perceptually.

La, the submediant, shares the same physiological and perceptual qualities as Fa, but differs somewhat in function. Although La is part of the same distinctive harmonic progressions that define Fa, Sol, and Ti, it is generally placed in a less noticeable position. For example, La will often lead up to Ti, which in turn resolves to Do, e.g. in IV-V-I. The predominant role of La, in this case, assists the dominant in its expected motion to tonic. La, though important, is certainly not the point of interest in this progression. Alternatively, La may resolve downward to Sol, e.g. in vi-V or IV-I. Whenever La steps downward to Sol, there is an equal expectation that it may instead move up to Ti. La may also resolve downward to Sol in a plagal cadence, or IV-I, but again is not the focus of attention during this progression. Because La has the ability to move upward or downward in roughly equal measure, there is no consistent directional expectation attached to it.

The supertonic, Re, lies midway between two members of the tonic triad. The most significant harmonic motion associated with Re is V-I, in which Re resolves either to Do or Mi/Me. Although both are valid, the powerful, stable attraction of Do may cause Re to have slightly stronger downward expectations than upward. Physiologically, Re is formed using a moderately closed tongue position, which closes further during the glide from [e] to [ɪ]. Consequently, Re is fairly conducive to forward placement, which helps prevent intonational drooping. Perceptually, [e] has moderately high and bright qualities,
which are reinforced as [e] naturally closes toward [i]. Because [eɪ] is a diphthong, there may be destabilization or lack of clarity in its vocalization or aural reception.

Mi, the mediant, is a member of the tonic triad, which often relegates it to the role of supporting Do. Mi tends to move either upward to Fa, or downward to Re; in either case, the harmonic associations are generally in motion away from tonic, e.g. I to V(7) or IV. Due to its position in the tonic chord, Mi is a functionally stable scale degree toward which other notes, such as Re and Fa, progress.\(^{19}\) Thus, the overall expectations associated with motion away from Mi are balanced and neutral. Physiologically and perceptually, Mi shares the same characteristics as Ti, which both result in significant upward tendencies.

The syllables Ri, Fi, Si, Li, and Di share similar functions. They are often associated with their diatonic counterparts, as reflected in their syllabic organization. For example, the thought process when finding Fi may be, “like Fa, but higher by a semitone.” However, the syllabic associations of Ri, Fi, Si, Li, and Di are less powerful than the harmonic contexts in which they are found. Typically, these four solfège syllables will be found as part of a secondary chord.\(^{20}\) Secondary dominants and diminished sevenths serve to augment the expected harmonic motion toward another chord, which is sometimes referred to as a secondary tonic. In a continuation of that terminology, Ri, Fi, Si, and Di can, in this context, be thought of as secondary leading

\(^{19}\) Although Mi may serve to strongly tonicize IV due to its placement a semitone below Fa, which would give Mi strong upward expectations, that event is highly situational.

\(^{20}\) Li is the exception. Li is generally found as part of some chromatic motion, rather than as part of a secondary chord. Although Li shares the same scale degree as Te, Li is usually approached in an upward ascent from below, whereas Te tends to be approached downward from Do. So, although Li does have associated upward motion, it is not as strong as the expectations which arise from a secondary dominant or secondary vii\(^7\) progression.
tones. For instance, V/vi strongly enhances the anticipation of vi, largely due to the upward semitone motion from Si to La which emulates the aforementioned Ti-Do relationship. Because of the similarity in function between the leading tone and these secondary leading tones, Ri, Fi, Si, and Di have strong upward mental expectations. Physiologically and perceptually, due to their shared vowel, all of these raised syllables, including Li, share the same strong upward qualities ascribed to Ti.

The subtonic, Te, is specifically associated with downward motion, as evidenced by the melodic minor scale. In this scale, Ti is used when ascending due to its strong association with Do, whereas Te is conversely used in descent away from Do. Therefore, there is a strong downward expectation associated with Te. Due to proximity, Te may retain some upward association with Do. However, the traditional downward context in which Te is generally found outweighs that upward connection. Physiologically and perceptually, Te shares the same upward qualities as Re, and contains the diphthongal destabilization caused by [eɪ].

Le and Me share similar functions as their counterparts, La and Mi. However, Me and Le are often resolved downward by semitone to Re and Sol, respectively, which are both members of V. Similar to the strong semitone tendencies described in the secondary dominant and Ti-Do relationships, Me and Le lead downward to Re and Sol. Although both Me and Le can move upward, the association is outweighed by the downward context.

---

21 The melodic minor scale was retroactively determined based on compositional patterns. That is, using Ti while ascending and Te while descending was common enough to merit their relative positions in the melodic minor scale. Thus, the functional expectation followed compositional practice.

22 In fact, the upper leading tone found in the Phrygian mode is how half cadences using a Le-Sol motion received the name of Phrygian Half Cadence.
Physiologically and perceptually, Me and Le share upward characteristics with Te and Re, and are potentially destabilized by the diphthong of [et].

The final solfège syllable, Ra, shares the downward physiological and perceptual associated with Fa. However, because of its close placement a semitone above Do, Ra functions as an upper leading tone. In a reversal of the upward pull found in Ti-Do, Ra likewise generally moves to Do, giving it equally strong downward expectations. Harmonically, Ra is often found in the context of an N\(^6\)-V-I progression. In this progression, Ra moves downward to Ti, which then resolves upward to Do. The unique sound of this progression utilizes both the upper and lower leading tones, which surround and reinforce the tonic. Within this common context, Ra “passes through” Ti on its way to Do, and the downward expectation is preserved. The strong functional pull associated with Ra, when combined with the dark and low physiological and perceptual qualities, give it maximal downward tendencies. Because of this, Ra may be particularly susceptible to flattening.

Figure 11 shows a summary of the information outlined in this chapter, and will be used as a reference for the analysis in chapter 7.

---

\(^{23}\) Le could move upward to either Te or Ti, depending on context. Regardless, the functional pull of Sol outweighs that motion.
<table>
<thead>
<tr>
<th>Solfège Syllables</th>
<th>IPA</th>
<th>Functional Tendencies</th>
<th>Physiological Tendencies</th>
<th>Perceptual Qualities</th>
<th>Additional Consideration</th>
<th>Collective Impact</th>
</tr>
</thead>
<tbody>
<tr>
<td>Do</td>
<td>[oʊ]</td>
<td>Neutral</td>
<td>Moderately upward</td>
<td>Moderately low</td>
<td>Diphthongal destabilization</td>
<td>Neutral, with destabilization</td>
</tr>
<tr>
<td>Sol</td>
<td>[oʊ]</td>
<td>Neutral</td>
<td>Moderately upward</td>
<td>Moderately low</td>
<td>Diphthongal destabilization</td>
<td>Neutral, with destabilization</td>
</tr>
<tr>
<td>Ti</td>
<td>[i]</td>
<td>Significantly upward</td>
<td>Significantly upward</td>
<td>Significantly high</td>
<td>–</td>
<td>Maximally upward</td>
</tr>
<tr>
<td>Fa</td>
<td>[a]</td>
<td>Moderately downward</td>
<td>Significantly downward</td>
<td>Slightly low</td>
<td>–</td>
<td>Significantly downward</td>
</tr>
<tr>
<td>La</td>
<td>[a]</td>
<td>Neutral</td>
<td>Significantly downward</td>
<td>Slightly low</td>
<td>–</td>
<td>Moderately downward</td>
</tr>
<tr>
<td>Re</td>
<td>[ei]</td>
<td>Slightly downward</td>
<td>Moderately upward</td>
<td>Moderately high</td>
<td>Diphthongal destabilization</td>
<td>Moderately upward, with destabilization</td>
</tr>
<tr>
<td>Mi</td>
<td>[i]</td>
<td>Neutral</td>
<td>Significantly upward</td>
<td>Significantly high</td>
<td>–</td>
<td>Significantly upward</td>
</tr>
<tr>
<td>Ri</td>
<td>[i]</td>
<td>Significantly upward</td>
<td>Significantly upward</td>
<td>Significantly high</td>
<td>–</td>
<td>Maximally upward</td>
</tr>
<tr>
<td>Fi</td>
<td>[i]</td>
<td>Significantly upward</td>
<td>Significantly upward</td>
<td>Significantly high</td>
<td>–</td>
<td>Maximally upward</td>
</tr>
<tr>
<td>Si</td>
<td>[i]</td>
<td>Significantly upward</td>
<td>Significantly upward</td>
<td>Significantly high</td>
<td>–</td>
<td>Maximally upward</td>
</tr>
<tr>
<td>Di</td>
<td>[i]</td>
<td>Significantly upward</td>
<td>Significantly upward</td>
<td>Significantly high</td>
<td>–</td>
<td>Maximally upward</td>
</tr>
<tr>
<td>Li</td>
<td>[i]</td>
<td>Moderately upward</td>
<td>Significantly upward</td>
<td>Significantly high</td>
<td>–</td>
<td>Significantly upward</td>
</tr>
<tr>
<td>Te</td>
<td>[ei]</td>
<td>Moderately downward</td>
<td>Moderately upward</td>
<td>Moderately high</td>
<td>Diphthongal destabilization</td>
<td>Slightly upward, with destabilization</td>
</tr>
<tr>
<td>Me</td>
<td>[ei]</td>
<td>Moderately downward</td>
<td>Moderately upward</td>
<td>Moderately high</td>
<td>Diphthongal destabilization</td>
<td>Slightly upward, with destabilization</td>
</tr>
<tr>
<td>Le</td>
<td>[ei]</td>
<td>Moderately downward</td>
<td>Moderately upward</td>
<td>Moderately high</td>
<td>Diphthongal destabilization</td>
<td>Slightly upward, with destabilization</td>
</tr>
<tr>
<td>Ra</td>
<td>[a]</td>
<td>Significantly downward</td>
<td>Significantly downward</td>
<td>Slightly low</td>
<td>–</td>
<td>Maximally downward</td>
</tr>
</tbody>
</table>

Figure 11. Summary of vowel impact on pitch
Chapter 7: Application in a Choral Setting

This chapter contains a chorale analysis to illuminate points at which solfège syllables used in the learning process would either exacerbate or counteract potential pitch finding or tuning issues. This is not a note-by-note examination, but rather a highlighting of prominent examples that may then be applied elsewhere. Each voice part will be discussed individually to investigate the horizontal effect of tuning issues over time, with occasional illustration of that line’s position within the vertical tuning among other voice parts. Additionally, it will be assumed that the singers will practice solfège during the learning process. This hypothetical approach encompasses both unison and polyphonic settings in which solfège is used, which closely represent the melodic and harmonic situations encountered in aural skills classrooms.

As shown in figure 12, the soprano line in this chorale is largely stepwise and diatonic. Most leaps occur between Do and Sol, which are given emphasis via repetition and their positions on almost every fermata. Due to the mobile nature of this line, it is essential for the sopranos to maintain an accurate sense of position within the tonal field. As mental reference points, the recurrence of Do and Sol help establish a strong connection with tonic. Their very nature as tonic and dominant helps to maintain accuracy when leaping between each other, due to their closely linked functions. However, inconsistency between singers concerning the treatment of the [oo] diphthong found in Do and Sol could cause a gradual destabilizing effect through lack of uniformity, perceptual clarity, and sense of tonal location. The resulting intonational drift,
8. Choral

Sopranino
Das Wort sie sol- len las- sen stahn und kein Dank da- zu
The word of God no foe can harm, not even know its

Alto
Das Wort sie sol- len las- sen stahn und kein Dank da- zu
The word of God no foe can harm, not even know its

Tenor
Das Wort sie sol- len las- sen stahn und kein Dank da- zu
The word of God no foe can harm, not even know its

Bass
Das Wort sie sol- len las- sen stahn und kein Dank da- zu
The word of God no foe can harm, not even know its

Pianoforte

4
ha- ben. Er ist bei uns wohl auf dem Plan mit sei- nem Geist und
mer- it. God guides us with his might- y arm with weap- ons of the

ha- ben. Er ist bei uns wohl auf dem Plan mit sei- nem Geist und
mer- it. God guides us with his might- y arm with weap- ons of the

ha- ben. Er ist bei uns wohl auf dem Plan mit sei- nem Geist und
mer- it. God guides us with his might- y arm with weap- ons of the

ha- ben. Er ist bei uns wohl auf dem Plan mit sei- nem Geist und
mer- it. God guides us with his might- y arm with weap- ons of the
Figure 12: Chorale from *Ein feste Burg ist unser Gott*. (from Bach [1715] 1985, 71-72.)
if reinforced through multiple readings, may prove detrimental to the independent solidification of this line among sopranos.

The alto line begins by strongly establishing tonic in a downward stepwise run from Sol to Do. Immediately in m. 2, though, there is an ascending octave jump followed by Fi, a chromatically altered note. The altos had just sung G on Fa, which has significant downward tendencies, and may potentially be slightly disoriented due to the octave jump at the beginning of m. 2. Consequently, finding G# may prove problematic for some of the singers. With Fa fresh in mind, some singers may sing G instead of G#, while others may simply guess on their way to the following note, Re. Although the G# in m. 2 only lasts an eighth note, it distinguishes the chord being sung as E major, rather than E minor.\(^2^4\) In a general reference to tuning, Hylton says “…ascending intervals should be conceptualized as somewhat larger than they actually are, and descending intervals need to be thought of as smaller” (1995, 71). When the tendency may be to sing under pitch, as in the case of the G# in question, the short step down from Sol to Fi is an excellent time to follow that suggestion. Fortunately, because G# is represented as Fi, it receives the upward physiological, functional, and perceptual tendencies ascribed to [i], which help maintain the intonational height required in this situation.\(^2^5\)

In measure 14, the altos sing D# using the syllable Di, which alters the very stable tonic note, Do. Do is placed throughout the alto line, and by this point tonic has been

---

\(^{24}\) This is an example, as mentioned in chapter 6, where Fi is part of a secondary dominant chord. Although the alto line does not resolve upward in this case, the G# is heard as moving upward to the soprano line’s A on the following beat.

\(^{25}\) The necessity of maintaining high intonation on this Fi is reinforced, due to its position as the 3rd of the chord, by the choral tuning study described by Sundberg (1987, 136). This is discussed in greater detail in the following paragraph concerning Di.
strongly established. Suddenly encountering a D#, therefore, may be initially confusing and would have a tendency to droop downward toward tonic, particularly when approached from above. The altos alone may not be confused by a low D#, but in harmonic context the effect of poor intonation would be greater. When unaccompanied, experienced choirs generally tune major intervals wider than those found in equal temperament (Sundberg 1987, 136). D# appears as the 3rd of a B major chord, which is held on a fermata. A flatted D# would be particularly noticeable due to its duration and deviation from standard choral tuning. Because this D# is sung on Di, however, the strong upward qualities of [i] will help the altos’ tendency to flatten. Additionally, it is also possible that because [i] is perceptually so dissimilar from the [oo] in Do, the vowel contrast may help singers more easily disassociate Di from tonic.

The tenor line in this chorale is quite disjunct; it changes directions often, covers a wide range, and jumps unexpectedly. Unlike the sopranos, the tenors rarely jump to Do or Sol, and thus are unable to use them as mental reference points. So, in general, the tenors may have more difficulty with intonation merely due to the construction of their line. The tenors begin m. 2 with a leap downward from Mi to La, which is the lowest note they have yet been expected to sing. The sudden extension of register combined with a large downward leap fosters tonal uncertainty, and becomes susceptible to intonational issues. Due to the [a] vowel’s strong downward physiological tendency, La may pull the note downward further, further exacerbating the potential for intonational uncertainty. In quick succession to La, the tenors then jump up to Re before returning downward to Ti. As the third of the A major chord on beat 3 of m. 2, Ti requires high, accurate pitch
according to choral tuning convention (Sundberg 1987, 136). Although Ti itself is conducive to maintaining upward intonation, the disjointed approach and varied vowel tendencies between [α] and the diphthong [ei] serve to destabilize the singers’ sense of tonal awareness and precision.

In m. 12, the tenors alternate between Sol and Fi, before singing Re on the downbeat of m. 13. Immediately after Re, though, the tenors jump up to Fa. After establishing Fi so strongly, finding Fa again can be difficult. However, the respective tendencies of Fi and Fa are conducive to accuracy in this excerpt. Because [α] is so much more open and perceptually lower than [i], the G (Fa) in m. 13 is much more easily differentiated that the preceding G#s (Fi). Articulating [α] pulls the singer downward, allowing them to more naturally find the correct pitch.

Following the tension between Fi and Fa in m. 13, the tenors immediately run upward diatonically to Do. On the way back down from Do, the lowered seventh scale degree, Te, is used in place of Ti. On the ascent, Ti leads strongly to Do due to its functional association, physiological generation, and perceptual height. However, the Te found during the descent from Do also has upward tendencies, albeit not as powerful as those found in Ti. Once the descent from D to C is established among the singers, maintaining an upward sense would likely help with intonation by keeping the distance between the two notes “small,” as prescribed by Hylton (1995, 71).26 Although this Te

---

26 However, because the [ei] in Te is so physiologically and perceptually similar to the [i] in Ti, the singers may have a difficult time distinguishing between the C# and C in the first place. In other words, if Te were sung instead as Ta, using [α], the two scale degrees would sound much more distinct from each other. Granted, in that scenario, [α] might cause the downward run to flatten. Nevertheless, the distinction between initially finding a pitch and singing that pitch in tune is interesting. Pitch finding and intonation, from an acoustic perspective, are gradients on the same scale. Their mental processes are different, however, and merit further consideration.
appears only briefly, lack of clarity in situations such as this can add up to more severe intonational issues.

The bass line is quite similar to the soprano line in that it is largely stepwise, diatonic, and jumps mostly to either Do or Sol. Although this chorale is not particularly difficult for the basses, maintaining a clear sense of location within the tonal framework is of particular importance since they provide the foundation upon which higher voice parts tune. As a result, intonational problems stemming from the bass line can have repercussive effects throughout other voice parts.

Although the above analysis ostensibly refers to a choir rehearsal, the same approach may be used to parse the specific difficulties of a sight-singing melody in an aural-skills class. Tonal constructs such as those described in this chapter appear in all types of music from the common practice period. Vowels are just one of many factors impacting the performance of a musician, but can be easily improved with only a cursory introduction to the material written here. Even though a deep understanding of vowel qualities requires a combination of interdisciplinary knowledge and personal singing experience, the concepts arising from that knowledge are easy to understand and communicate. For example, the following three guidelines could be easily incorporated into a music classroom, and are appropriate for all ages:

1) Sing brightly. Swallowing your tone leads to out of tune singing, so smile!

2) Avoid diphthongs. Sing only on pure vowels.

3) Listen carefully. Remember to reserve some of your attention for what you hear yourself singing.
These three instructions do not address important issues such as posture and breathing, but they are directly relevant to problems arising from vowel tendencies. Following these guidelines preemptively mitigates flatting caused by poor vowel placement and perceptual qualities, and helps to unify singers in a group setting. Those who use movable-Do are dependent upon the solfège syllables used within that system.

Implementation of guidelines such as those above into situations where solfège is taught would improve the intonational accuracy of those who use it, ultimately leading to more vocally confident and well-rounded musicians.
Chapter 8: Avenues for Further Research

Foremost among the ways in which this thesis could be expanded is the opportunity for controlled study. Given that this document is based entirely on connecting research across different musical fields of study, applied testing would illuminate the validity and severity of the speculative conclusions offered above. In direct relevance to this thesis, musicians could be asked to sing a given excerpt on a single, random, predetermined vowel. What would be the aggregate difference in accuracy between a test group using [i], for instance, as opposed to [o]? Other variables that could be isolated include testing the singing accuracy of vocalists compared to instrumentalists of similar educational levels, and testing intonational clarity (not pitch-finding accuracy) based on preparation time.

Another appropriate extension of this study would examine the effects of consonants within a solmization system. Is there a noticeable difference in accuracy of pitch between syllables beginning with voiced (such as “L”) or unvoiced (such as “S”) consonants? What about fricatives (such as “F”) compared to plosives (such as “T”)?

How do the concepts introduced in this thesis apply to different systems of solmization? How do they apply to different dialects? Are the “speech-like” tendencies of amateur singers present in other regions? If so, do those tendencies produce similar results? These questions, and others like them, deserve greater attention in the interest of more effectively structuring aural skills curricula.

One final broad topic that merits exploration is the psychological impacts of singing on solmization. The psychology of singing, and, by extension, solmization,
contains diverse branches of study. Although that diversity limited the psychological approach of this thesis, the research and thought conducted during its creation did open potentially significant questions for future interested parties:

1) What impact does self-confidence have on the success of those using solmization systems, particularly when singing for others? How can aural skills classes be formatted to better assist those who struggle due to a lack of self-assurance, rather than some deficiency in their musical ability?

2) How many technical problems do not result from insufficient vocal education, but a lack of vocal familiarity? That is, how many potential issues discussed in this thesis could be mitigated by merely introducing vocal warm-ups into the classroom?

3) How strongly are the aspects of vocal comfort and self-confidence linked to each other? How strongly are they linked to success in the aural skills curriculum?

4) What role does “muscle memory” and perceptual memory play in the movable-Do system? To what extent does a solfège syllable’s vowel reinforce that syllable’s function? How would a student fluent in movable-Do perform if asked to sing or identify intervals on random vowels compared to singing or identify intervals on vowels that coincided with solfège syllables of equal distance?
Reference List


